

**UNIVERSIDADE ESTADUAL DE CAMPINAS**  
**INSTITUTO DE QUÍMICA**  
**DEPARTAMENTO DE QUÍMICA ANALÍTICA**  
**LAQQA – LABORATÓRIO DE QUIMIOMETRIA EM QUÍMICA ANALÍTICA**



**UNICAMP**

“Utilização do sinal analítico líquido para validação de modelos de calibração multivariada através do cálculo de figuras de mérito e de cartas de controle”

DISSERTAÇÃO DE MESTRADO

AUTOR: WÉRICSON FORTUNATO DE CARVALHO ROCHA  
ORIENTADOR: Prof. Dr. RONEI JESUS POPPI

CAMPINAS – SP, JULHO DE 2007

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO  
INSTITUTO DE QUÍMICA DA UNICAMP

R582u Rocha, Wéricson Fortunato de Carvalho.  
Utilização do sinal analítico líquido para  
validação de modelos de calibração multivariada  
através do cálculo de figuras de mérito e de cartas  
de controle / Wéricson Fortunato de Carvalho  
Rocha. -- Campinas, SP: [s.n], 2007.

Orientador: Ronei Jesus Poppi.

Dissertação - Universidade Estadual de  
Campinas,  
Instituto de Química.

1. NAS. 2. Nimesulida. 3. Figuras de mérito.  
4. Carta de controle multivariada. I. Poppi, Ronei  
Jesus. II. Universidade Estadual de  
Campinas. Instituto de Química. III. Título.

**Título em inglês:** Utilization of net analyte signal for validation of models of multivariate calibration by using figure of merit and chart control

**Palavras-chaves em inglês:** NAS (Net Analyte Signal), Nimesulide, Figure of merit, Statistical quality control multivariate

**Área de concentração:** Química Analítica

**Titulação:** Mestre em Química na Área de Química Analítica

**Banca examinadora:** Prof. Dr. Ronei Jesus Poppi (orientador), Prof. Dr. Marco Flôres Ferrão (UNISC-DQF) e Jarbas José Rodrigues Rohwedder (UNICAMP-IQ)

**Data de defesa:** 18/07/2007

Dedico este trabalho aos meus queridos  
pais, Adonias Maciel Rocha  
e Sônia Maria de Carvalho Rocha.  
Obrigado pelo amor, pelo  
apoio e pela torcida em toda minha trajetória.

“Se quisermos progredir, não devemos repetir a história, mas fazer uma história nova.”

**(Hannibal)**

“Ou encontramos uma forma, ou a inventamos.”

**(Hannibal)**

“Quando a gente pensa que sabe todas as respostas, vem a vida e muda todas as perguntas.”

**(Emily Paul)**

"Não fiz o melhor, mas fiz tudo para que o melhor fosse feito. Não sou o que deveria ser, mas não sou o que era antes."

**(Martin Luther King)**

“É claro que meus filhos terão computadores, mas antes terão livros.”

**(Bill Gates)**

*"Este livro é apenas uma pequena parte; tenho material para pelo menos mais dez livros. Depende de você para que os outros dez sejam publicados - é uma responsabilidade que lhe passo. Durma bem."*

**(Leon Eliachar)**

## **AGRADECIMENTOS**

- Ao professor Dr. Ronei Jesus Poppi pela oportunidade de realização desse trabalho, pela orientação, confiança, paciência, convivência e amizade, principalmente nas horas ruins que não foram poucas.
- Aos colegas do grupo LAQQA, IQ-Unicamp e da Moradia Estudantil da Unicamp: Bruno Bernado Galindo Lopes, Jeferson Gomes Casseiro, André Luis Camilo, José Cicero Calheiros, Aldo Gomes Pereira, André Luis Rosa, André Murari, Deleon, Renato Andrade, Erick (Péricles), Heloisa Biscolla, Jocélia, João, Clebér Maia, Ebénezer, Marcelo Cordeiro, Diorginis, Kássio Lima, Marcello Garcia Trevisan, Paulo Henrique Março, Uanderso Resende, Julião, Mored, Fabiano, Ferrão, Paulo da Costa Filho (Pinto), Renato, Marcelo(P-2), Maurício(Professor de inglês), Marco Franco, Danilo, Alexandre Kisner e dentre outros nos quais seria impossível descrever aqui.
- Ao IIPF-Instituto Internacional de Pesquisas farmacêuticas por ter concedido a matéria-prima necessária na realização de todo o trabalho, por ter fornecido as informações e métodos de análise utilizado pela indústria e pelo apoio.
- Ao Dr. José Antônio Martins pela ajuda através do fornecimento da matéria-prima utilizada para a realização desse trabalho.
  - À minha família, meus pais Adonias Maciel Rocha e Sônia Maria de Carvalho Rocha, meu irmão Adson Batista de carvalho Rocha.
  - À minha tia Vanda que pagou a inscrição do meu vestibular da Unicamp em 2001.
  - À minha madrinha, Madalena, e ao meu padrinho, Silvério, pelo apoio quando precisei.
  - À minha tia Lindalva e ao seu marido Humberto pelas conversas pelo telefone na qual eu ligava à cobrar.
  - À minha avó, Dalva Maciel, a qual eu amo muito.

- Ao meu tio, Rondon Maciel Rocha, pelos R\$ 550,00 emprestados em 2001 para eu vir para Campinas o qual eu ainda não paguei.
- À família da minha namorada, Michele de Souza e Silva, pela amizade e companheirismo durante esses 2 anos de mestrado que tornaram os meus fins de semanas mais alegres.
- À minha namorada, Michele de Souza e Silva, pela amizade, carinho, incentivo e “puxões de orelha” nos momentos certos, principalmente, quando eu não fazia as tarefas de inglês.
- À vizinha da minha namorada Carina e seu marido, Alexandre, pela amizade.
- À CAPES pelo financiamento desse projeto.
- À Unicamp e ao Estado de São Paulo por fornecer toda estrutura física, tecnológica, e humana para realização deste trabalho.
- Aos funcionários, colegas, familiares e amigos do Instituto de Química da Unicamp aos quais me esqueci e a todas as pessoas que, de uma forma ou outra contribuíram para a realização deste trabalho e para esta nova etapa de minha vida concluída, os meus sinceros agradecimentos.

# Currículo

---

## 1. Dados pessoais

---

Wérickson Fortunato de Carvalho Rocha  
Naturalidade: Pompéu-MG

## 2. Formação Acadêmica (Graduação e Pós-Graduação)

---

- Curso Superior em Química (UNICAMP)
- Mestrado em Química Analítica sub-área: Quimiometria (UNICAMP)

## 3. Produção científica

---

### 3.1. Iniciação científica

Orientador: Dr. Ronei Jesus Poppi  
Instituição: Unicamp  
Nome do projeto:

Parâmetros de Qualidade de Lubrificantes e Óleo de Oliva através de Espectroscopia Vibracional, Calibração Multivariada e Seleção de Variáveis

### 3.2. Resumo do trabalho científico apresentado em congresso

- Modelos de regressão empregando iPLS e DRIFTS para determinação de Carbamazepina em medicamento.

ROCHA, Wérickson Fortunato de Carvalho; POPPI, Ronei Jesus; FERRÃO, Marco Flôres.

29<sup>o</sup> Reunião Anual da Sociedade Brasileira de Química, Águas de Lindóia, 2006.

Publicado no livro de resumos v.29, p.QA-031.

- Figuras de mérito em calibração multivariada na determinação de carbamazepina.

ROCHA, Wérickson Fortunato de Carvalho; Ferrão, Marco ; Franco, Marco Fernando ; Poppi, Ronei Jesus

3<sup>o</sup> Seminário Internacional Ciência e Tecnologia na América Latina - A Universidade como Promotora do Desenvolvimento Sustentável.

Evento ocorrido na cidade de Campinas na Universidade Estadual de Campinas entre os dias 14 e 15 de Setembro de 2006.

- Determination of Nimesulide Tablets employing IPLS and DRIFTS

ROCHA, Wérickson Fortunato de Carvalho; Rosa, André Luis; Martins, José Antônio; Poppi, Ronei Jesus

10th International Conference on Chemometrics in Analytical Chemistry.

- Livros didáticos A matéria: Estados Físicos, Estrutura e Comportamento  
R. JUNIOR, L.; SANTOS, M. M. ; ROCHA, Wérickson Fortunato de Carvalho  
V Encontro de Estudantes de Graduação dos Cursos de Formação de Professores da Unicamp. 2004  
4 de Dezembro de 2004.Evento ocorrido na Faculdade de Educação da Unicamp.
- Validação de modelos de calibração multivariada através do sinal analítico líquido e do cálculo de ANOVA  
30º Reunião Anual da Sociedade Brasileira de Química, Aguás de Lindóia, 2007.  
ROCHA, Wérickson Fortunato de Carvalho ; Franco,Marcos Fernando ; Poppi,Ronei Jesus
- Modelos de regressão empregando Bi-PLS e espectroscopia no infravermelho para determinação de Nimesulida em medicamento  
30º Reunião Anual da Sociedade Brasileira de Química, Aguás de Lindóia, 2007.  
ROCHA, Werickson Fortunato de Carvalho ; Franco,Marcos Fernando ; Poppi,Ronei Jesus
- Validação de modelo de calibração multivariada através do cálculo de figuras de mérito e de espectroscopia no infravermelho próximo.  
ROCHA, Werickson Fortunato de Carvalho; Rosa, André Luis; Martins, José Antônio; Poppi, Ronei Jesus  
30º Reunião Anual da Sociedade Brasileira de Química, Águas de Lindóia, 2007.
- Monitoramento no controle de qualidade de fármacos através de cartas de controle multivariada.  
ROCHA, Werickson Fortunato de Carvalho; Rosa,André Luis ; Martins,José Antônio ; Poppi,Ronei Jesus  
30º Reunião Anual da Sociedade Brasileira de Química, Águas de Lindóia, 2007.
- Validation of models of multivariate calibration: an application in the determination of nimesulide in tablets by near infrared spectroscopy  
ROCHA, Werickson Fortunato de Carvalho; Rosa, André Luis; Martins,José Antônio; Poppi, Ronei Jesus  
The 13th International Conference on Near Infrared Spectroscopy (13th ICNIRS) In Umeå-Vasa, Sweden & Finland 15-21 June 2007
- Monitoramento de fármacos através de espectroscopia no infravermelho próximo e de cartas de controle multivariadas  
ROCHA, Werickson Fortunato de Carvalho; Poppi,Ronei Jesus  
Escola Latino-americana sobre Matéria condensada.Corrientes, Argentina.

### **3.3. Publicação**

Título: Modelos de regressão empregando iPLS e DRIFTS para determinação de Carbamazepina em medicamento.

Autores: W. F.C. Rocha, M. F. Ferrão e R.J. Poppi

Palavras-chave: iPLS, Carbamazepina, DRIFTS, Quimiometria

Revista: TECNOLÓGICA vol. 10-n. 1-Jan./jun. 2006 ISSN. 1415-6229

UNISC- Universidade de Santa Cruz do Sul

## **4. Outros**

---

### **4.4 . Monitoria**

Monitor de Química e Física do colégio Anglo

Período: 1º Semestre de 2000

Monitor das disciplinas F-128 Física Geral 1\ F-129 Física experimental(Unicamp)

Período: 1º e 2º semestre de 2002

### **4.5. Estágios**

- Estágio na área de ensino de química CEES(Centro Estadual de Educação Supletiva) Paula Decourt Unicamp

Período: 1º Semestre de 2004

- Participação do programa PED(Programa de estágio docente) na disciplina QA 313(Química para Engenharia) na Unicamp

Período: 1º Semestre de 2006

### **4.3 Cursos científicos**

- Disciplina MEQVT(Metodologia do Ensino de Química via Telemática) oferecido pelo LAPEQ(Laboratório de Pesquisa em Ensino de Química e Telemática Educacional)-USP

Carga horária: 60 horas

Responsável: Prof. Dr. Marcelo Giordan

Período: 2º Semestre de 2005

- Basic Chemometrics

Carga horária: 10 horas

Responsável: Barry Wise (Eigenvector Research Inc. USA)

Período: 10/09/2006

Local: 10th International Conference on Chemometrics in Analytical Chemistry.

- Uncertainty estimation in multivariate and multiway calibration  
Carga horária: 10 horas  
Responsável: Klaas Faber (NL)  
Período: 11/09/2006  
Local: 10th International Conference on Chemometrics in Analytical Chemistry.
- QP-268 (Planejamento e Otimização de Experimentos)  
Carga horária: 60 horas  
Responsável: Prof.Dr. Roy Edwards Bruns  
Período: 2º Semestre de 2006  
Local:Universidade Estadual de Campinas
- QP-434 (Tópicos Especiais em Físico-Química II “Controle Multivariado de Processos Químicos”)  
Carga horária: 60 horas  
Responsável: Prof(a)Dr(a). Márcia Miguel Castro Ferreira  
Período: 1º Semestre de 2007  
Local:Universidade Estadual de Campinas
- Curso de Nanotecnologia realizado pelos palestrantes: Prof. Dr.Fernando Galembeck do Departamento de Físico-Química, e Prof. Dr. Oswaldo Luiz Alves, Laboratório de Química do Estado Sólido, Departamento de Química Inorgânica (IQ-UNICAMP).
- Curso de Química Computacional realizado pelos palestrante: Prof. Dr. Rogério Custodio, Prof. Dr. Nelson Henrique Morgon e Dr. Leandro Martinez, Departamento de Físico-Química (IQ – UNICAMP)
- Curso de Química Forense realizado pelo Palestrante: Prof. Dr. Valter Stefani, Laboratório de novos compostos orgânicos, Departamento de Química Orgânica (IQ – UFRGS).

### **4.3. Experiência profissional na área**

- Professor de Química Cursinho do Sindicato dos Trabalhadores de Campinas  
Período: 1º Semestre de 2003- 2º Semestre de 2004
- Professor de Química/Física do pré-vestibular da Moradia Estudantil da UNICAMP  
Período: 1º Semestre de 2002- 1º Semestre de 2005

- Professor de Física e Química ERAB (Escolas Reunidas Armas Brasileiras)

Período: 1º Semestre de 2005

- Autor da apostila de Química do Cursinho da Moradia Estudantil da Unicamp.

## Resumo

**Autor:** Wéricson Fortunato de Carvalho Rocha

**Orientador:** Prof. Dr. Ronei Jesus Poppi

Este trabalho teve como objetivo realizar a validação de modelos de calibração multivariada através do cálculo de figuras de mérito e de cartas de controle multivariadas do medicamento Nimesulida por meio do infravermelho próximo.

Foram preparadas 69 amostras sintéticas contendo o princípio ativo (Nimesulida) na faixa 10,38-39,47% (m/m) em excipiente (lactose, povidona-KV29-32, celulose 200, lauril sulfato de sódio, croscarmelose sódica e estereato de magnésio). Destas, 49 amostras foram utilizadas para a calibração e 20 para a validação as quais foram separadas através do algoritmo de Kennard-Stone. O tratamento utilizado nos espectros foi a correção do espalhamento multiplicativo. Na sequência, foram obtidos os seguintes valores para as figuras de mérito: limite de detecção (0,61), limite de quantificação (2,03), exatidão (RMESC 0,66 %(m/m), RMESCV 0,92 %(m/m), RMESP 1,05 %(m/m) de Nimesulida), seletividade média (0,0056), sensibilidade (0,0036), inverso da sensibilidade analítica (0,20 %(m/m)<sup>-1</sup> de Nimesulida) e razão sinal ruído (181,11).

Na segunda parte do trabalho foram utilizadas 113 amostras sintéticas para construção das cartas de controle multivariadas. Foram desenvolvidas três cartas de controle e calculado os limites de controle para cada carta. Os valores encontrados foram: - carta NAS (Superior =  $6,54 \times 10^{-4}$ ; Inferior =  $5,25 \times 10^{-4}$ ), carta interferente (7,45) e carta resíduo ( $2,38 \times 10^{-6}$ ). Através dessas cartas de controle multivariadas foi possível identificar as amostras que estavam dentro e fora de controle. Obteve-se 64 amostras fora de controle e 20 amostras dentro de controle de acordo com o planejamento experimental realizado. Sendo assim, foi possível identificar, de forma qualitativa, as amostras de Nimesulida que estavam dentro e fora de controle.

Portanto, a dissertação desenvolvida sugere um novo método analítico que pode ser usado para controle de qualidade em fármacos e para validação de modelos de calibração multivariada, pois os resultados obtidos, indicam que o modelo desenvolvido pode ser utilizado na indústria farmacêutica como uma alternativa ao método-padrão.

## Abstract

Author: Wéricson Fortunato de Carvalho Rocha

Adviser: Prof. Dr. Ronei Jesus Poppi

This study was mainly intended to elaborate the validation of multivariate calibration models based on the determination of figures of merit and multivariate control charts constructed with data from Nimesulide tablets upon near-infrared spectroscopy use.

A total of 69 synthetic samples were prepared containing the active principle (Nimesulide) in the range of 10,38-39, 47% (m/m) in excipients (Cellulose, Sodium Lauryl Sulphate, Magnesium Stearate, Carmellose Sodium, Povidone and Lactose). It was used 49 samples for the calibration and 20 for the validation that were separated by the Kennard-Stone algorithm. It was used the multiplicative scatter correction on the spectra set. Then, the following values for figures of merit were calculated: limit of detection (0,61) , limit quantification (2,03) , accuracy (RMESC 0,66 %(m/m), RMESCV 0,92%(m/m), RMESP 1,05%(m/m) of Nimesulide), mean selectivity (0,0056), sensitivity (0,0036), inverse analytical sensitivity (0,20 %(m/m)<sup>-1</sup> of the Nimesulide) and signal-to-noise ration (181, 11).

In the second application, 113 synthetic samples were used for the construction of multivariate control charts. Three control charts were designed and the control limits for each chart were calculated. The values found were: - chart NAS (upper =  $6,54 \times 10^{-4}$ ; below =  $5,25 \times 10^{-4}$ ), chart interferent (7,45) and chart residual ( $2,38 \times 10^{-6}$ ). From these charts it was possible to identify the samples that were in and out of control. It was obtained 64 samples out of control and 20 samples in control, according to the experimental design.

Therefore, it was possible to identify, in a qualitative way, the Nimesulide samples that were in control and the ones that were out of control.

Hence it follows that, this research project suggests a new analytical method that can be used for quality control in drugs and also for the validation of multivariate calibration models. Thus, the results obtained indicate that this model can be used in pharmaceutical industries as an alternative to the standard procedure.

## ÍNDICE

<b>Índice de figuras</b> .....	XXV
<b>Índice de tabelas</b> .....	XXIX
<b>Capítulo 1- Introdução</b> .....	1
1.1 Princípios da Espectroscopia no Infravermelho.....	3
1.2 Espectroscopia no Infravermelho por Reflectância Difusa.....	7
1.3 Objetivos.....	8
<b>Capítulo 2- Métodos Quimiométricos</b> .....	9
2.1 Organização dos Dados.....	11
2.2 Tratamento.....	12
2.2.1 Primeira Derivada.....	12
2.2.2 Correção do Espalhamento Multiplicativo.....	13
2.3 Análise de Componentes Principais – PCA.....	15
2.4 Regressão por Mínimos Quadrados Parciais.....	17
2.5 Validação Cruzada.....	19
2.6 Detecção de Amostras Anômalas em Calibração Multivariada.....	20
2.7 Previsão.....	22
2.8 Sinal Analítico Líquido.....	23
<b>Capítulo 3- Validação de Modelos de Calibração Multivariada Através do Cálculo de Figuras de Mérito</b> .....	25
3. Validação e Figuras de Mérito.....	27
3.1 Validação.....	27
3.2 Figuras de Mérito.....	28
3.2.1 Exatidão.....	28
3.2.2 Ajuste do Modelo.....	29

3.2.3 Linearidade.....	29
3.2.4 Sensibilidade.....	30
3.2.5 Sensibilidade Analítica.....	30
3.2.6 Seletividade.....	31
3.2.7 Razão Sinal/Ruído.....	31
3.2.8 Testes para erros Sistemáticos (“bias”).....	31
3.2.9 Limite de Detecção e Quantificação.....	32

**Capítulo 4-Cartas de Controle Multivariadas.....33**

4.1 Cartas de Controle Multivariadas.....	35
4.2 Carta de Controle Multivariada Baseado na Divisão do Espectro em NAS, Interferente e Resíduos.....	35
4.3 Construção de Cartas de Controle Baseada no Sinal Analítico Líquido.....	36
4.4 Cálculo dos Limites Estatísticos para as cartas NAS, Interferente e Resíduos.....	39
4.5 Carta de Controle NAS.....	39
4.6 Carta de Controle Interferente.....	40
4.7 Carta de Controle dos Resíduos.....	42
4.8 Validação das Cartas de Controle .....	44
4.8.1 Carta de Controle NAS Para Novas Amostras.....	44
4.8.2 Carta de controle Interferentes Para Novas Amostras.....	44
4.8.3 Carta de Controle Resíduos Para Novas Amostras.....	45

<b>Capítulo 5- Determinação de Nimesulida em Medicamentos Através do Cálculo de Figuras de Mérito e de Espectroscopia no Infravermelho Próximo.....</b>	<b>47</b>
5-Parte Experimental.....	49
5-1 Espectroscopia no Infravermelho Próximo.....	51
5-2 Modelos de Regressão PLS.....	51
<b>Capítulo 6-Controle de Qualidade da Nimesulida Através de Cartas de Controle Multivariada e de Espectroscopia no Infravermelho Próximo.....</b>	<b>61</b>
6- Parte Experimental.....	63
6-1 Espectros no Infravermelho Próximo.....	65
6-2 Construção das Cartas de Controle Multivariadas.....	66
6-3 Validação das Cartas de Controle Multivariadas.....	70
<b>Capítulo 7.Conclusões.....</b>	<b>77</b>
<b>Capítulo 8.Referências Bibliográficas.....</b>	<b>81</b>

## Índice de Figuras

- Figura 1.** Diferentes modos vibracionais de uma molécula de água.
- Figura 2.** Fórmula estrutural da Nimesulida.
- Figura 3.** Diagrama esquemático da configuração do acessório de reflectância difusa equipado com esfera de integração.
- Figura 4.** Organização dos dados em uma matriz **X** a partir dos espectros no infravermelho para construção de um modelo de calibração multivariada.
- Figura 5.** A) Espectros no infravermelho sem nenhum pré-processamento;  
B) Espectros corrigidos com a 1<sup>o</sup> derivada.
- Figura 6.** A) Espectros no infravermelho sem nenhum pré-processamento;  
B) Espectros corrigidos com o MSC.
- Figura 7.** Decomposição da matriz **X** nas matriz de escores (**T**), pesos (**P**) e resíduos (**E**).
- Figura 8.** Representação de duas componentes principais.
- Figura 9.** Representação geométrica do PLS com uma variável latente modelando os dados de **X** (variáveis independentes) e **Y** (variáveis dependentes). Os dados **X** formados por três variáveis e dados de **Y** por duas variáveis.
- Figura 10.** Gráfico do RMSECV em função da dimensionalidade do modelo de regressão.
- Figura 11.** Representação geométrica da decomposição do sinal analítico.
- Figura 12.** Representação da divisão do espectro (**r**) em quatro diferentes contribuições: NAS ( $r_k^*$ ), interferente ( $r_{INT}$ ), resíduos ( $r_{RES}$ ) e  $r^\perp$  (NAS + resíduos).
- Figura 13.** Representação da Carta de Controle NAS.
- Figura 14.** Representação da Carta de Controle Interferente.
- Figura 15.** Representação da Carta de Controle Resíduo.
- Figura 16.** Representação da Carta de Controle NAS para novas amostras.
- Figura 17.** Representação da Carta de Controle Interferente para novas amostras.
- Figura 18.** Representação da Carta de Controle resíduo para novas amostras.
- Figura 19.** Diagrama ternário para a mistura física do medicamento Nimesulida.

**Figura 20.** Metodologia utilizada para construção e validação do modelo para a mistura sintética do medicamento Nimesulida.

**Figura 21.** A) Espectros NIR do medicamento Nimesulida sem tratamento;  
B) Espectros NIR do medicamento Nimesulida com MSC.

**Figura 22.** Gráfico dos resultados obtidos pelo método de validação cruzada utilizado para a escolha do melhor número de variáveis latentes.

**Figura 23.** Gráfico dos valores de referência contra os valores previsto em (%) m/m estimado pelo o modelo PLS. (●) Calibração, (▲) Validação e (—) Reta ajustada

**Figura 24.** Representação do modelo na sua forma pseudo-univariada.

(\* Calibração, (●) Validação e (—) Reta ajustada.

**Figura 25.** Resíduos entre os valores reais e previstos pelo modelo PLS.

(●) Calibração e (▲) Validação.

**Figura 26.** Resíduos entre os valores reais e os previstos pelo modelo pseudo-univariado. A) (X) Calibração e B) (\*)Validação.

