



Universidade Estadual de Campinas
Instituto de Química – Depto. Química Analítica

Tese de doutorado

*Aplicação de métodos de análise multivariada em espectrometria
de massas para a determinação de compostos orgânicos voláteis.*

Aluna: Cleidiane Gorete Zamprônio
Orientador: Ronei Jesus Poppi
Co-orientador: Marcos Nogueira Eberlin

Campinas, novembro 2002

UNICAMP
BIBLIOTECA CENTRAL

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

UNIDADE	BC
Nº CHAMADA	T17UNICAMP
	Z14a
V	EX
TOMBO BC/	53364
PROC.	124103
C ☐	D ☒
PREÇO	R\$ 11,88
DATA	24/04/03
Nº CPD	

CM00181089-6

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE QUÍMICA
UNICAMP**

Z14a

Zamprônio, Cleidiane Gorete

Aplicação de métodos de análise multivariada em espectrometria de massas para a determinação de compostos orgânicos voláteis / Cleidiane Gorete Zamprônio. -- Campinas, SP: [s.n], 2002.

Orientador: Ronei Jesus Poppi.

Co-orientador: Marcos Nogueira Eberlin.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Química.

1. Isômeros. 2. MIMS. 3. Quimiometria.
I. Poppi, Ronei Jesus. II. Eberlin, Marcos Nogueira. III. Universidade Estadual de Campinas. IV. Título.

Dedico esta tese a toda minha família e
principalmente ao meu marido Steve

Agradecimentos

- Ao Prof. Ronei J. Poppi e ao Prof. Marcos N. Eberlin, pela excelente orientação, apoio, paciência, amizade e por todo o conhecimento transmitido.
- Aos amigos do LAQQA, Eduardo, Marcelo Sena, Paulo Henrique, Silvio, Fernando, Alessandra, Marcelo Trevisan, Chiquinho, Jez, Waldomiro e Juliana, pela convivência agradável do dia a dia.
- Aos amigos do Laboratório Thomson, Regina, Lillian, Fábio, Eduardo, Renato, Adriana, Daniela e Aline, pela convivência agradável e colaboração.
- Aos amigos do LEEDS, Miyuki, Tânia, Rosângela, Karin, Perse, Arnaldo, Jequié, pela convivência agradável do dia a dia.
- Ao grupo de Análises de Processos