**Figura 27.** Escore da 1ª variável latente contra as concentrações da calibração.

**Figura 28.** Gráfico de resíduos contra as concentrações de referência.

**Figura 29.** Diagrama ternário para a escolha das amostras de calibração e validação.

**Figura 30.** Espectros da mistura (-) e dos constituintes puros (-) sem tratamento e com a primeira derivada: a) Celulose, b) Estereato de Magnésio, c) Povidona e d) Nimesulida.

**Figura 31.** Espectros da mistura (-) e dos constituintes puros (-) sem tratamento e com a primeira derivada: a) Lactose, b) Amido e c) NaHCO<sub>3</sub>.

**Figura 32.** A) Espectros NIR do medicamento Nimesulida sem tratamento;  
B) Espectros NIR do medicamento Nimesulida com 1ª derivada.

**Figura 33.** Divisão espectral em NAS, Interferente e Resíduos.

**Figura 34.** Valor do NAS vs Nimesulida (% (m/m)) para as amostras da figura 19.

**Figura 35.** Cartas de controle a) NAS; b) Interferentes e c) Resíduos.

**Figura 36.** Cartas de controle com novas amostras a) NAS; b) Interferentes e c) Resíduos.

**Figura 37.** Vetores interferentes para as amostras sob controle (-) e para amostras 1, 11, 21, 22, 29, 36 e 42 que estão fora de controle(-).

**Figura 38.** Vetores resíduos para as amostras de calibração (-) e das novas amostras 1, 4, 7, 19, 21, 29, 32, 36, 40, 44 e 50 (-).

**Figura 39.** Vetores resíduos para região 4250-4650( $\text{cm}^{-1}$ ) da figura 38.

**Figura 40.** Cartas de controle: amostras sob controle (•) e amostras fora de controle (\*). a) NAS; b) Interferentes e c) Resíduos.

**Figura 41.** Vetores interferentes.

**Figura 42.** Vetores interferentes para região 4000-6000 ( $\text{cm}^{-1}$ ) da figura 41.

**Figura 43.** Vetores resíduos.

## Índice de Tabelas

**Tabela 1.** Resultados dos parâmetros de figuras de mérito para o modelo PLS.

**Tabela 2.** Descrição de todas amostras usadas para construção e validação das cartas de controle.

# **CAPÍTULO 1**

## **INTRODUÇÃO**

## 1.1 Princípios da espectroscopia no infravermelho

Determinações espectroscópicas são rápidas e não-destrutivas, têm ganho cada vez mais popularidade como método analítico em indústrias farmacêuticas, com destaque para o uso da espectroscopia no infravermelho próximo (NIR) para análise em matérias-primas<sup>1,2</sup>, produtos intermediários<sup>3,4,5</sup> e finais<sup>6,7</sup>. O NIR corresponde a região do espectro eletromagnético que vai de 12800 a 4000  $\text{cm}^{-1}$ . Durante o desenvolvimento do NIR, as medidas inicialmente eram realizadas com fotômetros e espectrofotômetros dispersivos, baseados em arranjos ópticos com filtros e redes de difração, respectivamente, com configurações semelhantes as utilizadas em espectrofotômetros que trabalham na região do ultravioleta e visível (UV-VIS), ou mesmo em equipamentos que possuíam fonte e detectores que se estendiam até essa região. Atualmente, devido ao reconhecimento do potencial da aplicação do NIR, principalmente para análises quantitativas, equipamentos modernos, em sua maioria interferômetros com transformada de Fourier, vêm sendo desenvolvidos especificamente para análises nesta região e já se encontram disponíveis com acessórios para análise de amostras sólidas, líquidas e gasosas, assim como os que possibilitam o desenvolvimento de aplicações em áreas específicas, como por exemplo, nas áreas farmacêutica e clínica<sup>8</sup>.

Diferentes ligações moleculares absorvem em número de onda específico na região do infravermelho, que possui uma faixa espectral conhecida como região de impressão digital. Os átomos em uma molécula nunca estão imóveis. Se, em um sistema, há " $N$ " átomos livres para se movimentarem nas três dimensões, o sistema terá " $3N$ " graus de liberdade. Se, no entanto, esses átomos estiverem ligados entre si, formando uma molécula, continuarão ainda existindo  $3N$  graus de liberdade, sendo três graus para a translação do centro de massa da molécula e, para uma molécula não linear, três graus para a rotação da mesma em torno dos três eixos, restando, assim,  $3N-6$  graus de liberdade para as vibrações. Para moléculas lineares, como não há rotação em torno do eixo internuclear, restam  $3N-5$  graus de liberdade para as vibrações. Esses graus de liberdade correspondem aos diferentes modos normais de vibração de uma molécula. Um modo normal de vibração é aquele

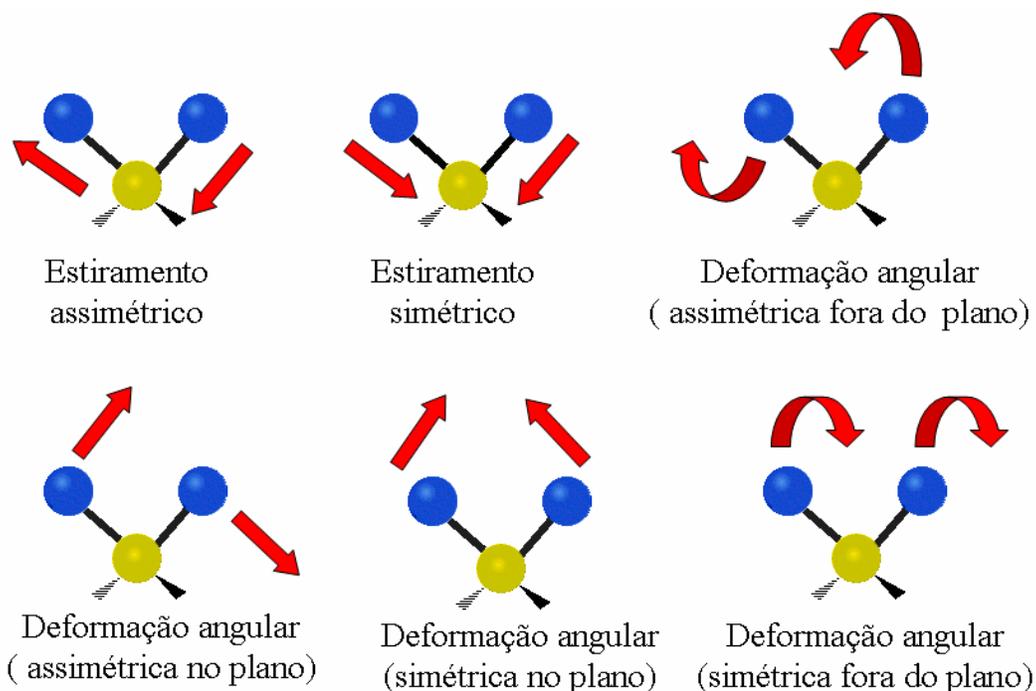
em que cada núcleo realiza uma oscilação harmônica simples em torno de sua posição de equilíbrio, todos os núcleos se movem com a mesma frequência e em fase e o centro de gravidade da molécula permanece inalterado<sup>9</sup>.

Na prática, nem sempre o número de modos normais de vibração corresponde ao número de bandas observadas no espectro. Isso ocorre devido à existência de vibrações de mesma energia, apresentando a mesma frequência e, conseqüentemente, a mesma posição no espectro. Além das frequências associadas às vibrações normais, frequências adicionais podem aparecer no espectro, resultantes dos seguintes fatores:

- Sobretons - bandas com valores de frequência correspondentes a múltiplos inteiros daqueles das vibrações normais. Por exemplo, seja  $\nu$  o valor da frequência de um dado modo normal de vibração, os sobretons vão aparecer com valores aproximados de  $2\nu$ ,  $3\nu$ , etc.

- Bandas de combinação – são combinações lineares das frequências normais ou múltiplos inteiros destas. Por exemplo, sejam  $\nu_a$  e  $\nu_b$  valores de frequência de modos normais de vibração, podem ocorrer as bandas de combinação  $(\nu_a + \nu_b)$ ,  $(\nu_a - \nu_b)$ ,  $(\nu_a + 2\nu_b)$ , etc.

As atividades dos sobretons e das bandas de combinação podem ser deduzidas pela teoria de grupos. Em todos os casos, as intensidades dessas bandas serão menores quando comparadas com as dos modos normais. As vibrações moleculares podem ser classificadas em deformação axial (ou estiramento) e deformação angular e podem ser simétricas ou assimétricas. As vibrações angulares podem ainda ser classificadas como no plano ou fora do plano. Os diferentes tipos de vibração para uma molécula de água são mostrados na figura 1.



**Figura 1:** Diferentes modos vibracionais de uma molécula de água.

Considerando a energia da radiação eletromagnética na faixa do NIR, essa possui energia suficiente para excitar sobretons e combinações de vibrações moleculares a altos níveis de energia. Logo é usada na medida quantitativa de grupos funcionais orgânicos, especialmente O-H, N-H, e C=O. Porém, uma desvantagem da técnica, é provavelmente, a baixa sensibilidade a constituintes menores em uma determinada amostra<sup>9</sup>.

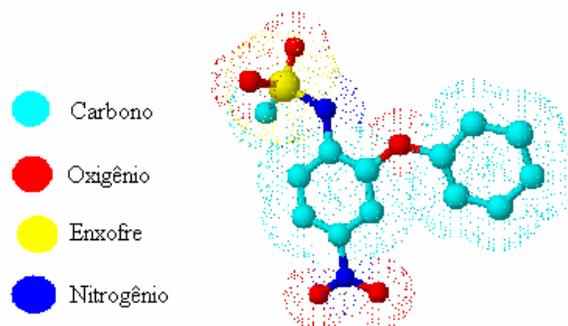
Uma importante aplicação do infravermelho é a análise quantitativa de misturas de compostos. Como o valor de absorvância de uma banda de absorção é proporcional a concentração do componente, a quantidade de um composto presente em uma amostra pode ser determinada através de uma curva de calibração (absorvância *versus* concentração) construída a partir de amostras com concentrações conhecidas do composto em questão. No entanto, quanto mais complexa é a amostra, ou seja, quanto maior o número de espécies presentes na matriz, mais difícil se torna a construção de uma calibração univariada confiável, sendo necessário lançar mão de cálculos estatísticos mais complexos, que permitam a utilização de vários comprimentos

de onda para determinação de uma única propriedade, obtendo-se, assim, uma calibração multivariada.

A análise do princípio ativo, contido em tabletes individuais de medicamentos, pode ser realizada pelo NIR no modo reflectância ou transmitância<sup>10</sup>. Para quantificar uma determinada quantidade de princípio ativo, modelos quantitativos de calibração construídos através do NIR devem ser desenvolvidos, a priori, para quantificar amostras futuras. Em muitos casos, um modelo quantitativo é desenvolvido com vários níveis da concentração do princípio ativo em um determinado medicamento, produzindo várias amostras. Essas são divididas em amostras de calibração e de validação com o intuito de construir o modelo e validá-lo, respectivamente.

Em análises de fármacos, o princípio ativo é usualmente a propriedade de interesse mais importante. Entretanto, já que amostras inteiras são medidas instantaneamente no NIR, o espectro também contém informação de outros compostos denominados excipientes, bem como propriedades físicas do medicamento. Dessa forma, a homogeneidade dos excipientes, assim como a água contida pode ser identificada. Especialmente, quando o princípio ativo é a menor parte do produto, isto é, medicamentos que contêm pequena quantidade de princípio ativo, os excipientes podem estar controlando importantes propriedades na qualidade como dissolução e dureza<sup>11</sup>.

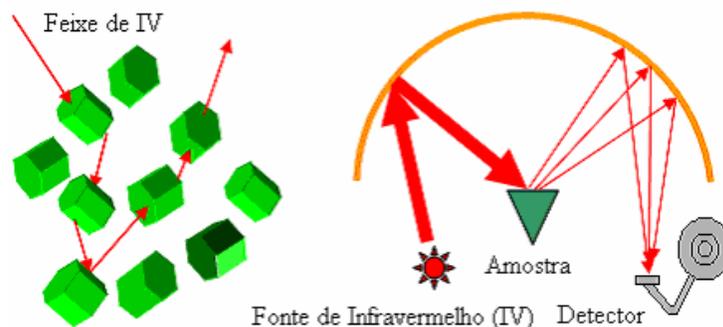
Nimesulida, cuja fórmula estrutural está apresentada na figura 2, é um agente antiinflamatório não-esteroidal, com atividade analgésica e antipirética cuja fórmula molecular é  $C_{13}H_{12}N_2O_5S$ .<sup>12,13</sup>. Sendo assim, é usada no tratamento de rinite alérgica e artrite.



**Figura 2:** Fórmula estrutural da Nimesulida.

## 1.2 Espectroscopia no infravermelho por Reflectância Difusa

Na espectroscopia de reflectância no infravermelho próximo, a amostra sólida finamente moída é irradiada com uma ou mais bandas estreitas de radiação com comprimento de onda na faixa do NIR ( $12800 - 4000 \text{ cm}^{-1}$ ). Ocorre reflectância difusa, na qual a radiação penetra na camada superficial das partículas, excita modos vibracionais da molécula do analito e é, então, espalhada em todas as direções. Um espectro de reflectância que depende da composição da amostra é, assim, produzido. Instrumentos para medidas de reflectância difusa estão disponíveis, podendo ser desde fotômetros de filtro até equipamentos baseados na transformada de fourier. Normalmente, as medidas de reflectância difusa são feitas em dois ou mais comprimentos de onda para cada espécie de analito que está sendo determinada<sup>14</sup>. A figura 3 mostra um esquema de uma medida por reflectância difusa utilizando esfera de integração.



**Figura 3:** Diagrama esquemático da configuração do acessório de reflectância difusa equipado com esfera de integração.

As paredes internas da esfera de integração da figura 3, que circunda a amostra, são recobertas com um material que reflete a radiação difusamente de modo quase perfeito, como sulfato de bário. A radiação refletida pela amostra chega aos detectores após múltiplas reflexões nas paredes da esfera. Isto faz com que a esfera seja iluminada uniformemente, de modo que a posição do detector não seja crítica, contudo, uma vez que a radiação difusa se espalha por toda a esfera, a intensidade do sinal é registrado pelo detector é menor que os obtidos por outros acessórios em que a radiação chega diretamente no detector.

### **1.3 Objetivos**

O objetivo deste trabalho consiste em construir e validar, utilizando espectroscopia no infravermelho próximo com medidas feitas por reflectância difusa:

- modelos de calibração multivariada, através do cálculo de figuras de mérito, para quantificação de Nimesulida em amostras de medicamentos genéricos, similares e de referência;
- cartas de controle multivariadas baseadas na divisão do espectro de infravermelho próximo em analito, resíduo e interferente para o fármaco Nimesulida.

# **CAPÍTULO 2**

## **MÉTODOS QUIMIOMÉTRICOS**

## Capítulo 2- Métodos Quimiométricos

O objetivo principal da química analítica quantitativa é estimar a concentração de uma ou mais espécies a partir dos valores de determinadas propriedades físico-químicas do sistema de interesse. Para isto, é necessário construir um modelo de calibração, ou seja, determinar a relação entre propriedades medidas e concentrações. A calibração multivariada é uma das mais bem sucedidas combinações de métodos estatísticos com dados químicos, tanto na química analítica, quanto na química teórica.

As notações de álgebra linear também se aplicam a quimiometria. Um vetor é uma matriz que tem somente uma linha ou uma coluna sendo chamado de vetor linha ou coluna respectivamente. Os vetores são designados por letras minúsculas em negrito (**y**). As matrizes são designadas por letras maiúsculas em negrito (**X**). Os elementos das matrizes, dos vetores, ou ainda constantes como coeficientes de regressão são representados por letras sem negrito ( $x$ )<sup>1</sup>.

### 2.1 Organização dos dados

Os dados em calibração multivariada podem ser dispostos em forma de matriz conforme a figura 4 no qual as linhas da matriz referem-se as amostras obtidas e as colunas referem-se às variáveis (absorbância de um composto em diferentes número de onda)<sup>2,3</sup>.

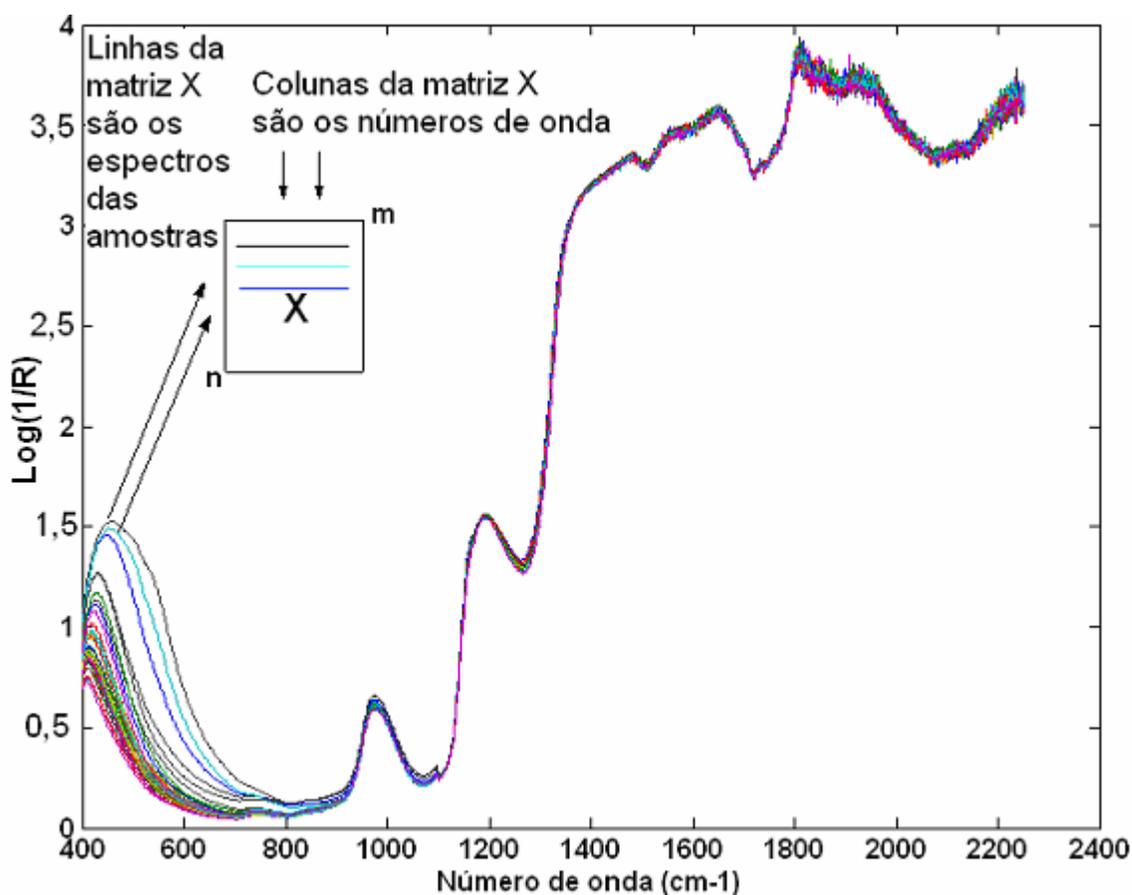


Figura 4: Organização dos dados em uma matriz  $X$  a partir dos espectros no infravermelho para construção de um modelo de calibração multivariada.

## 2.2 Tratamento

### 2.2.1- 1° Derivada

Variações na linha de base de um espectro produzem variações indesejáveis nas medidas de absorbância. Estes problemas de linha de base podem ser corrigidos tomando-se as derivadas do espectro. O método utilizado para essa tarefa foi a primeira derivada através do algoritmo Savistky-Golay<sup>4</sup>. A figura 5 abaixo demonstra espectros no infravermelho sem nenhum tratamento e com 1° derivada.

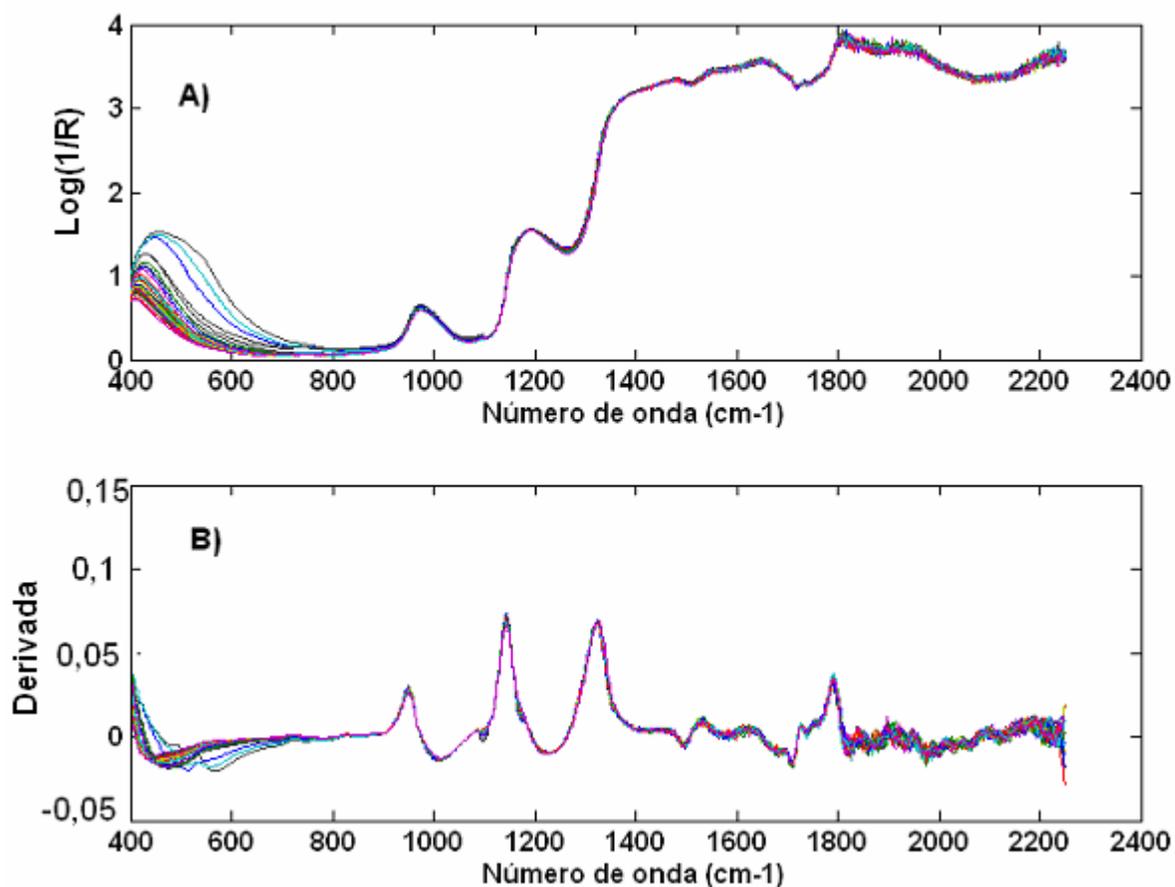


Figura 5: A) Espectros no infravermelho sem nenhum tratamento  
B) Espectros corrigidos com a 1<sup>o</sup> derivada.

### 2.2.2 Correção do Espalhamento Multiplicativo

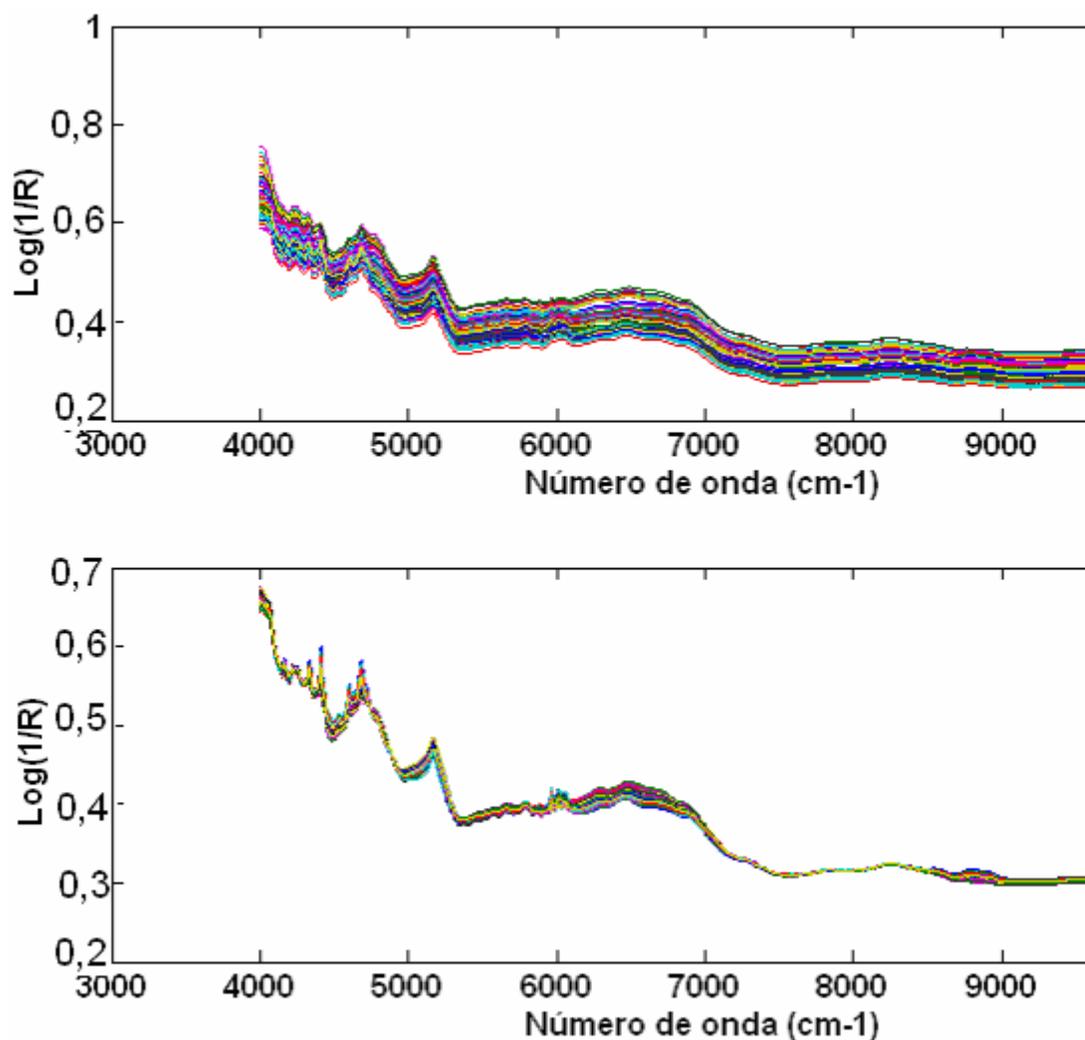
Uma outra alternativa para a correção da linha da base dos espectros, é a correção do espalhamento multiplicativo do inglês “Multiplicative scatter correction, MSC”<sup>5</sup>. Esta correção faz uma regressão de um espectro medido contra um espectro de referência (geralmente o espectro médio) e corrige o espectro medido usando uma inclinação. Para cada espectro individual  $x_i$ , os parâmetros  $a_i$  e  $b_i$  são estimados por regressão em  $\bar{x}$  (espectro de

referência/médio). Os parâmetros estimados  $a_i$  e  $b_i$  são utilizados para realizar a correção do espectro da seguinte maneira:

$$x_i = a_i + \bar{x}b_i \quad (\text{Equação 1})$$

$$x_{i,\text{corrigido}} = \frac{(x_i - a_i)}{b_i} \quad (\text{Equação 2})$$

A figura 6 mostra o efeito da correção com o MSC sobre espectros de infravermelho com diferentes linha-base.



**Figura 6:** A) Espectros no infravermelho sem nenhum tratamento  
B) Espectros corrigidos com o MSC.

## 2.3 Análise de Componentes Principais - PCA

Métodos baseados em fatores têm por objetivo a projeção dos dados originais de grande dimensão especial para dimensões menores. Esta projeção pode ser realizada principalmente pela Análise de Componentes Principais (PCA), Análises de Fatores (FA) e Decomposição de valor singular (SVD).

A Análise de Componentes Principais (PCA) foi introduzido na química por Malinowski no final dos anos 60, com o nome de Análise de fatores, e a partir da década seguinte um grande número de aplicações foram desenvolvidas, o que a tornou muito conhecida e explorada.

O objetivo da PCA é transformar dados complexos para que as informações mais importantes e relevantes se tornem mais fáceis de visualizar<sup>6,7,8</sup>. Para isso, a matriz original  $\mathbf{X}$ , que é uma tabela onde os dados químicos multivariados são reconstruída para 2 matrizes menores denominadas matrizes de escores ( $\mathbf{T}$ ) e de pesos ( $\mathbf{P}$ ) e mais uma matriz de resíduos que indica a parte não modelada como mostrado na figura 7.

$$\begin{array}{c}
 \mathbf{X} \\
 \begin{array}{c} n \\ \hline \square \\ \hline m \end{array}
 \end{array}
 =
 \begin{array}{c}
 \mathbf{T} \\
 \begin{array}{c} d \\ \hline \square \\ \hline m \end{array}
 \end{array}
 \begin{array}{c}
 \mathbf{P}^T \\
 \begin{array}{c} n \\ \hline \square \\ \hline m \end{array}
 \end{array}
 +
 \begin{array}{c}
 \mathbf{E} \\
 \begin{array}{c} n \\ \hline \square \\ \hline m \end{array}
 \end{array}$$

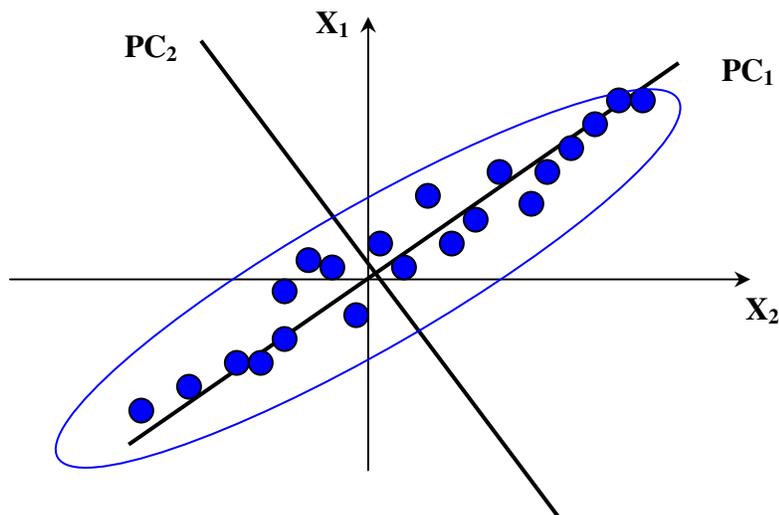
**Figura 7:** Decomposição da matriz  $\mathbf{X}$  nas matriz de escores ( $\mathbf{T}$ ), pesos ( $\mathbf{P}$ ) e resíduos ( $\mathbf{E}$ ).

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

As componentes principais são determinadas baseadas no critério de variância máxima. Cada componente principal subsequente descreve um máximo

de variância contido nos dados, sendo que a primeira componente principal determina a maior variância dos dados. No segundo componente principal, há menos informação do que no primeiro e, assim, sucessivamente.

A figura 8 representa graficamente duas componentes principais no caso de 2 variáveis ( $x_1$  e  $x_2$ ).



**Figura 8:** Representação de duas componentes principais.

Note pela figura 8 que a maior variância dos dados está sendo explicada pela primeira componente principal (PC1)- esferas azuis.

As novas coordenadas-PC1 e PC2 (que podem ser denominadas de Fatores ou Componentes Principais) são combinações lineares das variáveis originais. Por exemplo, os escores do primeiro componente principal são definidos como:

$$\begin{aligned} t_{11} &= x_{11}p_{11} + x_{12}p_{21} + \dots + x_{1n}p_{n1} \\ t_{21} &= x_{21}p_{11} + x_{22}p_{21} + \dots + x_{2n}p_{n1} \\ t_{m1} &= x_{m1}p_{11} + x_{m2}p_{21} + \dots + x_{mn}p_{n1} \end{aligned} \quad (\text{Equação 3})$$

Por causa da grande fração de variância usualmente descrita pelo primeiro, segundo e terceiro componentes, os dados podem ser visualizados através dos escores de um componente contra outro. O método mais simples usado para o cálculo dos escores e pesos em química analítica é o algoritmo NIPALS<sup>8</sup> (mínimo

quadrados parciais não linear iterativo). A análise de componentes principais (PCA) é muito aplicada para obter informações espectrais utilizadas no controle de qualidade de matérias-primas, produtos acabados bem como no monitoramento do produto durante a produção. Engenheiros de controle de processos industriais têm aplicado PCA, por exemplo, na detecção de falhas durante a operação de processos industriais<sup>9</sup>.

## 2.4 Regressão por Mínimos Quadrados Parciais

O método de regressão por mínimos quadrados parciais (PLS, do inglês, Partial Least Squares) pode ser considerado como o método mais utilizado para a construção de modelos de calibração multivariada. Esse método realiza uma análise de fatores em que a matriz de espectros é decomposta em matrizes de variações do espectro (pesos) onde os escores representam a contribuição de cada amostra<sup>10</sup>.

No PLS, tanto a matriz das variáveis independentes  $\mathbf{X}$ , como a das variáveis dependentes  $\mathbf{Y}$  são decompostas pela Análise de Componentes Principais:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (\text{Equação 4})$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F} \quad (\text{Equação 5})$$

Uma relação entre os dois blocos pode ser realizada correlacionando-se os escores para cada componente de cada vez, utilizando-se um modelo linear:

$$\mathbf{u}_h = b_h \mathbf{t}_h \quad (\text{Equação 6})$$

Onde :

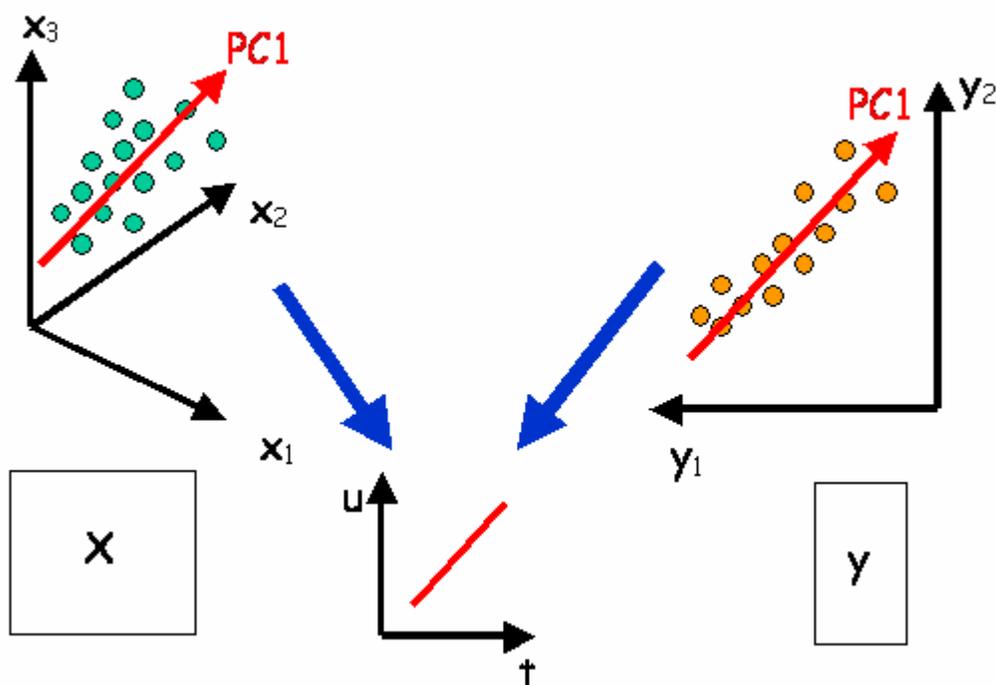
$$b_h = \frac{\mathbf{u}_h \mathbf{t}_h^T}{\mathbf{t}_h^T \mathbf{t}_h} \quad (\text{Equação 7})$$

Para cada  $h=1,2,\dots,A$  componentes principais

Esse modelo, entretanto, não é o melhor possível. Isto porque a Análise de Componentes Principais é realizada em cada matriz separadamente, podendo resultar numa relação não muito satisfatória (não linear) entre os escores dos dois

blocos. Deve-se buscar um modelo onde as matrizes de resíduos  $E$  e  $F$  sejam as menores possíveis e, ao mesmo tempo, conseguir uma relação linear entre  $t$  e  $u$ .

No PLS isto é realizado por uma leve mudança nos valores dos escores, de forma a produzir a melhor relação possível, ou seja, há uma leve rotação no eixo das componentes principais para obtenção de uma maior correlação com o vetor de dados  $y$ . Então, essas novas componentes principais passam a se chamar variáveis latentes<sup>11</sup>. Como pode ser notado, há um compromisso entre a habilidade dos componentes principais em descrever as amostras nos espaços individuais (modelagem dos blocos  $X$  e  $Y$ ), e o aumento na correlação entre  $u$  e  $t$ . A figura 9 mostra uma representação geométrica do PLS com uma variável latente modelando os dados do bloco de variáveis independentes ( $x$ ) e o das variáveis dependentes ( $y$ ).



**Figura 9:** Representação geométrica do PLS com uma variável latente modelando os dados de  $X$  (variáveis independentes) e  $Y$  (variáveis dependentes). Os dados  $X$  formados por três variáveis e dados de  $Y$  por duas variáveis.

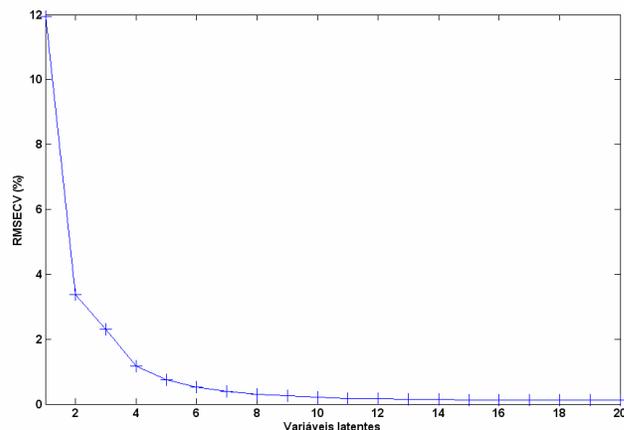
## 2.5 Validação Cruzada

Para a determinação do número correto de PC o método mais utilizado consiste no método de Validação Cruzada (CV – do inglês, Cross Validation), o qual se baseia na habilidade de previsão de um modelo construído por parte de um conjunto de dados seguido pela previsão do restante do conjunto de dados, que é realizada pelo modelo construído. A validação cruzada pode ser realizada em blocos, ou seja, um número determinado de amostras é deixado de fora no processo de construção do modelo e a seguir essas amostras são previstas pelo modelo construído, ou ainda por um caso conhecido como “leave one out” (deixe um fora), onde uma amostra é deixada de fora no processo de construção do modelo e a seguir essa amostra é prevista pelo modelo construído. Em ambos os casos, o processo é repetido até que todas as amostras tenham sido previstas e a raiz quadrada da soma do quadrado dos erros da validação cruzada (RMSECV – do inglês, Root Mean Square Error of Cross validation) é calculado<sup>12</sup>

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n-1} (\hat{y}_i - y_i)^2}{n}} \quad (\text{Equação 8})$$

Onde  $\hat{y}_i$  e  $y_i$  são, respectivamente, os valores previstos e de referência para a propriedade de interesse.

O cálculo é feito para o número de componentes principais de 1 até A, e os resultados de RMSECV são mostrados em um gráfico de RMSECV contra o número de variáveis latentes. Escolhe-se então o valor mínimo ou quando se tem variação significativa desse valor.



**Figura 10:** Gráfico do RMSECV em função da dimensionalidade do modelo de regressão.

O comportamento típico observado nesses gráficos é o aparecimento de um mínimo ou de um patamar, que indica a melhor dimensionalidade do modelo de regressão, ou seja, o menor número de variáveis latentes ou componentes principais que produziu o menor erro de previsão. Nesse caso pode ser escolhido 5 variáveis latentes.

## 2.6 Detecção de amostras anômalas em calibração multivariada

Como já dito, na construção de um modelo de calibração multivariada, geralmente são utilizados dois conjuntos de amostras, um para a calibração e outro para a validação do modelo. Entre essas amostras, há possibilidade da ocorrência de amostras anômalas, isto é, amostras que possuem comportamento muito diferente, quando comparado com o restante do conjunto.

A ocorrência desse tipo de amostras, entre as amostras de calibração, pode conduzir a modelos com uma baixa capacidade de previsão, isto é, que produzem altos valores de erro<sup>13</sup>. Quando presentes entre as amostras de validação, podem influenciar os resultados dos testes de validação, geralmente levando a resultados que indicam que o modelo não é adequado ou que a sua capacidade é inferior à que ele pode apresentar na ausência de tais amostras.

Hoje já existem diversas técnicas para a identificação de amostras anômalas<sup>14,15,16,17</sup>, sendo que esta área constitui uma linha de pesquisa que ainda é de grande importância, tendo em vista a variedade de tipos de equipamentos e amostras que geram dados de diferentes magnitudes e complexidade.

No presente trabalho, como forma de detecção de amostras anômalas, utilizou-se um dos métodos recomendados pela norma E1655-00 da ASTM (American Society for Testing and Materials)<sup>14</sup>, que se baseia na análise da influência da amostra dentro do conjunto.

O “peso relativo” de uma amostra em relação às demais presentes em um mesmo conjunto, pode ser calculado com base na distância que uma amostra está do centro dos dados. Essa característica que é descrita na literatura pelo termo inglês “leverage”, é determinada como:

$$h_i = \sum_{a=1}^A \frac{\hat{t}_{i,a}^2}{\hat{t}_a^T \hat{t}_a} \quad (\text{Equação 9})$$

Qualitativamente, tomando como por exemplo dados espectrais, o termo “h” mede o quanto o espectro de uma amostra difere dos espectros das demais amostras presentes nas amostras de calibração, amostras com altos valores de h são consideradas como amostras anômalas.

Para modelos de regressão linear, como por exemplo PLS, um valor médio de h para todos os espectros referentes às amostras de calibração é definido como a razão :  $A/n_{\text{cal}}$ , onde A é o número de variáveis do modelo, e  $n_{\text{cal}}$  é o número de amostras utilizadas no conjunto de calibração. Em média cada amostra de calibração contribui com  $A/n_{\text{cal}}$  para a definição das variáveis espectrais e dos coeficientes de regressão. Segundo a norma E1655-00 da ASTM<sup>14</sup>, amostras com h maior que três vezes esse valor devem ser retiradas do conjunto de amostras de calibração de um modelo.

As amostras anômalas identificadas segundo o procedimento descrito acima correspondem a amostras que possuem os dados experimentais diferentes do restante do conjunto.

Um outro tipo de amostra anômala são aquelas que possuem o valor da propriedade de interesse (**y**) significativamente diferente do valor esperado ou

apresentado por um método de referência. Essa espécie de amostras pode ser identificada por meio dos resíduos estudentizados. O resíduo estudentizado ( $e_{stud}$ ) de uma amostra  $i$ , pode ser determinado como:

$$e_{stud} = \frac{(y_i - \hat{y}_i)}{RMSEC - \sqrt{(1-h_i)}} \quad (\text{Equação 10})$$

Onde o RMSEC é a raiz quadrada do erro médio quadrático de calibração (do inglês, root mean square error of calibration), calculado como:

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n-1} (\hat{y}_i - y_i)^2}{n_{GL}}} \quad (\text{Equação 11})$$

## 2.7 Previsão

A parte fundamental da regressão é a verificação de sua validade. Ou seja, se o modelo desenvolvido é capaz de prever corretamente ou com pequena margem de erro os valores de novas amostras. Os espectros de infravermelho de amostras desconhecidas são dispostos na forma de matriz ( $\mathbf{X}_{teste}$ ). Após isto, os dados passam por pré-tratamentos matemáticos assim como os dados de calibração<sup>18</sup>.

O passo seguinte consiste em se obter o vetor de escores ( $t_{teste}$ ) e os resíduos ( $e$ ) pelo ajuste das matrizes dos novos espectros,  $\mathbf{X}_{teste}$  para  $p$  (pesos do bloco  $X$  calculados na fase de treinamento):

$$\mathbf{X}_{teste} = \sum_{a=1}^A \mathbf{T}_a \mathbf{P}_a^T + \mathbf{e} \quad (\text{Equação 12})$$

Para “A” – variáveis latentes na modelagem.

A partir dos coeficientes de regressão  $b$  calculados na fase de calibração, os escores  $\mathbf{T}$  teste produzem uma estimativa do vetor de concentrações, já que<sup>18</sup>:

$$\mathbf{Y} = \sum_{a=1}^A \mathbf{t}_a \mathbf{b}_a \mathbf{q}_a^T \quad (\text{Equação 13})$$

Para "A" variáveis latentes.

Onde  $\mathbf{q}_a$  (os pesos do bloco  $\mathbf{Y}$ ) foi calculado na fase de treinamento.

## 2.8 Sinal Analítico Líquido

O sinal analítico líquido, do inglês, "Net analyte signal" foi definido por Lorber<sup>19</sup> como sendo a parte do sinal analítico que é ortogonal às contribuições de outros possíveis constituintes presentes na amostra como mostrado na figura 11.

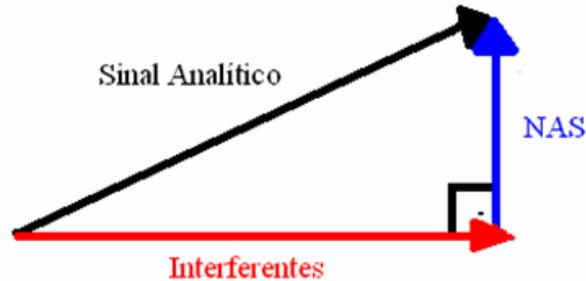


Figura 11: Representação geométrica da decomposição do sinal analítico.

Com o uso do NAS é possível calcular um valor escalar livre de interferentes a partir do vetor sinal analítico, o que torna possível a construção de uma nova forma de calibração, em que o modelo multivariado pode ser representado em uma forma pseudo-univariada. Para o cálculo do NAS usa-se a matriz de dados reconstruída com  $A$  variáveis latentes ( $\hat{\mathbf{x}}_A$ ) através do algoritmo SIMPLS<sup>20,21</sup>. Em seguida é determinada a matriz que contém a informação de todas as espécies presentes na amostra, exceto da espécie de interesse  $k$ , como<sup>22</sup>:

$$\hat{\mathbf{x}}_{A,-k} = [\mathbf{I} - \hat{\mathbf{y}}_{A,k} \hat{\mathbf{y}}_{A,k}^+] \hat{\mathbf{x}}_A \quad (\text{Equação 14})$$

em que

$\hat{\mathbf{x}}_{A,-k}$ : matriz que contém a informação de todas as espécies, exceto da espécie de interesse

$\hat{y}_{A,k}$  : vetor de concentrações da espécie de interesse  $k$  estimado com  $A$  variáveis latentes e o índice "+" sobrescrito indica a pseudo-inversa do vetor em questão.

**I**: matriz identidade

$\hat{x}_A$  : matriz de dados reconstruída com  $A$  variáveis latentes ( $\hat{x}_A$ ) através do algoritmo SIMPLS

Isso faz com que a matriz ( $\hat{x}_{A,-k}$ ) fique livre de qualquer contribuição da espécie  $k$ . Logo, o vetor NAS é obtido conforme a equação 2.

$$\hat{x}_{A,k}^{nas} = \left[ I - \hat{x}_{A,-k}^T \left( \hat{x}_{A,-k}^T \right)^+ \right] \hat{x}_A \quad (\text{Equação 15})$$

em que:

$\hat{x}_A$  : vetor de respostas instrumentais de uma amostra estimado com  $A$  variáveis latentes

$\hat{x}_{A,k}^{nas}$  : vetor NAS

Uma vez que  $\hat{x}_{A,k}^{nas}$  é livre de interferentes, é possível substituí-lo por uma representação escalar sem perda de informação. Assim temos:

$$n\hat{a}s = \left\| \hat{x}_{A,k}^{nas} \right\| \quad (\text{Equação 16})$$

onde  $\| \cdot \|$  representa a norma Euclidiana do vetor  $\hat{x}_{A,k}^{nas}$ . Com a possibilidade de calcular um valor escalar livre de interferentes, a partir de um vetor contendo contribuições de constituintes desconhecidos, torna-se possível a construção de uma nova forma de calibração multivariada, em que o modelo pode ser representado em uma forma univariada. Primeiramente o cálculo do NAS é feito para as  $i$  amostras de calibração; em seguida o coeficiente de regressão é determinado por mínimos quadrados entre o vetor **nâs** e o vetor de concentrações **y**:

$$\hat{b}_{nas} = (\hat{n}\hat{a}s^T \hat{n}\hat{a}s)^{-1} \hat{n}\hat{a}s^T y \quad (\text{Equação 17})$$

E o modelo de regressão pode, então, ser representado por:

$$\hat{y} = \hat{b}_{nas} \hat{n}\hat{a}s + e \quad (\text{Equação 18})$$

# **CAPÍTULO 3**

## **VALIDAÇÃO DE MODELOS DE CALIBRAÇÃO MULTIVARIADA ATRAVÉS DO CÁLCULO DE FIGURAS DE MÉRITO**

### 3. Validação e Figuras de Mérito

#### 3.1 Validação

Quando um novo método analítico é desenvolvido existe a necessidade das agências reguladoras tais como INMETRO (Instituto Nacional de Metrologia, Normalização e Qualidade Industrial), ANVISA (Agência Nacional de Vigilância Sanitária), Farmacopéia Norte-Americana (USP), *International Conference on Harmonization* (ICH) e *American Society for Testing & Materials* (ASTM) para certificar se o método apresenta uma performance adequada nas condições em que ele será aplicado. Esse processo de averiguação é denominado validação<sup>1</sup>. A validação se dá através da determinação de diversos parâmetros conhecidos como figuras de mérito. As principais figuras de mérito são<sup>1,2,3</sup>:

- Exatidão
- Ajuste do modelo
- Linearidade
- Sensibilidade
- Sensibilidade Analítica
- Seletividade
- Razão sinal/ruído
- Teste para erro sistemático ("bias")
- Limite de detecção
- Limite de quantificação

A maneira pela qual essas figuras de mérito devem ser determinadas é estabelecida pelos órgãos de fiscalização e encontra-se descrita em normas específicas<sup>4</sup>, guias de validação<sup>1,2,3,5,6</sup> e trabalhos científicos<sup>7,8,9</sup>.

Determinações quantitativas utilizando espectroscopia no infravermelho próximo tem crescido a cada ano, e estas determinações fazem uso de modelos de calibração multivariada nas quais são utilizado o cálculo de figuras de mérito para validação. Exemplos, referem-se a determinação de parâmetros

no controle de qualidade de indústrias alcooleiras<sup>10</sup> e determinações de propriedades de fármacos<sup>11</sup>.

A ausência de validação desses modelos, devido à carência de normas oficiais que descrevem como essa validação deve ser realizada tem restringido sua implementação. Nos últimos anos, uma considerável atenção vem sendo direcionada para a elaboração de guias, normas e trabalhos científicos que enfocam a necessidade da validação de modelos de calibração multivariada, de modo que, atualmente, já se encontram disponíveis trabalhos que descrevem procedimentos e propostas de como essa validação deve ser realizada. Alguns documentos apresentam caráter geral, enquanto que outros, como por exemplo a norma E1655-00 da ASTM<sup>4</sup> descrevem especificamente o desenvolvimento e validação de modelos de calibração multivariada a partir da espectroscopia no infravermelho. Muitos trabalhos científicos tratam de problemas específicos, e portanto, abordam a determinação de apenas algumas figuras de mérito específicas relacionadas ao trabalho realizado. São exemplos destes trabalhos: a determinação de limite de detecção, estimativa de intervalos de confiança, determinação de sensibilidade e sensibilidade analítica<sup>14</sup>, determinação da seletividade, precisão no nível de repetibilidade e precisão intermediária<sup>15</sup>, sendo que, estes trabalhos mostram que figuras de mérito, raramente abordadas durante a validação de modelos multivariados, podem ser determinadas. São exemplos ainda, trabalhos que propõem melhora de definições e relações entre diferentes definições para uma mesma figura de mérito.

## **3.2 Figuras de Mérito**

### **3.2.1 Exatidão**

Expressa o grau de concordância entre o valor estimado ou medido e o valor tido como verdadeiro ou de referência<sup>2</sup>. Comumente em aplicações com calibração multivariada, a exatidão é estimada através da raiz quadrada do erro

médio quadrático de previsão (RMSEP – "Root Mean Squares Error of Prediction"):

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (\text{Equação 19})$$

onde  $n$  é o número de amostras de validação,  $y_i$  é o vetor referência que contém a concentração do analito de interesse e  $\hat{y}_i$  é o vetor que contém os valores previstos da concentração dada pelo modelo. Segundo o protocolo emitido pela ICH<sup>2</sup>, a exatidão deve ser determinada usando um mínimo de nove determinações em no mínimo três níveis de concentração que cubram a faixa de aplicabilidade do modelo. Nessa dissertação ela foi determinada com base nas 20 amostras utilizadas no conjunto de validação.

### 3.2.2 Ajuste do modelo

O grau de ajuste do modelo pode ser estimado a partir da correlação entre os valores de referência e os valores estimados da propriedade de interesse pelo modelo para as amostras de calibração. Isso é feito determinando-se, por mínimos quadrados, a reta que melhor se ajusta aos valores de referência e os estimados pelo modelo<sup>15, 16</sup>. Uma vez que o escalar "nas" é determinado, o ajuste do modelo também pode ser determinado através da melhor reta que se ajusta ao gráfico do "nas" contra a concentração, para as amostras de calibração.

### 3.2.3 Linearidade

Em modelos de calibração multivariada uma medida quantitativa da linearidade não corresponde a uma tarefa simples, ou mesmo possível. Qualitativamente, gráficos de resíduos e dos escores contra a concentração, os quais devem ter comportamento aleatório e linear, respectivamente, podem indicar se os dados seguem ou não o comportamento linear.

### 3.2.4 Sensibilidade

É definida como fração de sinal responsável pelo acréscimo de uma unidade de concentração da propriedade de interesse. Para modelos de calibração multivariada, como a Regressão por Mínimos Quadrados Parciais-PLS, pode ser determinada como<sup>17,18</sup>

$$\text{S\hat{E}N} = \frac{1}{\|\mathbf{b}\|} \quad (\text{Equação 20})$$

ou, quando o NAS é determinado, o vetor de sensibilidade líquida para cada amostra  $i$  presente no conjunto de calibração pode ser determinado a partir do vetor  $\hat{\mathbf{x}}_{i,x}^{\text{cal}}$  como:

$$\hat{s}_i^{\text{cal}} = \frac{\hat{\mathbf{x}}_{i,x}^{\text{cal}}}{y} \quad (\text{Equação 21})$$

Onde o vetor de sensibilidades  $\hat{s}_i^{\text{cal}}$  deve ser igual para todas as amostras de calibração. O escalar S\hat{E}N pode ser determinado por:

$$\text{S\hat{E}N} = \frac{1}{\|\mathbf{s}^{\text{cal}}\|} \quad (\text{Equação 22})$$

### 3.2.5 Sensibilidade analítica

É definida como a razão entre a sensibilidade e uma estimativa do desvio-padrão para a flutuação do sinal analítico ( $\delta x$ )<sup>19,20</sup>:

$$\gamma = \frac{\text{S\hat{E}N}}{\|\hat{\Delta}c\|} \quad (\text{Equação 23})$$

onde o inverso desse parâmetro ( $\gamma^{-1}$ ) permite estabelecer a menor diferença de concentração entre amostras, a qual pode ser distinguida pelo método, e  $\delta x$  é estimado por meio do desvio-padrão para 15 espectros do sinal do branco (espectro obtido sem a presença de medicamento).

### 3.2.6 Seletividade

É a medida do grau de sobreposição entre o sinal da espécie de interesse e os interferentes presentes na amostra e indica a parte do sinal que é perdida por essa sobreposição<sup>21</sup>. Para modelos de calibração multivariada esse parâmetro é definido como<sup>12, 13</sup>:

$$SEL_i = \frac{n\hat{a}s_i}{\|x_i\|} \quad (\text{Equação 24})$$

onde  $n\hat{a}s_i$  é o escalar NAS estimado para amostra 'i' e  $x_i$  o vetor de dados original. A seletividade calculada a partir da equação 24 não se refere ao sentido geralmente empregado para o termo em química analítica, com modelos univariados, e sim a uma forma de estimar quanto do sinal é perdido por ortogonalidade.

### 3.2.7 Razão sinal/ruído

Indica quanto da intensidade do NAS da espécie de interesse está acima da estimativa do desvio-padrão para a flutuação do sinal analítico ( $\delta x$ ). É definido através do NAS como<sup>17,18</sup>:

$$S/R_i = \frac{n\hat{a}s_i}{\delta x} \quad (\text{Equação 25})$$

### 3.2.8 Teste para erro sistemático ("bias")

Segundo a IUPAC, "bias" é definido como a diferença entre a média populacional e o seu valor verdadeiro. Erros sistemáticos são todas as componentes de erros que não são aleatórios, logo, eles correspondem à parcela fixa do "bias" em uma medida química<sup>7</sup>. A existência desse tipo de erro afeta a exatidão e a determinação dos intervalos de confiança. A norma E1655-

00 da ASTM<sup>4</sup> sugere um teste-t para avaliar se o "bias" incluso no modelo é significativo. O "bias" médio pode ser calculado como:

$$\text{bias} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (\text{Equação 26})$$

Em seguida, o desvio-padrão dos erros de validação (SDV, "standard deviation of validation") é estimado por:

$$\text{SDV} = \sqrt{\frac{\sum [(y - \hat{y}) - \text{bias}]^2}{n - 1}} \quad (\text{Equação 27})$$

O valor de t é então obtido:

$$t = \frac{|\text{bias}| \sqrt{n}}{\text{SDV}} \quad (\text{Equação 28})$$

Caso o valor de t calculado seja menor que o seu valor crítico para n graus de liberdade com 95% de confiança, o "bias" incluído no modelo pode ser considerado insignificante e desprezado.

### 3.2.9 Limite de Detecção e Quantificação

O limite de detecção (LD) e o limite de quantificação (LQ) de um procedimento analítico, expressam as menores quantidades da espécie de interesse que podem ser detectadas e determinadas quantitativamente, respectivamente<sup>3</sup>. Para um conjunto de dados que apresenta comportamento no qual a variância é constante ao longo da faixa de trabalho, erros com previsão não correlacionados e que seguem uma distribuição normal, os LD e LQ na calibração multivariada podem ser calculados por<sup>15</sup>:

$$LD = 3\delta_x \|\mathbf{b}_k\| = 3\delta_x \frac{1}{S\hat{E}N} \quad (\text{Equação 29})$$

$$LQ = 10\delta_x \|\mathbf{b}_k\| = 10\delta_x \frac{1}{S\hat{E}N} \quad (\text{Equação 30})$$

em que,  $\delta_x$  é o desvio padrão do sinal de referência,  $\mathbf{b}_k$  é o vetor dos coeficientes de regressão do modelo PLS para a espécie k,  $S\hat{E}N$  corresponde ao valor de sensibilidade obtido através das equações 20 e 22.

# **CAPÍTULO 4**

## **CARTAS DE CONTROLE MULTIVARIADAS**

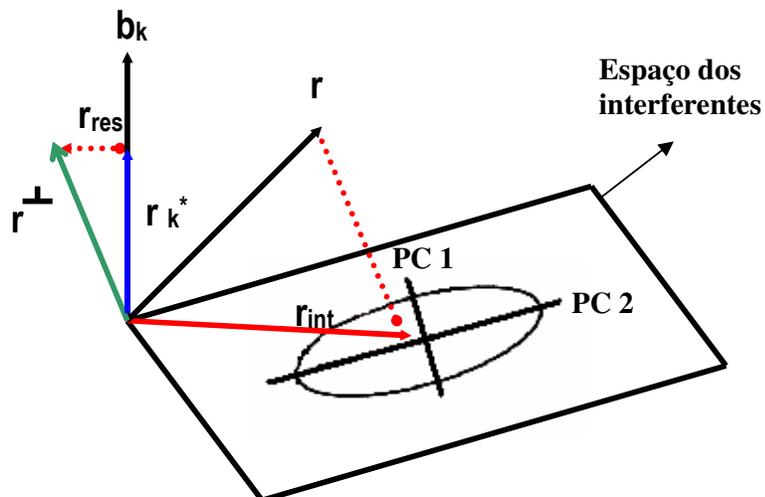
## 4.1 Cartas de Controle Multivariadas

O controle de processo estatístico multivariado utilizado para checar o desempenho de processos no controle de qualidade está se tornando cada vez mais importante, diante do crescimento do número de variáveis que podem ser medidas<sup>1,2,3,4</sup>. Um dos exemplos da implementação de ferramentas estatísticas multivariadas são as cartas de controle estatísticas baseadas no sinal analítico líquido que levam em conta a divisão do espectro em quatro constituintes: (NAS+resíduos), NAS, resíduos e interferentes. Essas cartas junto com as cartas Hotteling<sup>5,6,7,8,9</sup> vêm sendo usadas como uma ferramenta nova para substituição das cartas de controle univariadas, tais como as cartas Shewart<sup>10,11,12,13,14</sup>.

## 4.2 Carta de Controle Multivariada baseado na divisão do espectro em NAS, Interferente e Resíduos

Na figura 12 o vetor “ $\mathbf{r}$ ” representa o vetor de dados, enquanto que, o vetor “ $\mathbf{r}^\perp$ ” representa o vetor resultante da decomposição do vetor “ $\mathbf{r}$ ” que é perpendicular ao vetor interferente “ $\mathbf{r}_{int}$ ”. Realizando a decomposição do vetor “ $\mathbf{r}^\perp$ ” encontra-se como resultado o vetor NAS “ $\mathbf{r}_k^*$ ” e o vetor resíduos ( $\mathbf{r}_{res}$ ). As linhas pontilhadas em vermelho, na figura 12, representam a projeção do vetor ( $\mathbf{r}$ ) no espaço dos interferentes e o vetor resíduo ( $\mathbf{r}_{res}$ ). Enquanto que ( $\mathbf{b}_k$ ) representa a direção na qual foi projetado o vetor analito ( $\mathbf{r}_k^*$ ).

A figura geométrica que se aproxima de uma elipse mostrada no espaço dos interferentes da figura 12 indica o limite tolerável nesse espaço para dizer se determinada amostra está sob ou fora de controle, enquanto que, os eixos representados dentro dessa figura representam as duas primeiras componentes principais obtidas pelo PCA.



**Figura 12 :** Representação da divisão do espectro ( $r$ ) em quatro diferentes contribuições:

$r^\perp$  (NAS + resíduos), NAS ( $r_{k^*}$ ), resíduos ( $r_{RES}$ ) e interferente ( $r_{INT}$ ).

Após essa divisão, são criadas três cartas de controle (NAS, interferentes e resíduos) para cada parte da divisão espectral, com o intuito de monitorar se as amostras estão dentro ou fora de controle. Para isso são necessárias três fases: construção do modelo, cálculo dos limites estatísticos e validação do modelo

### 4.3 Construção das Cartas de Controle Baseada no Sinal Analítico Líquido

Para a construção das cartas de controle são necessárias duas partes. A primeira refere-se à criação do espaço dos interferentes que descreve todas as variações dos constituintes da amostra, exceto do analito, enquanto a segunda baseia-se no cálculo do vetor de regressão NAS ( $b_k$ )<sup>7,15</sup>. Para isso, dois grupos de informação são necessários: amostras que contêm informação somente do excipiente, denominada de placebo, e amostras que contêm a informação do medicamento. Para o cálculo do espaço dos interferentes utiliza-se a matriz placebo ( $R_k$ ). Para isso é feito o PCA (Análise de componentes principais)<sup>16,17</sup> da matriz placebo com “A” componentes principais conforme a equação 31.

$$\mathbf{R}_{-k} = \mathbf{TP}^T + \mathbf{E} \quad (\text{Equação 31})$$

Em que:

**R<sub>-k</sub>**: matriz de dimensão (j×l<sub>b</sub>), onde “j” refere-se ao número de variáveis e “l<sub>b</sub>”, ao número de amostras do excipiente utilizadas para construir o espaço dos interferentes. O subscrito “-k” denota que a matriz não possui informação do analito.

**P**: matriz “loading” de dimensão (j×A) onde “j” refere-se ao número de variáveis e “A”, ao número de componentes principais usado pelo PCA.

**T**: matriz score de dimensão (l<sub>b</sub>×A).

**E**: matriz resíduos de dimensão (j×l<sub>b</sub>).

Após essa fase, utiliza-se a matriz de “loading” para definir o espaço dos interferentes conforme a equação.

$$\mathbf{r}_{\text{int}} = \mathbf{PP}^+ \mathbf{r} \quad (\text{Equação 32})$$

Onde

**r<sub>int</sub>**: vetor interferente de dimensão (j×1), onde “j” refere-se ao número de variáveis.

**r**: vetor de dados que contém a informação do analito junto com a informação do excipiente com dimensão (j×1).

**P<sup>+</sup>**: pseudo-inversa da matriz “loading”.

A segunda parte da construção do modelo leva em conta o cálculo do vetor de regressão NAS de acordo com as equações 33 e 34.

$$\mathbf{B}_k = \left( \mathbf{I}_j - \mathbf{PP}^T \right) \mathbf{R}_{\text{mod}} \quad (\text{Equação 33})$$

Onde

**B<sub>k</sub>**: matriz que contém os vetores que são ortogonais ao espaço dos interferentes de dimensão (j×l<sub>mod</sub>) em que “l<sub>mod</sub>” é o número de amostras dentro de controle utilizadas para o cálculo do vetor de regressão NAS, enquanto o subscrito “k” refere-se à matriz que contém a informação do analito de interesse.

De acordo com Smilde e Skibsted<sup>7</sup>, é recomendado usar cinco amostras para o cálculo da matriz **B<sub>k</sub>**, pois não há diferença se for usado um número maior. Logo, o número de amostras (l<sub>mod</sub>) utilizadas nessa dissertação é cinco.

**I<sub>j</sub>**: matriz identidade de dimensão (j×j)

$\mathbf{R}_{\text{mod}}$ : matriz de dados na qual existem somente amostras dentro de controle e com dimensão  $(j \times I_{\text{mod}})$ .

Após isso, é calculado a média dos vetores da matriz  $\mathbf{B}_k$ , com o intuito de achar uma direção única para os vetores ortogonais ao espaço dos interferentes. Esse novo vetor é denominado vetor de regressão NAS e é calculado da seguinte forma:

$$\mathbf{b}_k = \frac{\sum_{i=1}^{I_{\text{mod}}} \mathbf{B}_{k,i}}{I_{\text{mod}}} \quad (\text{Equação 34})$$

$\mathbf{b}_k$ : vetor de regressão NAS com dimensão  $(j \times 1)$ .

Agora o vetor NAS ( $\mathbf{r}_k^*$ ) pode ser computado pela projeção do espectro na direção do vetor de regressão NAS, conforme mostrado na figura 12, através da equação 35.

$$\mathbf{r}_k^* = \mathbf{b}_k \left( \mathbf{b}_k^T \mathbf{b}_k \right)^{-1} \mathbf{b}_k^T \mathbf{r} \quad (\text{Equação 35})$$

É importante notar, na figura 12, que existe uma pequena diferença entre  $\mathbf{r}^\perp$  e  $\mathbf{r}_k^*$  que é o vetor resíduo. Então, como:

$$\mathbf{r} = \mathbf{r}_{\text{int}} + \mathbf{r}_k^* + \mathbf{r}_{\text{res}} \quad (\text{Equação 36})$$

logo,

$$\mathbf{r}_{\text{res}} = \mathbf{r} - \mathbf{r}_{\text{int}} - \mathbf{r}_k^* \quad (\text{Equação 37})$$

onde:

$\mathbf{r}_{\text{res}}$  : vetor resíduo

$\mathbf{r}_k^*$  : vetor NAS

$\mathbf{r}_{\text{int}}$  : vetor interferente

$\mathbf{r}$  : vetor de dados

#### 4.4 Cálculo dos limites estatísticos para as cartas NAS, Interferente e Resíduos

Para o cálculo dos limites estatísticos das três cartas citadas acima utiliza-se somente amostras que estão dentro de controle. Essas amostras são organizadas em uma matriz denominada(  $\mathbf{R}_{noc}$  ) e possuem dimensão  $(j \times I_{noc})$  em que “ $I_{noc}$ ” é o número de amostras dentro de controle e “ $j$ ” o número de variáveis espectrais. É importante ressaltar que essas amostras não foram usadas para construir o espaço dos interferentes, nem para fazer o cálculo do vetor de regressão NAS. O subscrito “noc” (do inglês, “normal operating condition”<sup>17</sup>) refere que essa matriz só possui amostras dentro de controle. Nesta dissertação usou-se  $I_{noc}$  igual a 19.

#### 4.5 Carta de Controle NAS

A projeção do espectro da amostra no vetor de regressão NAS é proporcional à concentração do analito na amostra. Essa projeção é denominada NAS<sup>18</sup>. Os limites estatísticos da carta NAS para cada amostra dentro de controle são calculados a seguir:

$$nas_{noc} = \mathbf{R}_{noc}^T \mathbf{b}_k \quad (\text{Equação 38})$$

onde  $\mathbf{nas}_{noc}$  é o vetor com os valores do NAS para cada amostra sob controle da matriz  $\mathbf{R}_{noc}$ .

Calculando a média e o desvio-padrão do vetor  $\mathbf{nas}_{noc}$ , pode-se computar e graficar os limites estatísticos da carta NAS, como mostra a equação 39.

Considerando o limite com 95% de confiança<sup>19</sup>, temos:

$$Limite_{95\%} = \overline{nas}_{noc} \pm 2s_{noc} \quad (\text{Equação 39})$$

onde:

$\overline{nas}_{noc}$  : média do vetor  $\mathbf{nas}_{noc}$  que possui dimensão  $(1 \times 1)$

$s_{noc}$  : desvio-padrão do vetor  $\mathbf{nas}_{noc}$  que possui dimensão  $(1 \times 1)$

Assume-se que os valores do NAS seguem uma distribuição normal, que pode ser detectada por um teste de normalidade. A figura 13 abaixo representa a carta NAS.

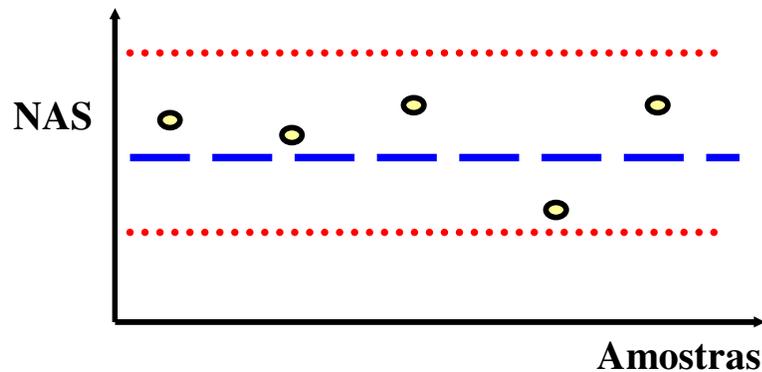


Figura 13 : Representação da Carta de Controle NAS

#### 4.6 Carta de Controle Interferente

Todos os vetores que fazem parte da matriz  $\mathbf{R}_{noc}$  têm uma fração projetada no espaço dos interferentes, como mostrado na figura 12. Amostras futuras também devem ser projetadas nesse espaço. Dessa forma, amostras projetadas dentro do limite representado pela figura geométrica, que se aproxima de uma elipse, que está graficada no espaço dos interferentes segundo a figura 12, são denominadas dentro de controle, enquanto as amostras projetadas além desses limites, são denominadas fora de controle. Vale lembrar que o espaço dos interferentes foi construído através do modelo PCA, no qual se usa a matriz de *loadings*, conforme mostrado na equação 32. Novamente, na figura 12, temos que as duas retas que se encontram dentro do limite referem-se às duas primeiras componentes principais do PCA. Sendo assim, pode-se dizer que a carta interferente é baseada na distância da projeção das amostras ao centro dos eixos formados pelas duas primeiras componentes principais. Para construir a carta interferente, primeiramente, deve-se projetar os vetores que são os espectros das amostras que estão dentro de controle no espaço dos interferentes, de acordo com a equação 40.

$$\mathbf{R}_{int} = \mathbf{P}\mathbf{P}^+ \mathbf{R}_{noc} \quad (\text{Equação 40})$$

Onde:

$\mathbf{R}_{\text{int}}$ : matriz interferente com dimensão  $(j \times I_{\text{noc}})$

Então, é necessário calcular as novas coordenadas para esses interferentes em um novo sistema de eixos (matriz score), da seguinte forma:

$$\mathbf{T}_{\text{noc}} = \mathbf{R}_{\text{int}}^T \mathbf{P} (\mathbf{P}^T \mathbf{P})^{-1} \quad (\text{Equação 41})$$

Agora, a distância das amostras ao centro do eixo das duas componentes principais é calculada

$$d_{\text{noc},i} = \left( \mathbf{t}_{\text{noc},i} - \bar{\mathbf{t}}_{\text{noc}} \right)^T \mathbf{S}^{-1} \left( \mathbf{t}_{\text{noc},i} - \bar{\mathbf{t}}_{\text{noc}} \right) \quad (\text{Equação 42})$$

onde

$\mathbf{T}_{\text{noc}}$ : matriz score com dimensão  $(I_{\text{noc}} \times A)$ , em que “A” é o número de componentes principais retidos no modelo PCA.

$\bar{\mathbf{t}}_{\text{noc}}$ : valor médio para todos os vetores contidos na matriz  $\mathbf{T}_{\text{noc}}$ .

$\mathbf{S}$ : denota a matriz de covariância centrada na média dos scores com dimensão  $(A \times A)$ .

$\mathbf{t}_{\text{noc},i}$ : vetor score com dimensão  $(1 \times A)$

$d_{\text{noc},i}$ : distância dos interferentes ao centro dos eixos formados pelas duas primeiras componentes principais.

Dessa forma, o limite com 95% confiança para a carta interferente pode ser calculado e graficado na figura 14 :

$$D_{\text{limite},95\%} = F_{0,95} \left( A, I_{\text{noc}} - A \right) \left[ \frac{A \left( I_{\text{noc}}^2 - I_{\text{noc}} \right)}{I_{\text{noc}} \left( I_{\text{noc}} - A \right)} \right] \quad (\text{Equação 43})$$

onde

$F_{0,95}$ : valor da distribuição F com 95% de confiança em que o índice  $(I_{\text{noc}} - A)$  refere-se ao número de graus de liberdade.

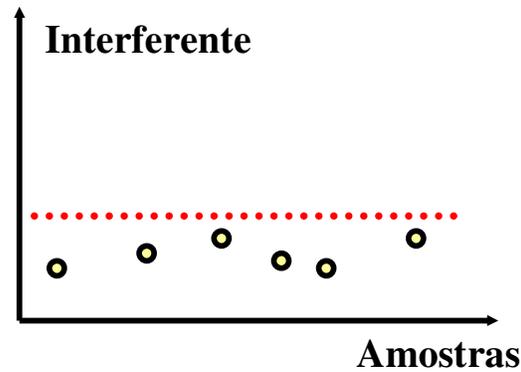


Figura 14 : Representação da Carta de Controle Interferente

#### 4.7 Carta de Controle do Resíduo

A carta resíduo baseia-se na parte do sinal que não é modelada pelo modelo e pelo limite estatístico- $Q^{20}$ . De acordo com a figura 12, temos que o resíduo provém da decomposição do vetor que é ortogonal ao vetor interferente ( $\mathbf{r}^T$ ). Vale ressaltar que, a partir da decomposição desse vetor, temos como resultado o vetor NAS e o vetor resíduo.

Para o cálculo do resíduo é necessário o cálculo da matriz NAS (equação 44) e da matriz interferente (equação 40). Dessa forma, calcula-se a matriz resíduo segundo a equação 45.

$$\mathbf{R}_{noc}^* = \mathbf{b}_k \left( \mathbf{b}_k^T \mathbf{b}_k \right)^{-1} \mathbf{b}_k^T \mathbf{R}_{noc} \quad (\text{Equação 44})$$

Onde:

$\mathbf{R}_{noc}^*$  : matriz NAS que possui dimensão  $(j \times l_{noc})$

$$\mathbf{R}_{res} = \mathbf{R}_{noc} - \mathbf{R}_{int} - \mathbf{R}_{noc}^* \quad (\text{Equação 45})$$

Então, de acordo com Jackson e Mudholkar<sup>26</sup>, é calculado o resíduo  $Q_{noc}$  da seguinte forma:

$$\mathbf{Q}_{noc} = \mathbf{R}_{res}^T \mathbf{R}_{res} \quad (\text{Equação 46})$$

onde:

$\mathbf{Q}_{noc}$ : matriz com dimensão  $(l_{noc} \times l_{noc})$ .

A diagonal da matriz  $\mathbf{Q}_{noc}$  contém os valores Q-estatístico que serão graficados na carta resíduo.

O limite de confiança para carta  $Q^{20}$  é calculado e plotado na figura 15 conforme as equações 47, 48 e 49.

$$Q_\alpha = \theta_1 \left[ \frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (\text{Equação 47})$$

Onde

$c_\alpha$ : desvio-padrão normal

$$\theta_i = \sum_{j=k+1}^n \lambda_j^i \quad i: 1, 2 \text{ e } 3 \quad (\text{Equação 48})$$

“k” corresponde ao número de componentes principais retidos no modelo e “n” ao número total de componentes principais (ou o número total de amostras).

$$h_0 = 1 - \left( \frac{2\theta_1\theta_3}{3\theta_2^2} \right) \quad (\text{Equação 49})$$

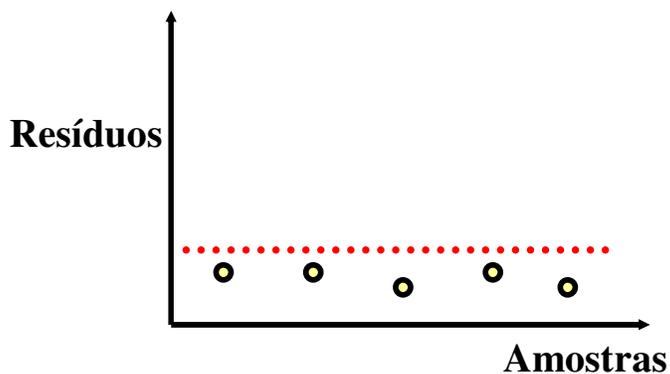


Figura 15 : Representação da Carta de Controle Resíduo

## 4.8 Validação das Cartas de Controle (Modelo)

### 4.8.1 Carta de controle NAS para novas amostras

O valor NAS é calculado de acordo com a equação 50 e graficado na carta NAS de calibração segundo a figura 16.

$$nas_{novo} = R_{novo}^T b_k \quad (\text{Equação 50})$$

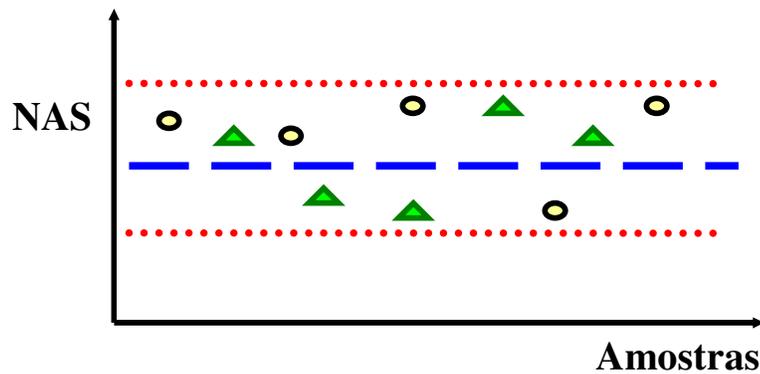


Figura 16 : Representação da Carta de Controle NAS.

● Amostras de Calibração; ▲ Amostras de Validação

### 4.8.2 Carta de controle do interferente para novas amostras

O valor dos interferentes para as novas amostras são calculados e graficados na carta interferente de calibração de acordo com a figura 17 através das equações 51, 52 e 53.

$$R_{int,novo} = PP^+ R_{novo} \quad (\text{Equação 51})$$

$$T_{novo} = R_{novo}^T P(P^T P)^{-1} \quad (\text{Equação 52})$$

$$d_{novo,i} = \left( t_{novo,i} - \bar{t}_{novo} \right)^T S^{-1} \left( t_{novo,i} - \bar{t}_{novo} \right) \quad (\text{Equação 53})$$

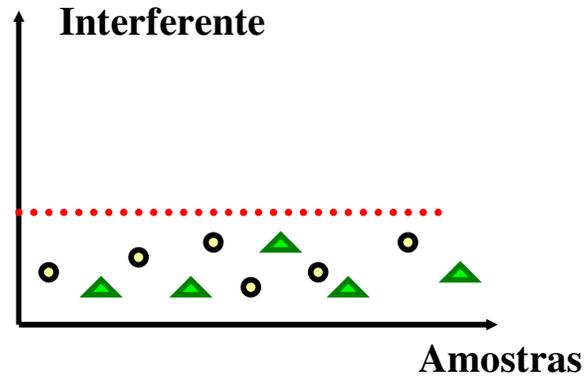


Figura 17 : Representação da Carta de Controle Interferente.

● Amostras de Calibração; ▲ Amostras de Validação

#### 4.8.3 Carta de controle dos resíduos para novas amostras

Também os resíduos das novas amostras são calculados e graficados na carta resíduo de calibração, como mostrado na figura 18 através das equações 54, 55 e 56 abaixo.

$$\mathbf{R}_{novo,k}^* = \mathbf{b}_k \left( \mathbf{b}_k^T \mathbf{b}_k \right)^{-1} \mathbf{b}_k^T \mathbf{R}_{novo} \quad (\text{Equação 54})$$

$$\mathbf{R}_{res,novo} = \mathbf{R}_{novo} - \mathbf{R}_{int,novo} - \mathbf{R}_{novo,k}^* \quad (\text{Equação 55})$$

$$\mathbf{Q}_{novo} = \mathbf{R}_{res,novo}^T \mathbf{R}_{res,novo} \quad (\text{Equação 56})$$

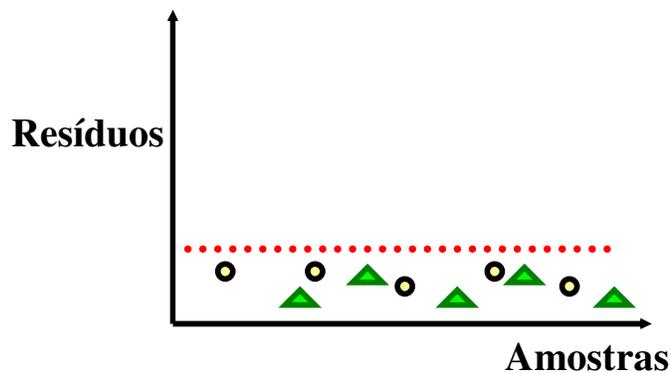


Figura 18 : Representação da Carta de Controle Resíduos.

● Amostras de Calibração; ▲ Amostras de Validação

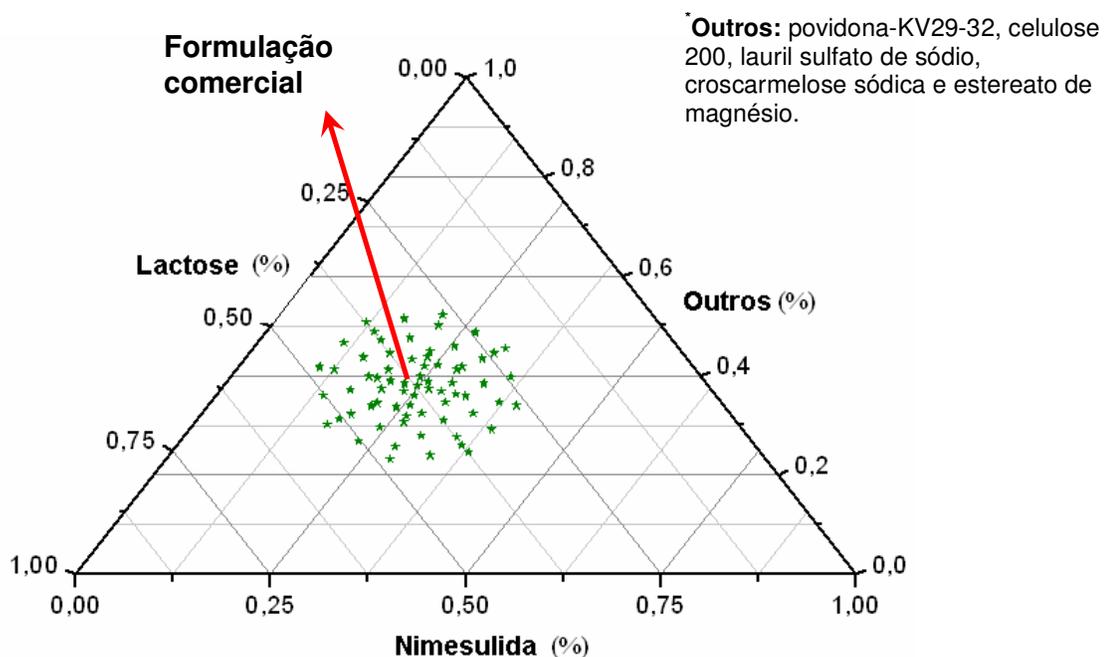
# **CAPÍTULO 5**

**DETERMINAÇÃO DE NIMESULIDA EM  
MEDICAMENTOS ATRAVÉS DO CÁLCULO DE  
FIGURAS DE MÉRITO E DE ESPECTROSCOPIA NO  
INFRAVERMELHO PRÓXIMO**

Na primeira etapa desta dissertação aplicou-se espectroscopia no infravermelho próximo e cálculo de figuras de mérito, baseado no cálculo do sinal analítico líquido, para construir e validar um modelo de calibração multivariada linear por mínimos quadrados parciais (PLS).

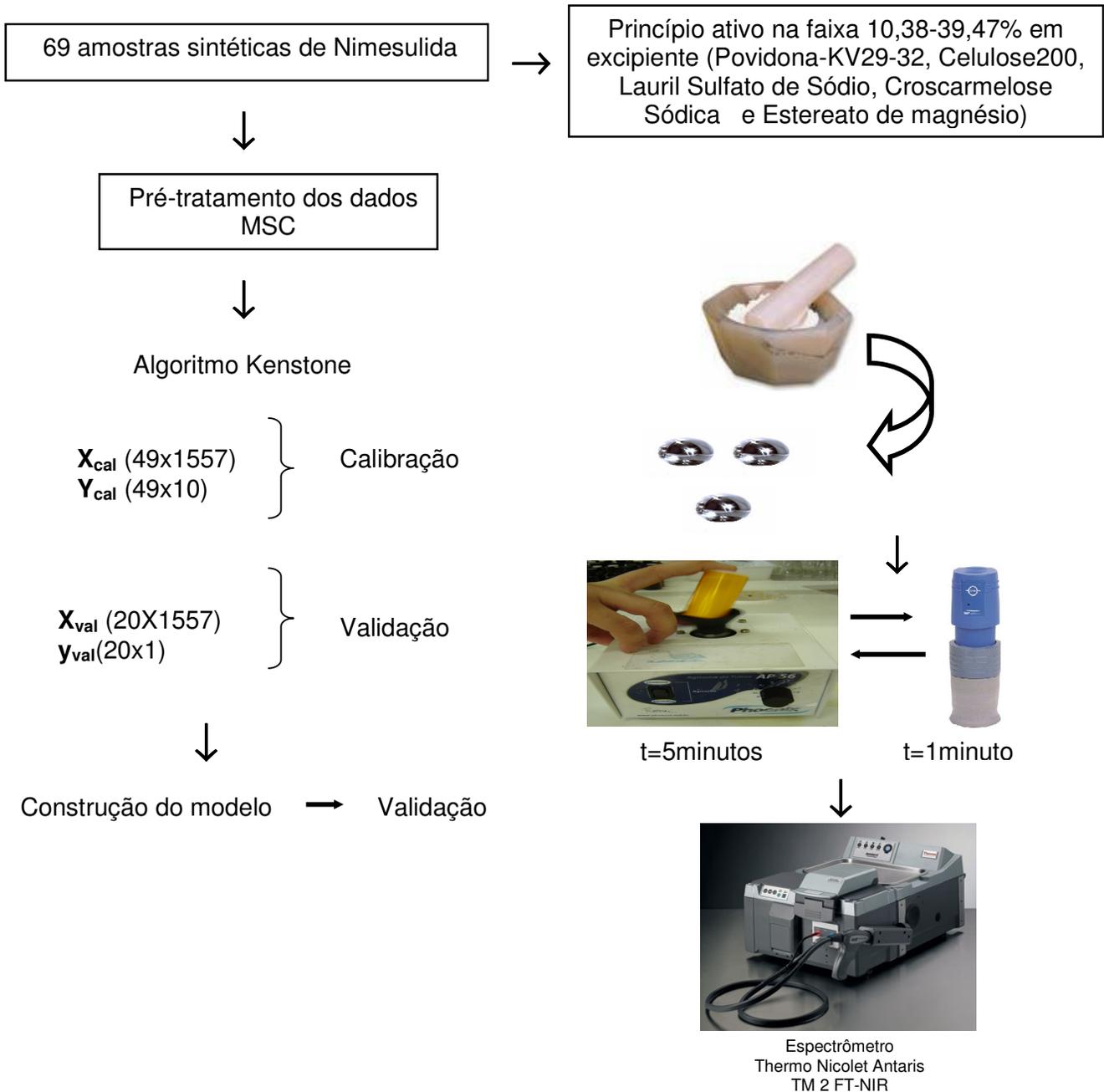
## 5-Parte Experimental

Foram preparadas 69 amostras sintéticas contendo o princípio ativo (Nimesulida) na faixa 10,38-39,47% (m/m) em excipiente (lactose, povidona-KV29-32, celulose 200, lauril sulfato de sódio, croscarmelose sódica e estereato de magnésio), conforme mostrado na figura 19 abaixo.



**Figura 19:** Diagrama ternário para a mistura física do medicamento Nimesulida.

Todas as amostras foram pesadas em uma balança analítica com precisão de 0,00001g e misturadas através de vórtex e almofariz por cinco minutos em recipientes de 5mL com a ajuda de três esferas metálicas. Após isso, todas as amostras de Nimesulida foram colocadas em um moinho de facas por um minuto. Depois as amostras foram recolocadas no vórtex por mais um minuto.



**Figura 20:** Metodologia utilizada para construção e validação do modelo da mistura física do medicamento Nimesulida.

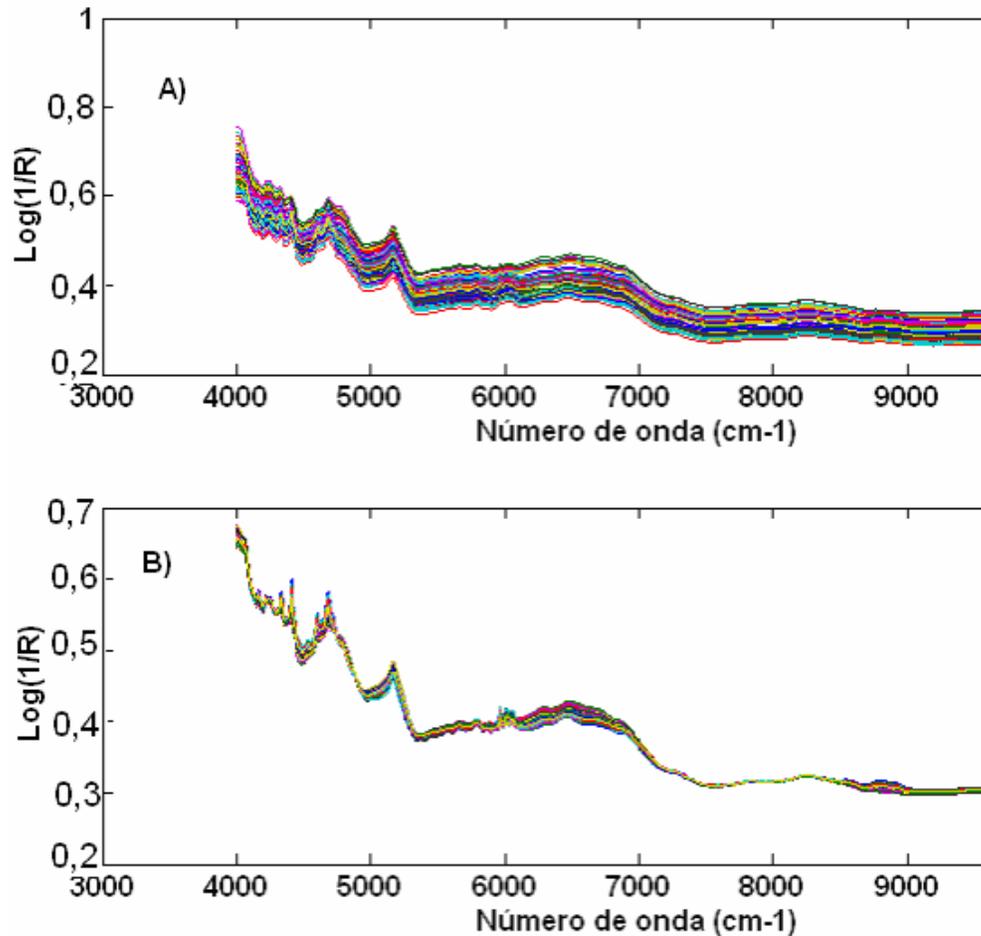
Também foram utilizados, nas mesmas condições experimentais, três lotes de amostras comerciais do medicamento Nimesulida cuja concentração foi previamente determinada pelo método padrão que utiliza diluição seguida de medida espectrofotométrica na região do UV.

## **5-1 Espectroscopia no infravermelho próximo**

Os espectros de reflectância foram obtidos num espectrômetro Thermo Nicolet Antaris TM II FT-NIR com módulo de esfera de integração de ouro para reflectância difusa, fonte de tungstênio-halogênio, 32 varreduras, resolução de  $8\text{cm}^{-1}$  e detector de arseneto de gálio.

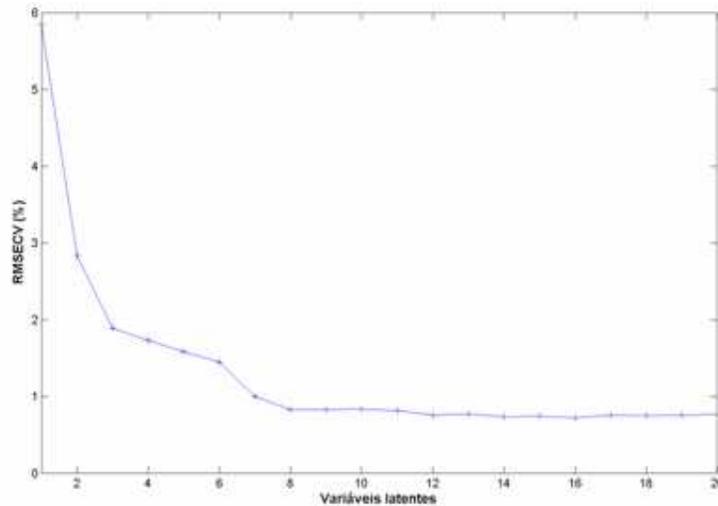
## **5-2 Modelo de Regressão PLS**

Na construção dos modelos foi utilizado o PLS-ToolBox 3.5<sup>1</sup> para MATLAB 6.5 (MathWorks). O primeiro passo para a construção do modelo foi realizar o pré-processamento dos dados, pois os espectros das amostras apresentaram desvio da linha base. O pré-processamento utilizado para a matriz **X** foi o MSC, do inglês “multiplicative spectra correction” enquanto para o vetor **y** não foi realizado nenhum tratamento. Após essa etapa, utilizou-se o algoritmo de Kennard-Stone<sup>2</sup> para separar as amostras de calibração e validação, ou seja, após o tratamento de todas as amostras realizou-se a separação das amostras em dois grupos: amostras de calibração e validação. A figura 21 mostra os espectros das amostras do medicamento Nimesulida sem nenhum tratamento e com o tratamento MSC.



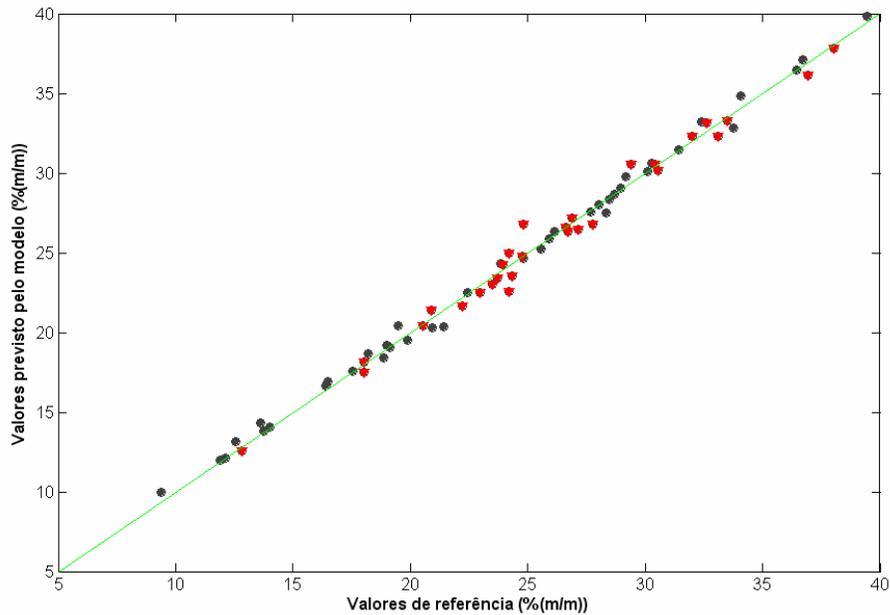
**Figura 21:** A) Espectros NIR do medicamento Nimesulida sem tratamento;  
B) Espectros NIR do medicamento Nimesulida com MSC.

Para o modelo construído foram empregadas 6 variáveis latentes. Isso ocorre porque esse número apresenta o menor valor da raiz quadrada da soma do quadrado dos resíduos (RMSECV) que não leva a sobre-ajuste ("overfit") e explica 100% da variância de  $\mathbf{X}$  e 99,93% de  $\mathbf{y}$ . A figura 22 demonstra o gráfico dos valores de RMSECV versus o número de variáveis latentes na qual foi utilizado para a escolha do número de variáveis latentes.



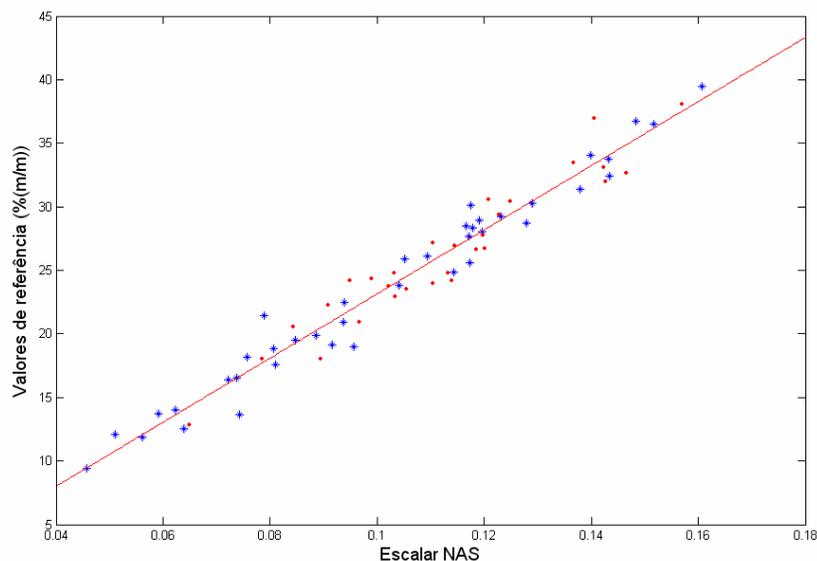
**Figura 22:** Gráfico dos resultados obtidos pelo método de validação cruzada utilizado para a escolha do melhor número de variáveis latentes.

Assim, foi construído um modelo de calibração tanto na sua forma pseudo-univariada utilizando o NAS como multivariada e validado através do cálculo de figuras de mérito. Nas figuras 23, 24, 25 e 26 a seguir estão mostrados os valores de referência contra os valores previstos pelos modelos multivariado e pseudo-univariado e os seus respectivos erros. Tanto no modelo multivariado como no pseudo-univariado. Nota-se uma boa concordância entre os valores de referência e os valores previstos pelos modelos. Isso ocorre porque nas figuras 23 e 24 os pontos graficados nessas figuras, correspondentes a relação entre os valores obtidos pelo modelo e os valores de referência, estão bem sobrepostos nas retas ajustadas pelo método dos mínimos quadrados.



**Figura 23:** Gráfico dos valores de referência contra os valores previstos do medicamento Nimesulida em (%) m/m estimado pelo o modelo PLS.  
(•) Calibração, (▲) Validação e (–) Reta ajustada.

Outra maneira de avaliar o ajuste do modelo é através do gráfico dos valores escalares do sinal analítico líquido-NAS, determinados pela norma do vetor de sinal analítico líquido em função da concentração da espécie de interesse. Desta forma utilizou-se o modelo pseudo-univariado através da equação 18 no qual encontrou-se a seguinte equação:  $\hat{y} = (252,52)nas - 2,084$



**Figura 24:** Representação do modelo na sua forma pseudo-univariada.

(\* ) Calibração, (•) Validação e (–) Reta ajustada.

Os resíduos referentes a esses ajustes, mostrados nas figuras 25 e 26, são menores que 3% tanto para as amostras de calibração quanto para validação. O cálculo dos resíduos, mostrado nestas figuras, foi feito através da equação 57 abaixo:

$$R_{\text{resíduos}} = \left( \frac{y_{\text{ref}} - y_{\text{pred}}}{y_{\text{ref}}} \right) \times 100 \quad (57)$$

Onde:

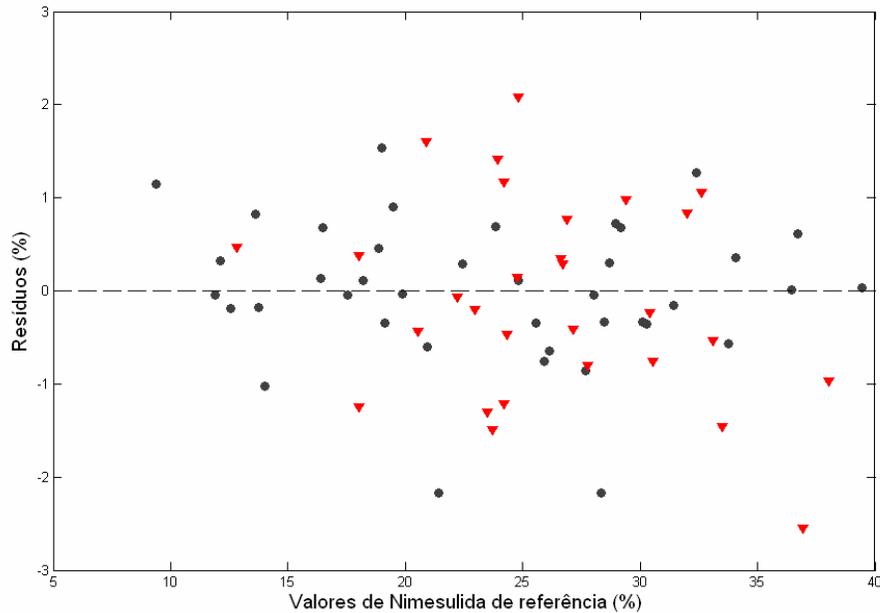
$R_{\text{resíduos}}$  : Resíduos

$y_{\text{ref}}$  : Vetor referência que contém a concentração do analito de interesse.

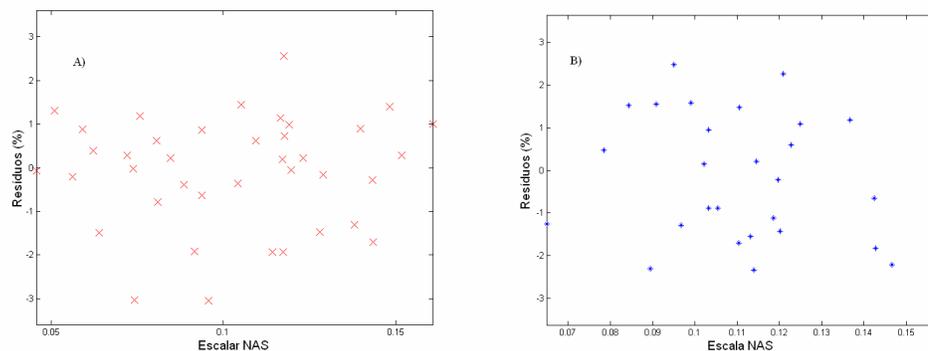
$y_{\text{pred}}$ : Vetor que contém os valores previstos da concentração dada pelo modelo.

Considerando a farmacopéia USP<sup>3</sup>, pode-se considerar esse modelo como aceitável, pois a maioria das amostras de previsão possuem erros relativos menores que 5% e a farmacopéia aceita até  $\pm 10$  (%) de diferença do valor tabelado para a maioria dos medicamentos. Não foi encontrado nessa farmacopéia o erro relativo tolerável para o medicamento Nimesulida. Sendo assim, considerou o erro máximo aceitável para a maioria dos medicamentos que é igual a  $\pm 10$  (%) do valor comercial.

Outro fato notório que deve-se levar em conta é que a indústria na qual foi realizado o procedimento experimental considera o erro aceitável para o medicamento Nimesulida igual a 5%. Já que as amostras apresentaram erros inferiores à 5% pode-se considerar o modelo como satisfatório.



**Figura 25:** Resíduos entre os valores reais e previstos pelo modelo PLS. (●) Calibração e (▲) Validação.



**Figura 26:** Resíduos entre os valores reais e os previstos pelo modelo pseudo-univariado. A) (X) Calibração e B) (\*) Validação.

A tabela 1 apresenta os resultados para os parâmetros de figuras de mérito. Pode-se verificar que o uso do NAS para o cálculo de algumas figuras de mérito foi de essencial importância para validação dos modelos apresentados.

**Tabela 1:** Resultados dos parâmetros de figuras de mérito para o modelo PLS.

Exatidão <sup>a</sup>	RMSECV <sup>d</sup>	0,984
	RMSEC <sup>d</sup>	0,653
	RMSEP <sup>d</sup>	1,122
	R <sup>2</sup> <sub>calibração</sub>	0,993
	R <sup>2</sup> <sub>validação</sub>	0,988
Seletividade média		0,0056
Sensibilidade <sup>b</sup>		0,0036
Sensibilidade analítica <sup>a</sup>		4,9
Inverso da sensibilidade Analítica <sup>b</sup>		0,2
Ajuste NAS	Coefficiente angular	252,52 ± 18,11
	Coefficiente linear	-2,08 ± 0,02
	Correlação	0,986
Ajuste PLS	Coefficiente angular	0,99 ± 0,01
	Coefficiente linear	0,23 ± 0,25
	Correlação	0,996
Erro sistemático de acordo com a ASTM	Bias	0,03
	Desvio Padrão	0,66
	Graus de liberdade	49
	t <sub>bias</sub>	0,31
	t <sub>crítico</sub>	2,10
Limite de quantificação <sup>a</sup>		2,03
Limite de detecção <sup>a</sup>		0,61
Razão sinal ruído <sup>c</sup>		181,11

<sup>a</sup>valores em % (m/m) de Nimesulida, <sup>b</sup>valor em % (m/m)<sup>-1</sup> de Nimesulida e <sup>c</sup>valor máximo. <sup>d</sup>Raiz quadrada do erro médio quadrático (RMSE) de validação cruzada (RMSECV), calibração (RMSEC), previsão (RMSEP), respectivamente onde  $n$  é o número de amostras,  $y_i$  é o vetor referência que contém a concentração do analito de interesse e  $\hat{y}_i$  é o vetor que contém os valores previstos da concentração dada pelo modelo.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

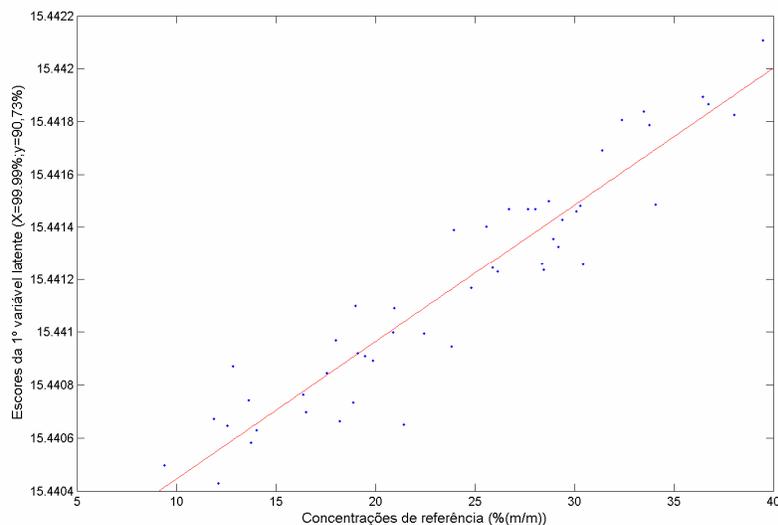
No que se refere à exatidão do modelo, essa figura de mérito é representada pelo coeficiente de correlação da calibração ( $R_{cal}$ ), coeficiente de correlação da validação ( $R_{val}$ ), raiz quadrada do erro quadrático médio da calibração (RMSEC) e da previsão (RMSEP). Dessa forma, pode-se dizer que o modelo pode ser considerado exato, pois possui valores de correlação tanto na calibração como na validação de 0,99. Além disso, possui valores de RMSEP e RMSEC, respectivamente 1,051 % e 0,655%. No que diz respeito à relação sinal/ruído, foi encontrado o valor de 181,11, o qual significa que o escalar NAS está aproximadamente 181 vezes maior que o desvio-padrão da flutuação do sinal instrumental. Esse valor mostrou-se satisfatório, porém, quanto maior essa razão, melhor será o valor do sinal obtido em relação ao ruído.

Considerando os valores de limite de detecção e quantificação, pode-se inferir que o modelo NIR construído só pode detectar valores iguais ou superiores a 0,61 % (m/m) de Nimesulida, enquanto para quantificação, o modelo não consegue determinar valores menores que 2,03% (m/m) de Nimesulida. Como a concentração de Nimesulida varia de 10,38-39,47% (m/m), pode-se considerar o modelo eficaz para detectar e quantificar esse princípio ativo. Isso ocorre porque os limites, tanto de detecção como de quantificação encontrado, possuem valores médios menores do que as porcentagens de Nimesulida presentes nas amostras.

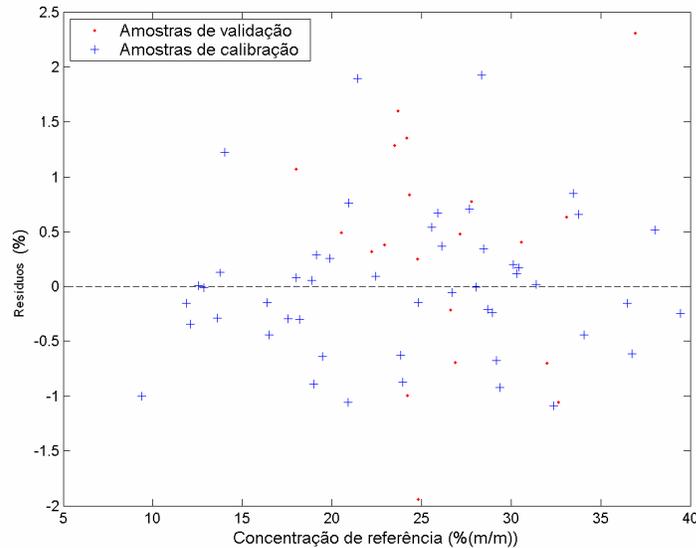
O modelo apresentou um pequeno valor numérico de sensibilidade ( $0,0036(m/m)^{-1}$ ), que em parte é decorrência do pré-processamento utilizado. O valor para o inverso da sensibilidade analítica, apresentado na tabela 1, pode ser interpretado de forma mais clara, por sua relação direta com a concentração. Segundo esse valor, o modelo é capaz de distinguir entre amostras com diferença de concentração da ordem de 0,2% (m/m). Já no que se refere à seletividade do modelo, o valor encontrado foi de 0,0056. Isso significa que, em média, menos que 1% da informação do analito contido no espectro da amostra não é ortogonal aos interferentes. Deve-se ressaltar, que esse valor de seletividade não se refere à seletividade no seu significado físico, que é em geral empregado em Química Analítica. Ele indica que cerca de 1 % do sinal é retirado na etapa de aniquilação de posto realizado durante o cálculo do NAS.

O ajuste do modelo foi avaliado com base nas figuras 23 e 24. Analisando a tabela 1, pode-se inferir que tanto o ajuste para o modelo PLS quanto para o modelo pseudo-univariada possui uma grande correlação com os dados de referência, pois ambos possuem correlação aproximadamente 0,99 entre os dados de referência e os dados calculados pelo modelo.

A medida quantitativa de linearidade em modelos de calibração multivariada não é simples, ou mesmo possível. Qualitativamente o gráfico dos resíduos das amostras de calibração e validação pode indicar se os dados seguem ou não o comportamento linear<sup>4</sup>. A distribuição aleatória desses resíduos é um indicativo de comportamento linear. Analisando as figuras 27 e 28, observa-se uma tendência a um comportamento linear na figura 27 e erros com comportamento aleatório na figura 28 para as amostras de calibração como de previsão. Logo, pode-se concluir que o modelo não possui falta de linearidade.



**Figura 27:** Escore da 1ª variável latente contra as concentrações da calibração.



**Figura 28:** Gráfico de resíduos contra as concentrações de referência.

Em relação aos erros sistemáticos, segundo as normas da American Society for Testing and Materials (ASTM), há uma sugestão de realização de um teste-t para avaliar se o “bias” incluso no modelo é significativo. Como o valor do  $t_{\text{bias}}$  calculado (0,31) é menor que o valor de  $t_{\text{crítico}}$  com  $n-1$  graus de liberdade (2,10) pode-se considerar que o bias incluído no modelo é insignificante. Logo, o modelo não apresenta erros sistemáticos, razão pela qual se pode dizer que a construção do modelo apresentado, da fase experimental até o cálculo das figuras de mérito, foi realizada de forma adequada.

Ao que refere-se as amostras comerciais de três lotes diferentes obtiveram-se resíduos (erros relativos) calculados através da equação 35 com valores iguais à 3 %, 5,5 % e 6,7 %. Não foi realizada uma análise de novos lotes ou mesmo de mais amostras do mesmo lote porque a empresa na qual realizou-se essas análises estava mudando os locais de produção e de análise do medicamento Nimesulida. Esse fato impossibilitou a realização de novas análises, porém os resultados obtidos podem ser considerados satisfatórios, pois de acordo com USP<sup>3</sup> os erros encontrados foram menores que 10%.

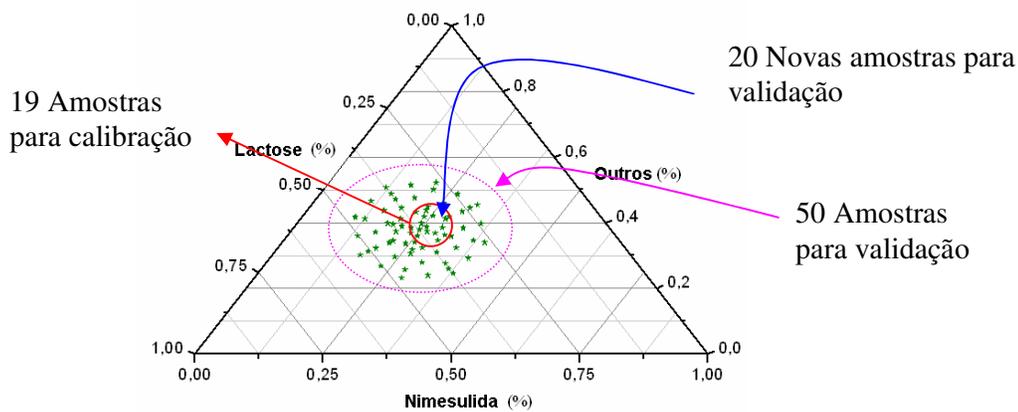
# **CAPÍTULO 6**

**CONTROLE DE QUALIDADE DO FÁRMACO  
NIMESULIDA EM MEDICAMENTO ATRAVÉS DE  
CARTAS DE CONTROLE MULTIVARIADA E DE  
ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO**

A segunda parte desta dissertação refere-se ao controle de qualidade em fármacos através da utilização de cartas de controle multivariadas. A vantagem do método é fundamentada no monitoramento simultâneo da divisão espectral em analito (NAS), excipientes (interferentes) e resíduos. Dessa forma, temos dois casos: amostras que estão dentro ou fora de controle. As amostras que possuem seus valores de NAS, interferente e resíduos dentro dos limites estabelecidos pelas amostras em controle são consideradas dentro de controle. Então, qualquer amostra que possuir pelo menos um valor fora do limite de qualquer um dos três parâmetros (NAS, interferente e resíduos) será denominada fora de controle<sup>1</sup>.

## 6-Parte Experimental

Na segunda etapa da dissertação foram utilizadas 113 amostras sintéticas, nas quais utilizou-se das mesmas 69 amostras mostradas da figura 19. Dessas 69 amostras (figura 29), 19 possuem concentração dentro da faixa de 22,33-27,29% (m/m) de Nimesulida e foram escolhidas para a calibração, enquanto que as restantes, ou seja, as outras 50 amostras foram escolhidas para validação (figura 29). Deve-se ressaltar, que essas 50 amostras não possuem concentração do analito dentro da faixa de concentração das amostras de calibração. Também foram preparadas mais 20 amostras (figura 29) para a validação contendo o princípio ativo (Nimesulida) dentro da faixa 22,33-27,29% (m/m) com o mesmo excipiente das amostras da figura 19, totalizando desta maneira 89 amostras.



**Figura 29:** Diagrama ternário para escolha das amostras de calibração e validação

Foram também utilizadas 10 amostras para construção do espaço dos interferentes. Essas amostras foram feitas somente com os excipientes do medicamento Nimesulida, ou seja, não foram acrescentadas o princípio ativo. Essas amostras foram feitas dividindo os excipientes em três partes. Uma contendo lactose, outra contendo povidona e uma terceira parte denominada “outros” no qual é formado por celulose, lauril sulfato de sódio, croscarmelose sódica e estereato de magnésio.

Dessa forma, misturou-se esses constituintes da mesma forma das amostras anteriores variando a concentração de lactose, excipiente que contém a maior quantidade no medicamento comercial, entre 10-80% (m/m). Essas amostras compõem a matriz de dados denominada placebo.

As 14 amostras restantes que faltam para completar as 113 amostras sintéticas são amostras feitas para a validação das quais 9 foram adicionadas NaHCO<sub>3</sub> e 5 amido, além de todos excipientes que contém a formulação da Nimesulida. Essas 14 amostras foram feitas ajustando a concentração do princípio ativo quando se acrescentava esses novos excipientes, ou seja, calculava-se a nova concentração de Nimesulida após o acréscimo desses novos excipientes.

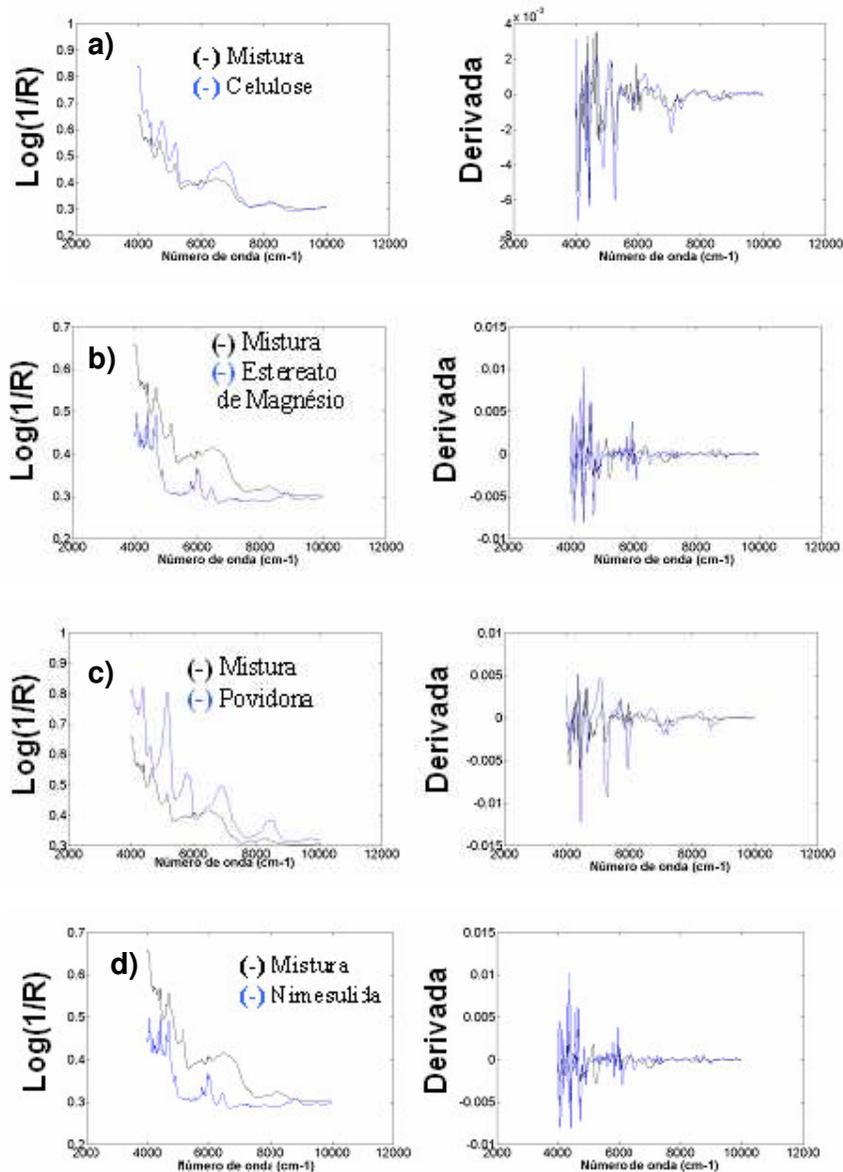
A tabela 2 mostra um resumo de todas as amostras utilizadas nessa segunda parte desta dissertação.

**Tabela 2** – Descrição de todas amostras usadas para construção e validação das cartas de controle.

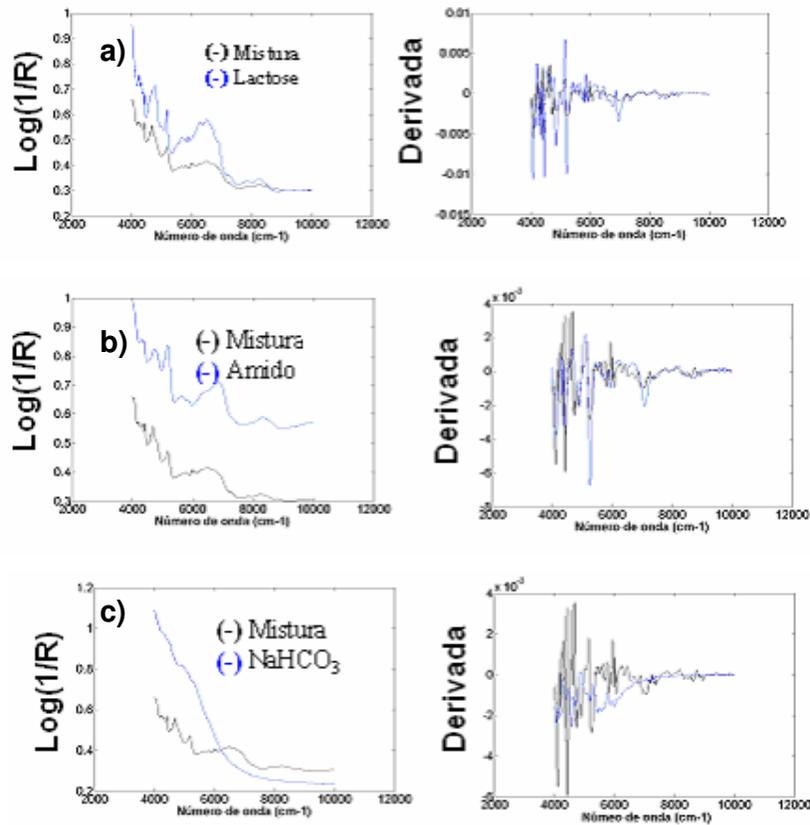
<b>Descrição</b>	<b>Usada para</b>
- 10 amostras para o placebo (Amostras com todos os constituintes exceto o analito)	- Construção do espaço dos interferentes
- 19 amostras sob controle	- Construção dos limites para a carta NAS
- 70 amostras para validação (50 amostras com concentrações fora e 20 com concentrações dentro do limite estipulado pela carta NAS)	- Validação dos limites para a carta NAS
- 14 amostras fora de controle ( Amostras com presença de Amido e Lactose. Constituintes ausentes nas amostras sob controle)	- Validação dos limites para a carta Interferente
- <b>Total</b> : 113 amostras	

## 6-1 Espectros no infravermelho próximo

Os espectros de todos os constituintes puros presentes no medicamento Nimesulida junto com a mistura sintética estão mostrados nas figuras 30 e 31.



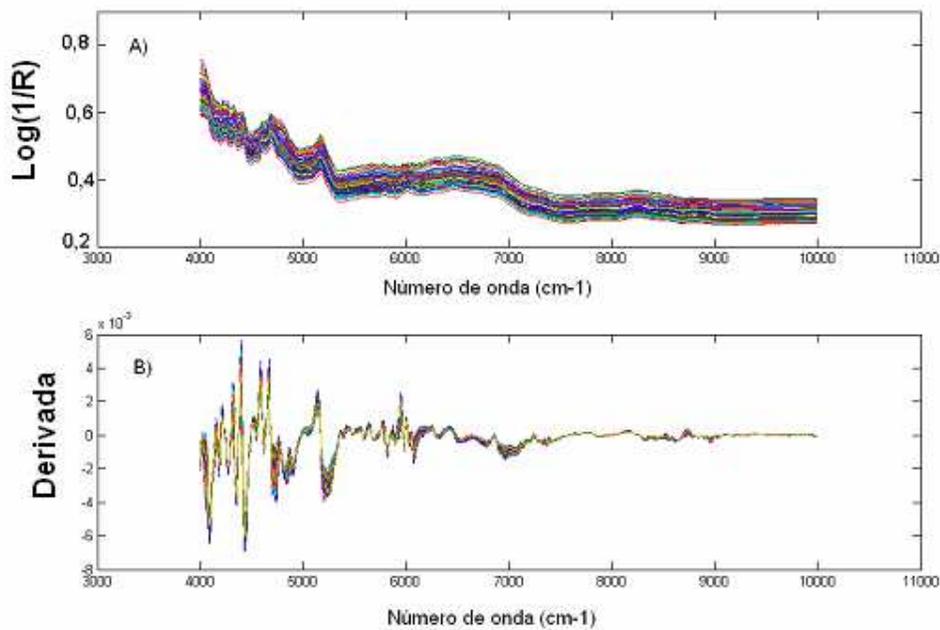
**Figura 30:** Espectros da mistura (-) e dos constituintes puros (-) sem tratamento e com a primeira derivada :a) Celulose, b) Estereato de Magnésio, c) Povidona e d) Nimesulida.



**Figura 31:** Espectros da mistura (-) e dos constituintes puros(-) sem tratamento e com a primeira derivada: a) Lactose, b) Amido e c)  $\text{NaHCO}_3$ .

## 6-2 Construção das Cartas de Controle Multivariada

Para a construção das cartas de controle, primeiramente, realizou-se o pré-processamento dos dados. Para os conjuntos de espectros de calibração, validação e placebo (povidona, croscarmelose sódica, lactose, estereato de magnésio e lauril sulfato), utilizou-se a primeira derivada.



**Figura 32:** A) Espectros NIR do medicamento Nimesulida sem pré-processamento;  
B) Espectros NIR do medicamento Nimesulida com 1<sup>o</sup> derivada.

Após isso, realizou-se uma Análise de Componentes Principais (PCA) da matriz placebo para a construção do espaço dos interferentes no qual foram utilizadas duas componentes principais. Através do PCA consegue-se explicar 99,98% da variância total dos dados. Construído o espaço dos interferentes, dividiram-se os espectros dos dados em três constituintes: NAS, interferente e resíduos, com o intuito de realizar a construção das cartas de controle. Para isso, utilizou-se das equações 40, 44 e 45. A figura 33 apresenta os sinais referentes ao NAS, interferentes e resíduos da amostra 15.

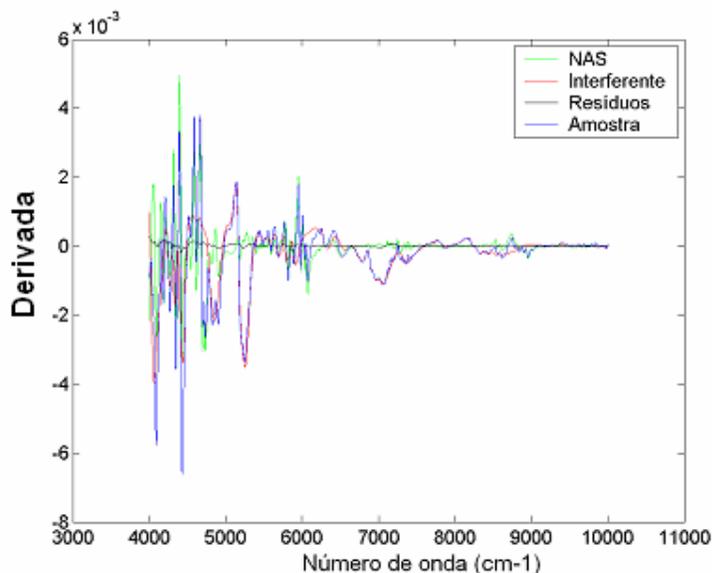
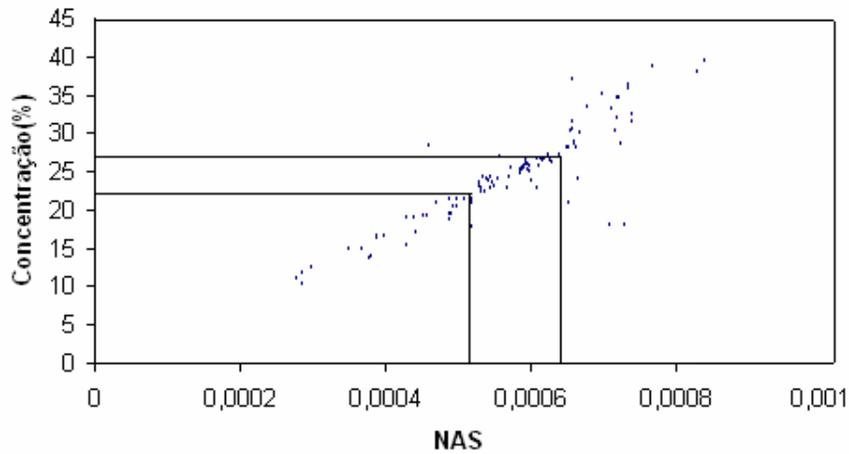


Figura 33: Divisão espectral em NAS, Interferente e Resíduos.

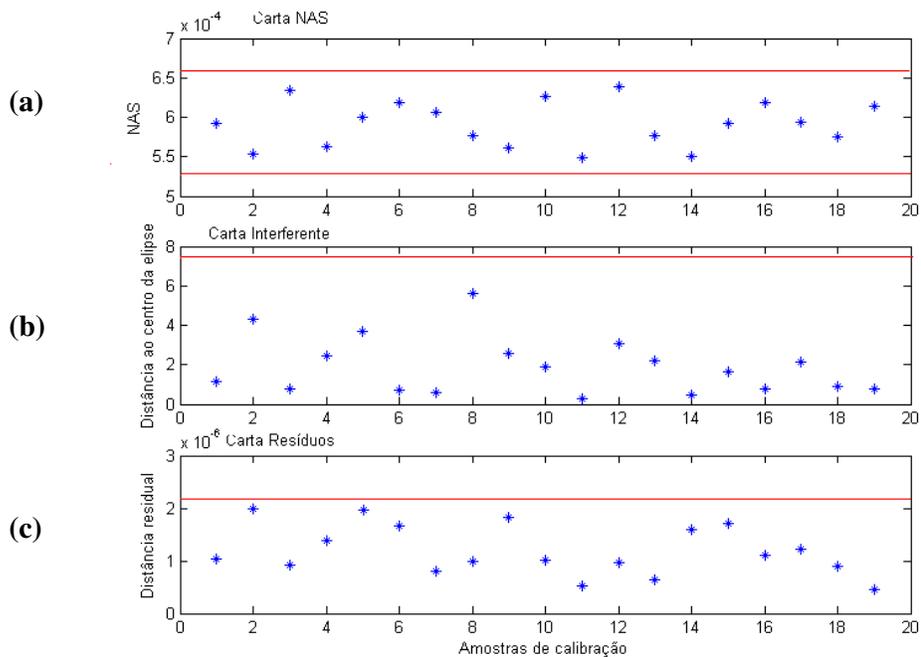
Após a etapa da divisão espectral de todas as amostras, foi necessário calcular os limites de confiança para cada carta. Para isso, utilizaram-se somente as amostras sob controle -19 amostras escolhidas para a calibração. O limite de confiança para a carta resíduo foi feito de acordo com a equação 47, obtendo-se o valor de  $2,38 \times 10^{-6}$ , considerando 95% de confiança. Para a carta interferente, o limite de confiança foi calculado conforme a equação 43, obtendo-se o valor de 7,5.

Utilizando-se da carta NAS, foram determinados os limites de confiança inferior e superior de concentração de Nimesulida nas amostras, estipulados em 22,35% (m/m) e 27,13% (m/m), respectivamente, levando-se em conta a farmacopéia USP<sup>2</sup> que considera a variação de  $\pm 10\%$  da concentração do princípio ativo no medicamento comercial, que é de 24,83%(m/m). Amostras que possuem concentração entre esses limites estão dentro do controle em relação a essa variável. Pensando em um modelo pseudo-univariado, podem-se entender esses limites como valores de NAS. Assim sendo, esses limites são  $5,26 \times 10^{-4}$  e  $6,53 \times 10^{-4}$ . Entretanto, nesse caso, esses limites estão em valores de NAS, e não mais de concentração. A figura 34 mostra a relação entre a concentração de Nimesulida e o valor de NAS. As linhas horizontais e verticais que passam pelos pontos dos gráficos demarcam os limites superior e inferior.



**Figura 34:** Valor do NAS vs Nimesulida (% (m/m)) para as amostras da figura 19.

A figura 35 abaixo mostra as cartas de controle desenvolvidas. Como todas as 19 amostras estão dentro de controle (amostras que foram escolhidas para a calibração), os valores de NAS, interferente e resíduos desses medicamentos estão dentro dos limites estabelecidos pelas três cartas representados pelas linhas vermelhas nas figuras 35, 36 e 40.

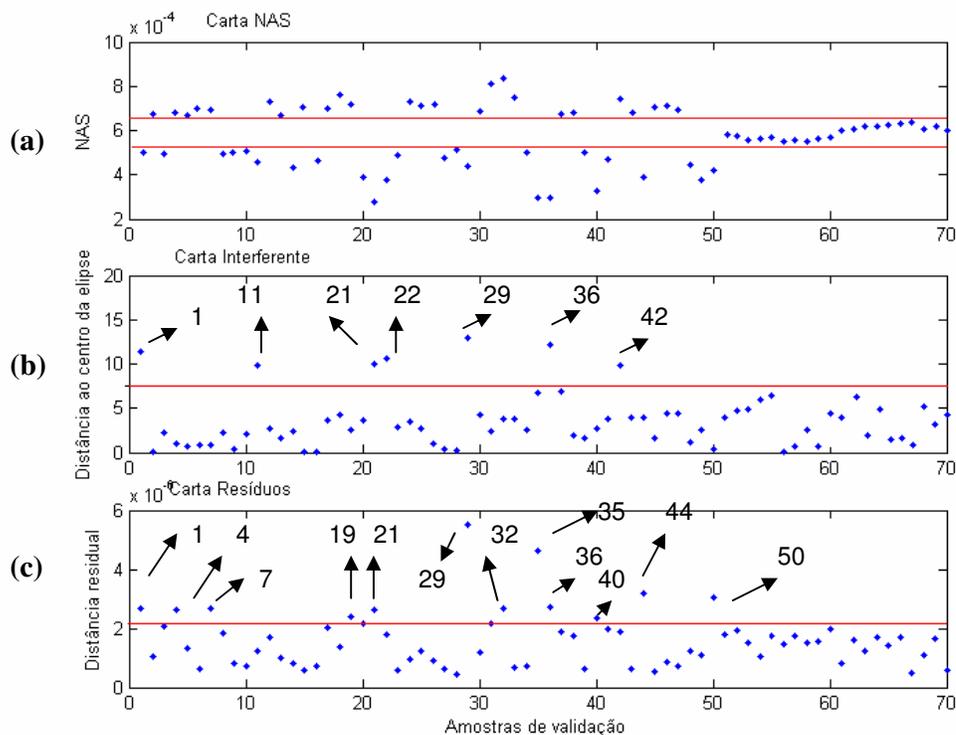


**Figura 35:** Cartas de controle a) NAS; b) Interferentes e c) Resíduos.

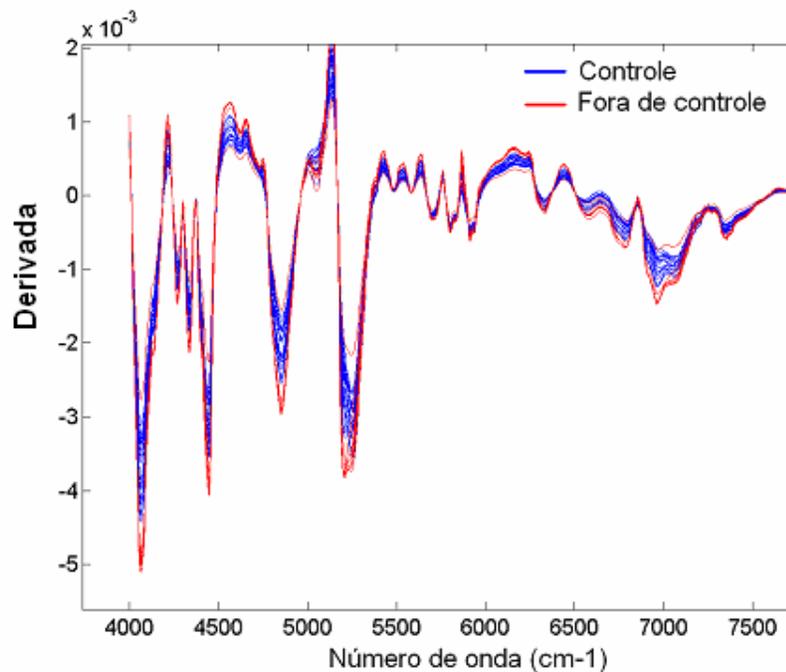
### 6-3 Validação das Cartas de Controle Multivariada

Outra etapa do trabalho foi avaliar nessas cartas novas amostras que estavam dentro e fora de controle. Na figura 36 utilizaram-se 70 amostras para essa tarefa, das quais as 50 primeiras são as amostras de validação mostradas na figura 29. Nota-se, pela figura 36, que todas essas amostras estão fora dos limites da carta NAS, o que está de acordo com o esperado, pois essas amostras foram preparadas em laboratório com concentrações diferentes da faixa de calibração (22,35% a 27,13%). Dessa forma, verificou-se com essa metodologia que 25 amostras estão fora do limite superior e 25, fora do limite inferior, o que está de acordo com o planejamento realizado.

Considerando a carta interferente para essas amostras, pode-se inferir que as amostras 1, 11, 21, 22, 29, 36 e 42 estão fora de controle, o que pode ser explicado pelo fato de que as amostras 1, 11, 21, 22, 29 e 36 possuem uma concentração de lactose maior do que a faixa de concentração de lactose do modelo, enquanto a amostra 42 possui uma concentração de lactose menor do que a do modelo construído.



**Figura 36:** Cartas de controle com novas amostras a) NAS; b) Interferentes e c) Resíduos.

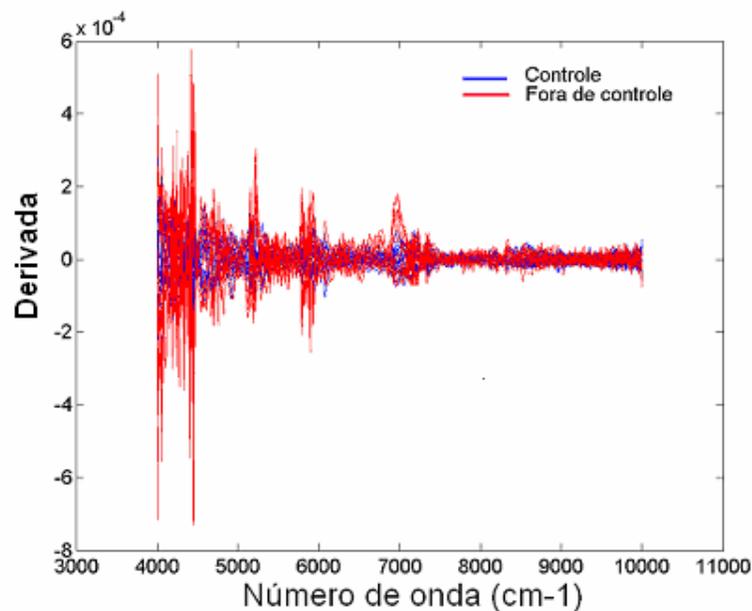


**Figura 37:** Vetores interferentes para as amostras sob controle (-) e para amostras 1, 11, 21, 22, 29, 36 e 42 que estão fora de controle(-).

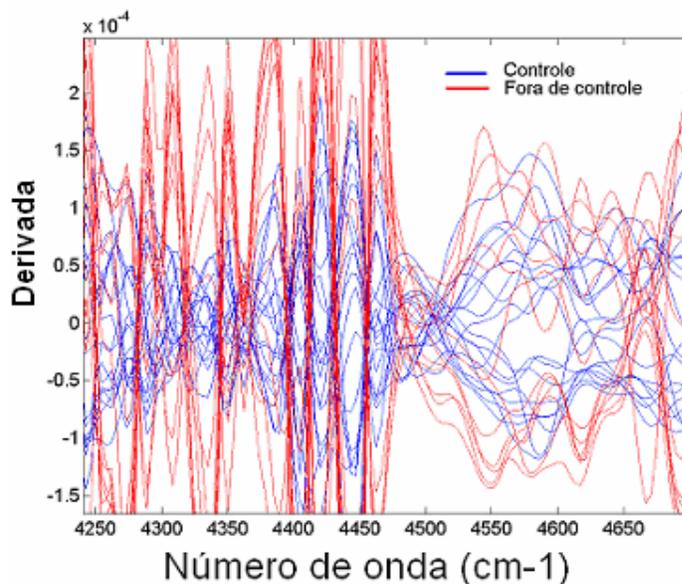
Na figura 37 é mostrado os vetores relativos aos interferentes das amostras sob controle e das amostras 1, 11, 21, 22, 29, 36 e 42 fora de controle, assinaladas na parte “b” da figura 36. Nota-se nesse gráfico que há várias regiões espectrais com intensidades diferentes entre as amostras dentro e fora de controle. Não foi mostrado o espectro dos interferentes das novas amostras que estão dentro de controle, pois dificultaria a visualização do espectro das amostras que estão fora de controle.

No que refere-se à carta resíduos, as amostras 1, 4, 7, 19, 21, 29, 32, 35, 36, 40, 44 e 50, assinaladas na parte “c” da figura 36 estão fora de controle. Isso significa que essas amostras não foram modeladas adequadamente, pois a diferença entre os valores previstos e os de referência estão acima do resíduo padrão estipulado no modelo. Investigando essas amostras, descobriu-se que essas não foram submetidas ao moinho. Dessa forma, pode-se considerar a hipótese de que essas amostras não são homogêneas e que possuem diferentes tamanhos de partículas entre os constituintes presentes. Essa hipótese pode ser considerada uma vez que, antes de as amostras serem submetidas ao NIR, os constituintes dos medicamentos tiveram de ser

homogeneizados por meio de vórtex, moinho e almofariz, pois os tamanhos das partículas dessas diversas substâncias são diferentes. O vórtex foi usado com o objetivo de misturar as partículas dos diferentes constituintes, enquanto o almofariz e o moinho foram utilizados com o intuito de triturar os sólidos para que essas partículas tenham aproximadamente a mesma granulometria. As duas tarefas juntas têm como objetivo homogeneizar as misturas sintéticas e, conseqüentemente, reduzir o espalhamento da luz no momento da aquisição dos espectros. Logo, essas amostras podem não ter ficado homogêneas, o que explica os valores fora do limite na carta resíduo. Isso pode ser visualizado nas figuras 38 e 39 pela diferença espectral entre essas amostras e aquelas que estão dentro de controle.



**Figura 38:** Vetores resíduos para as amostras de calibração (-) e das amostras 1, 4, 7, 19, 21, 29, 32, 36, 40, 44 e 50 (-).



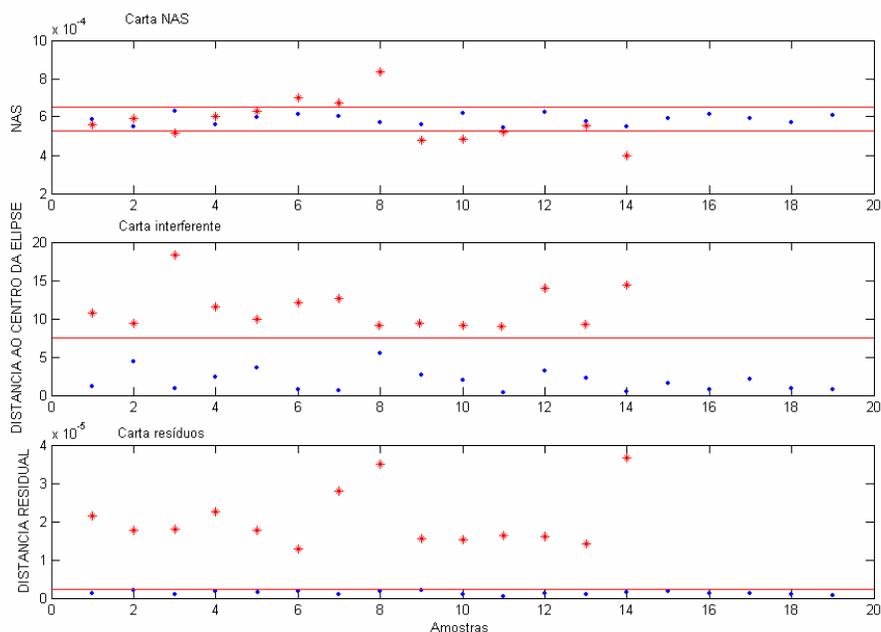
**Figura 39:** Vetores resíduos para região 4250-4650( $\text{cm}^{-1}$ ) da figura 38.

Outro fato notado é que essas amostras foram feitas com massa total de 2g e em dias diferente daqueles das amostras de controle. Todas amostras em controle foram pesadas com massa total 4g. Devido à falta de reagentes foi necessário diminuir a massa total para 2g. Isso pode também ter influenciado nos altos valores de Q obtidos. As amostras 51 a 70 são classificadas como sob controle, pois, em todas as cartas, essas possuem valores dentro dos limites estipulados. Isso está de acordo com o esperado, porque essas amostras foram preparadas com os mesmos interferentes das amostras em controle e com concentração dentro da faixa de  $\pm 10\%$  do valor estipulado para o medicamento comercial. Além do mais, pode-se dizer que essas amostras são homogêneas, pois não estão fora de controle nas cartas interferente e resíduo. Além disso, todas essas amostras foram preparadas da mesma forma que as amostras de controle, ou seja, no mesmo dia, quantidade, tempo de agitação e homogeneização.

As últimas amostras da figura 36-amostras 51 à 70 são as 20 amostras que foram feitas dentro da faixa de concentração do princípio ativo (Nimesulida) estabelecido na calibração e com o mesmo excipiente das amostras da figura 19. Além disso, da mesma forma e modo de preparo das amostras de calibração. Nota-se pela figura 36 que em todas as cartas essas foram assinaladas dentro dos limites estabelecidos. Logo, pode-se concluir que

essas amostras estão dentro de controle o que também está de acordo com o planejamento realizado mostrado na figura 29.

Em seguida, com o intuito de estudar o espaço dos interferentes, foi feita uma nova validação envolvendo mais 14 amostras de constituição idêntica à citada acima, exceto pelo fato de que em quatro delas foi adicionado amido e nas outras dez,  $\text{NaHCO}_3$ , diferença essa que tem como objetivo confirmar que o modelo, do ponto de vista do interferente, funciona adequadamente. Os resultados obtidos estão mostrados na figura 40, a seguir.



**Figura 40:** Cartas de controle: amostras sob controle (●) e amostras de validação (\*).

a) NAS; b) Interferentes e c) Resíduos.

Pode-se notar pela figura 40 que todas as amostras de validação estão fora do limite estipulado nas cartas interferentes e resíduos, portanto estão fora de controle, uma vez que essas tiveram acréscimo de amido e de  $\text{NaHCO}_3$ , os quais não estão presentes nas amostras de controle. As quatro primeiras amostras tiveram acréscimo de amido. Nota-se pela carta NAS que essas amostras estão dentro dos limites estabelecidos pela calibração. Isso está de acordo com o planejamento realizado, pois todas essas amostras foram feitas com concentrações dentro da faixa de controle. Considerando ainda a carta NAS, as próximas dez amostras tiveram acréscimo de  $\text{NaHCO}_3$ . Pela análise

de referência, sete amostras foram feitas com concentração fora dos limites estipulados pela carta NAS e somente três, dentro do limite estipulado pela carta NAS. Isso também está de acordo com os resultados obtidos, o que confirma mais uma vez a eficiência do modelo.

Analisando as figuras 41 e 42 abaixo, nota-se que os vetores interferentes das amostras que possuem amido e  $\text{NaHCO}_3$  são diferentes dos vetores interferentes das amostras sob controle. Isso refere-se ao fato da existência de novas espécies (Amido e  $\text{NaHCO}_3$ ) nessas amostras o que reflete na mudança dos espectros de infravermelho próximo para essas amostras. Dessa forma, o modelo consegue identificar que essas amostras possuem algum constituinte diferente nesse espaço.

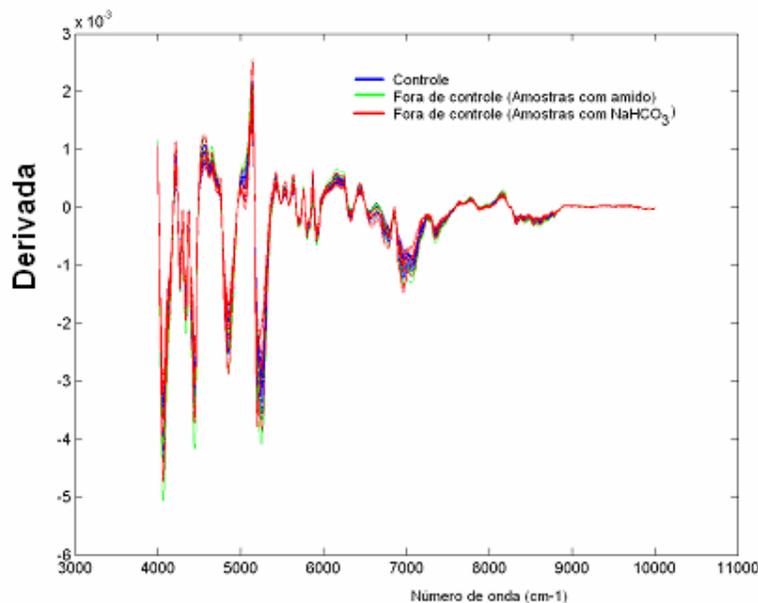


Figura 41: Vetores interferentes.

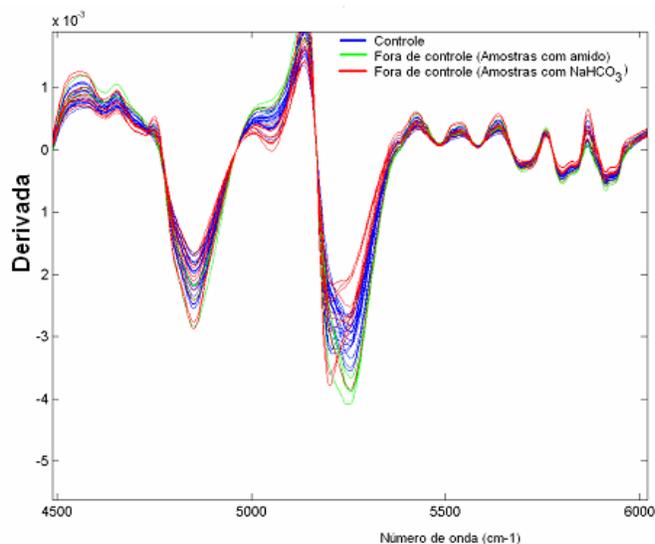


Figura 42: Vetores interferentes para região 4000-6000 ( $\text{cm}^{-1}$ ) da figura 41.

No que se refere aos resíduos para essas novas amostras submetidas ao modelo, nota-se pela figura 43, a presença de novos picos, principalmente, na região entre 5000-6000 ( $\text{cm}^{-1}$ ). Isso pode ser explicado por vários fatores, entre os quais estão as características de absorção de compostos não modelados pelo espaço dos interferentes, ou parte do vetor NAS, os quais ocasionam, em muitos casos, mudanças de pico característicos na região do infravermelho<sup>28</sup>. A causa exata desses picos é difícil de ser identificada, mas o resultado obtido é eficaz para diferenciar amostras que possuem constituintes e modo de preparo diferentes daqueles do modelo construído.

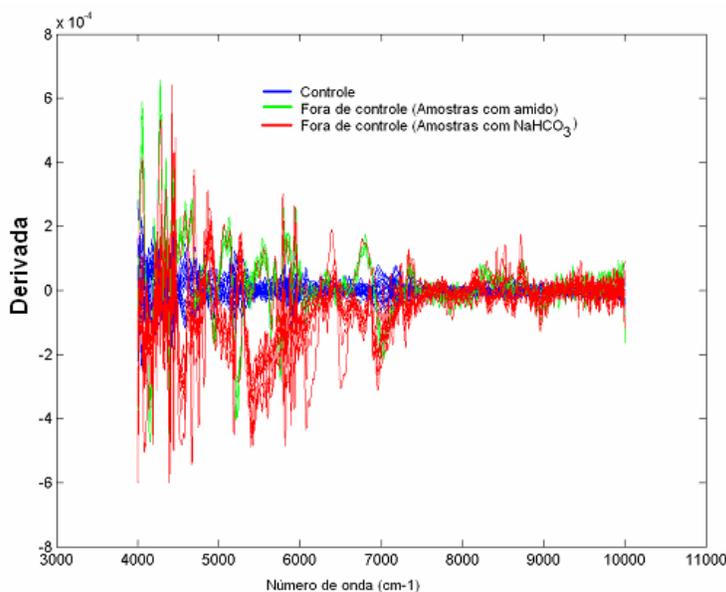


Figura 43: Vetores resíduos.

# **CAPÍTULO 7**

## **CONCLUSÕES**

A determinação de Nimesulida em medicamento genérico foi realizada através de espectroscopia no infravermelho próximo, regressão por mínimos quadrados parciais (PLS), juntamente com a determinação do NAS (Sinal Analítico Líquido) para validar o modelo construído através do cálculo das figuras de mérito e das cartas de controle multivariada. Em relação ao cálculo das figuras de mérito através do sinal analítico líquido, conseguiu-se obter, para os valores de limites de detecção, como de quantificação, resultados inferiores ao que se refere a concentração de Nimesulida presente nos medicamentos o que torna o modelo eficaz em relação a esse requisito. O modelo pode também ser considerado exato e eficiente tanto para a calibração como para a previsão, pois possuem valores de RMSEC, RMSECV, RMSEP,  $R_{\text{validação}}$ ,  $R_{\text{calibração}}$ , ajuste (PLS e NAS) e erros segundo a ASTM aceitáveis para quantificação de Nimesulida, ou seja, alta correlação e baixos erros. O modelo apresentou um alto valor da relação sinal/ruído o que torna o modelo eficaz para discernir o que é ruído e o que é sinal das amostras do sinal instrumental. Também o modelo consegue diferenciar valores muito baixos de concentração, entre os medicamentos, representado pelo seu reduzido valor do inverso da sensibilidade analítica (0,2%).

Pode-se afirmar também que, com a utilização do NAS, consegue-se construir modelos de calibração multivariada em uma forma pseudo-univariada, conforme mostrado nas figuras 16 e 24, tornando os modelos de calibração multivariada mais práticos em termos de visualização, pois se consegue relacionar diretamente o valor da concentração da Nimesulida com o sinal instrumental, de forma semelhante ao que se realiza em modelos de calibração univariada.

A utilização do NAS, no desenvolvimento desta dissertação, foi de fundamental importância para validação e caracterização de modelos de calibração multivariada, conforme mostrado na tabela 1. Em relação aos resíduos para as amostras comerciais encontrou-se valores de 3 %, 5,5 % e 6,7 % que estão dentro dos valores aceitos pela farmacopéia USP. Então, pode-se inferir que a metodologia descrita nessa dissertação é válida para determinação de Nimesulida em formulações farmacêuticas.

No que se refere às cartas de controle de qualidade, foi possível demonstrar um novo método multivariado para controle de qualidade na qual o

espectro do medicamento é dividido em três contribuições independentes: NAS (analito), interferente e resíduo. Então foi desenvolvido três cartas de controle para cada contribuição e calculado os limites de controle para cada carta. Através dessas cartas de controle multivariadas foi possível identificar as amostras que estavam dentro e fora de controle. Obteve-se 64 amostras fora de controle e 20 amostras dentro de controle de acordo com o planejamento experimental realizado. Sendo assim, foi possível identificar, de forma qualitativa, as amostras que estavam dentro e fora de controle.

Logo, a dissertação desenvolvida sugere um novo método analítico que pode ser usado para controle de qualidade em fármacos, principalmente, no que se refere ao controle da homogeneidade dos constituintes.

# **CAPÍTULO 8**

## **REFERÊNCIAS BIBLIOGRÁFICAS**

## 8 REFERÊNCIAS BIBLIOGRÁFICAS

### Capítulo 1- Introdução

1. PLUGGE, W.; VLIES, C. **Journal of Pharmaceutical and Biomedical Analysis**. 1996, 14, p. 891-898.
2. LANGKILDE, F. W.; SVANTESSON, A. **Journal of Pharmaceutical and Biomedical Analysis**. 1995, 13, p. 409-414.
3. HAN, S. M.; FAULKNER, P. G. **Journal of Pharmaceutical and Biomedical Analysis**. 1996, 14, p. 1681-1689.
4. FRAKE, P.; GREENHALGH, D.; GRIERSON, S. M.; HEMPENSTALL, J. M.; RUDD, D. R. **International Journal of Pharmaceutics**. 1997, 151, p. 75-80.
5. DUONG, N.-H.; ARRATIA, P.; Muzzio, F. J.; LANGE, A.; TIMMERMANS, J.; REYNOLDS, S. **Drug Development and Industrial Pharmacy**. 2003, 29, p. 679-687.
6. GOTTFRIES, J.; DEPUI, H.; FRANSSON, M.; JONGENELEN, M.; JOSEFSON, M.; LANGKILDE, F. W.; WITTE, D. T. **Journal of Pharmaceutical and Medical Analysis**. 1996, 14, p. 1495-1503.
7. PLUGGE, W.; VAN der Vlies, C. **Journal of Pharmaceutical and Biomedical Analysis**. 1993, 11, p. 435-442.
8. MILLER, C.E.; WILLIAMS, P.; NORRIS, K. **In Near-Infrared Technology**, St. Paul, Minnesota, USA, 2001, p. 19.
9. PASQUINI, C. **Journal of the Brazilian Chemical Society**. 2003, 14, p. 198-219.
10. DONOSO, M.; KILDSIG, D. O.; GHALY, E. S. **Pharmaceutical Development and Technology**. 2003, 8, p. 357-366.
11. FÉVOTTE, G., Calas, J.; PUEL, F.; HOFF, C. **International Journal of Pharmaceutics**. 2004, 273, p. 159-169.
- 12- <http://www.rxlist.com/cgi/generic/cetiriz.htm>, acessada em Agosto 2006.
- 13- <http://www.gtamart.com/mart/products/chemical/mgi2.htm>, acessada em Abril 2007.
- 14- Berntsson, O. **Characterization and Application of Near Infrared Reflection Spectroscopy for Quantitative Process Analysis of Powder Mixtures**. Kungliga Tekniska Högskolan (KTH), Stockholm, Sweden, 2001. Tese (doutorado).

### Capítulo 2- Métodos Quimiométricos

- 1- KALIVAS, J.H. **Analytical Letters**. 2005, 38, p. 2259-2279.
- 2- BRERETON, R. G.; **Analyst**. 2000, 125, p. 2125.
- 3- MILLER, J. N.; MILLER, J. C.; **Statistics and Chemometrics for Analytical Chemistry**, 4ª ed., Prentice Hall, United Kingdom, 2000.
- 4- SAVITSKY, A; GOLAY, M. J. E. **Analytical Chemistry**. 1964, 36, p. 1627-1639
- 5- MALEKI, M. R.; MOUAZEN, A. M.; RAMON, H. **Biosystems Engineering**, 2007, 96, p. 427-433.

- 6- RAWLINGS, O. J. **Applied Regression Analysis-A Research Tool**, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, Califórnia, 1988.
- 7- MARTENS, H.; NAES, T. **Multivariate Calibration**. Wiley, New York, 1989.
- 8- GELADI, P.; KOWALSKI, B. **Partial Least-Squares Regression: A Tutorial**. **Analytical Chimica Acta**. 1986, 185, p. 1.
- 9- ANDERSSON, M.; FOLESTAD, S.; GOTTFRIES, J., Johansson, M. O.; JOSEFSON, M.; WAHLUND, K. G. **Analytical Chemistry**, 2000, 72, p. 2099-2108.
- 10- ESCANDAR, G.M.; DAMIANI, P. C; GOICOECHEA, H. C. **Microchemical Jornal**, 2006, 82, p. 29-42.
- 11- FOULK, S. **American Laboratory**, 2005, p. 37.
- 12- Otto, M. **Chemometrics**. Wiley, Weinheim, 1999.
- 13- VANDEGNISTE, B. G. M.; MASSART, D. L.; BUYDENS, L. M. C.; Jong, S.; LEWI, P. J.; VERBEQUE, J.S. **Handbook of Chemometrics and Qualimetrics - Part B**, Elsevier, Amsterdam, 1998.
- 14- **Annual Book of ASTM Standards Predctices for Infrared, Multivariate, Quantitative Analysis**, ed. 1655, v. 03/06, ASTM International, West Conshohocken, Pennsylvania, USA, 2000.
- 15- PIEMA, J. A. F., L; DASZYKOWSKI, M.; WAH, F., MASSART, D. L. **Chemometrics and Intelligent Laboratory Systems**. 2003, 68, p. 17.
- 16- PELL, R. J., **Chemometrics and Intelligent Laboratory Systems**, 2000, 52, p. 87.
- 17- WALCZAK, B.; MASSART, D. L. **Chemometrics and Intelligent Laboratory Systems**. 1998, 41, p. 1.
- 18- BARTHUS, R.C. **Aplicação de Métodos Quimiométricos para Análises de Controle de Qualidade de Óleos Vegetais Utilizando Espectroscopia no Infravermelho e Raman**. 1999. Tese (mestrado).
- 19- LORBER , A., **Analytical Chemistry**,1986, 58, p. 1167.
- 20- BRAGA, J. W. B.; POPPI, R. J.; **Journal of Pharmaceutical Science**. 2004, 93, p. 2124.
- 21- JONG, S. **Simpls. Chemometrics and Intelligent Laboratory Systems**. 1993, 18, p. 251-263.
- 22- BRAGA, J. W. B.; POPPI, R. J.; **Química Nova**. 2004, 27, p. 1004.

### **Capítulo 3- Validação de Modelos de Calibração Multivariada através do Cálculo de Figuras de Mérito**

- 1- Eurachem/ Citac- Work Group. **Guide of Quality in Analytical Chemistry – An Aid to Accreditation**, 2002.
- 2- International Conference on Harmonization, ICH Harmonised Tripartite Guideline – **Validation of Analytical Procedures: Metodology**. 1997, 62, p. 27463.
- 3- International Conference on Harmonization, ICH Harmonised Tripartite Guideline – **Text on Validation of Analytical Procedures**.1995, 60, p. 11260.
- 4- Annual Book of ASTM Standards - **Practices for Infrared Multivariate Quantitative Analysis**, vol. 03/06, 2000.

- 5- Pharmaceutical Analytical Sciences Grup (PASG). **Guidelines for the Development and Validation of Near Infrared (NIR) Spectroscopic Methods**. <http://www.pasg.org.uk>.
- 6- Eurachem/ Citac- Work Group. **Quantifying Uncertainty in Analytical Measurement**. 2 ed., 2000.
- 7- MIRMOHSENI, A.; ABDOLLAHI, H.; ROSTAMIZADEH, K. **Analytica Chimica Acta**. 2007, 585, p. 179-184.
- 8- REN, M.; ARNOLD, M. A. **Analytical and Bioanalytical Chemistry**. 2007, 387, p. 879-888.
- 9- BLANCO M; CASTILLO, M; PEINADO, A. **Analytica Chimica Acta**. 2007, 581, p. 318-323.
- 10- VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. **Journal of the Brazilian Chemical Society**. 2007, 18, p. 259-266.
- 11- KAUFFMAN, J. F.; DELLIBOVI, M.; CUNNINGHAM, C. R. **Journal of Pharmaceutical and Biomedical Analysis**. 2007, 43, p. 39-48.
- 12- **Pure and Applied Chemistry**. 2006, 78, p. 633-661.
- 13- GABRIELSSON, J.; TRYGG, J. **Critical Reviews in Analytical Chemistry**, 2006, 36, p. 243-255.
- 14- BURNS, D. A., CIURCZAK, E. W. **Handbook of Near-Infrared Analysis**, Marcel Dekker, New York, 2001.
- 15- MOFFAT, A. C.; TRAFFORD, A. D.; JEE, R. D.; GRAHAM, P.; **Analyst** 2000, 125, p. 1341.
- 16- LAASONEN, M.; PULKKINEN, T. H.; SIMARD, C.; RASANEN, M.; VUORELA, H.; **Analytical Chemistry**, 2003, 75, p. 754.
- 17- LORBER, A., **Analytical Chemistry**, 1986, 58, p. 1167-1172.
- 18- FERRÉ, J.; BROWN, S. D.; RIUS, F. X.; **Journal of Chemometrics**. 2000, 15, p. 537.
- 19- RODRIGUEZ, L. C.; Campanã, A. M. G.; Linares, C. J.; Ceba, M. R. **Analytical Letters**, 1993, 26, p. 243.
- 20- PENA, A.M.; MANSILLA, A. E.; VALENZUELA, M. I. A; GOICOECHEA, H. C., OLIVIERI, A. C., **Analytica Chimica Acta**, 2002, 463, p. 75.
- 21- VESSMAN, J.; STEFAN, R. I.; STADEN, J. F., DANZER, K.; LINDNER, W.; BUMS, D. T.; FAJGELJ, A.; MULLER, H. **Pure and Applied Chemistry**, 2001, 73, p. 1381.

## CAPÍTULO 4 - Cartas de Controle Multivariadas

- 1- NIJHUIS, A.; JONG, S.; VANDEGINSTE, B. G. M.; **Chemometrics and Intelligent Laboratory Systems**, 1999, 47, p. 107-125.
- 2- . BARTHUS, R. C.; POPPI, R. J. **Spectroscopy Letters**, 2002, 35, p. 729-739.
- 3- NIJHUIS, A.; JONG, S.; VANDEGINSTE, B. G. M. **Chemometrics and Intelligent Laboratory Systems**, 1997, 38, p. 51-62.
- 4- SKIBSTED, E. T. S.; WESTERHUIS, J. A.; BOELEN, H. F.; SMILDE, A. K.; WITTE, D. T. **Applied Spectroscopy**. 2004, 58, p. 264-271.
- 5- RAMAKER, H. J.; SPRANG, N. M. **Statistical Batch Process Monitoring**. University of Amsterdam, Process Analysis & ChemometricsGroup, 2004. Tese (doutorado)

- 6- SEKULIC, S. S.; WARD, H. W.; BRANNAGAN, D. R.; STANLEY, E. D.; EVANS, C. L.; SCIAVOLINO, S. T.; HAILEY, P. A.; ALDRIDGE, P. K. **Analytical Chemistry**, 1996, 68, p. 509-513.
- 7- SKIBSTED, E. T. S.; BOELEN, H. F. M.; WESTERHUIS, J. A.; SMILDE A. K.; BROAD, N. W.; REES, D. R., Witte; D. T. **Analytical Chemistry**. 2005, 77, p. 7103 - 7114;
- 8- PLUGGE, W.; VLIES, C. **Journal of Pharmaceutical and Biomedical Analysis**. 1996, 14, p. 891-898.
- 9- Langkilde, F. W; Svantesson, A. **Journal of Pharmaceutical and Biomedical Analysis**. 1995, 13, p. 409- 414.
- 10- HAN, S. M; FAULKNER, P. G. **Journal of Pharmaceutical and Biomedical Analysis**. 1996, 14, p. 1681-1689.
- 11- FRAKE, P.; GREENHALGH, D.; GRIERSON, S. M.; HEMPENSTALL, J. M.; RUDD, D. R. **International Journal of Pharmaceutics**. 1997, 151, p. 75-80.
- 12- DUONG, N. H.; ARRATIA, P.; MUZZIO, F. J.; LANGE, A.; TIMMERMANS, J.; REYNOLDS, S. **Drug Development and Industrial Pharmacy**. 2003, 29, p. 679-687.
- 13- GOTTFRIES, J.; DEPUI, H.; FRANSSON, M.; JONGENELEN, M.; LANGKILDE, F. W.; WITTE, D. T. **Journal of Pharmaceutical and Biomedical Analysis**. 1996, 14, p. 1495-1503.
- 14- PLUGGE, W.; VLIES, C. **Journal of Pharmaceutical and Biomedical Analysis**. 1993, 11, p. 435-442.
- 15- LORBER, A. **Analytical Chemistry**. 1986, 58, p. 1167-72.
- 16- MARTENS, H.; NAES, T. **Multivariate Calibration**. John Wiley & Sons, Chichester, 1989.
- 17- WOLD, S.; ESBENSEN, K. H.; GELADI, P. **Chemometrics and Intelligent Laboratory Systems**. 1987, 2, p. 37-52.
- 18- LORBER, A.; FABER, N. M.; KOWALSKI, B. R. **Analytical Chemistry**. 1997, 69, p. 1620-1626.
- 19- MILLER, J. N.; MILLER, J. C. **Statistics and Chemometrics for Analytical Chemistry**. Prentice Hall, New York, 2000.
- 20- JACKSON, J. E.; MUDHOLKAR, G. S. **Technometrics**. 1979, 21, p. 341-49.

## **CAPÍTULO 5- DETERMINAÇÃO DE NIMESULIDA EM MEDICAMENTOS ATRAVÉS DO CÁLCULO DE FIGURAS DE MÉRITO E DE ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO**

- 1 – WISE, B. M.; GALLAGHER, N. B.; BRO, R. SHAVER, J. M.; WINDIG, W. **PLS Toolbox 3.5 For Use With Matlab**. Manson, WA, Eigenvector Research Inc. 2005.
- 2- KENNARD, R.W.; Stone, L. A., **Technometrics**, 1969, 11, p. 137–148.
- 3- **The United States Pharmacopoeia**. 25<sup>a</sup> ed. rev., U.S.P. Convention, Rockville, 2002.
- 4- GOICOECHEA, H. C.; OLIVIER, A. C.; **Chemometrics and Intelligent Laboratory Systems**. 2001, 56, p. 73.

## **CAPÍTULO 6- CONTROLE DE QUALIDADE DO FÁRMACO NIMESULIDA EM MEDICAMENTO ATRAVÉS DE CARTAS DE CONTROLE MULTIVARIADA E DE ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO**

1- SKIBSTED, E. T. S., BOELEN, H. F. M.; WESTERHUIS, J. A., A. K. SMILDE, N. W.; BROAD, D. R.; WITTE, D. T. **Analytical Chemistry**. 2005, 77, p. 7103 - 7114;

2- **The United States Pharmacopoeia**. 25<sup>a</sup> ed. rev., U.S.P. Convention: Rockville, 2002.

## **CAPÍTULO 7- CONCLUSÕES**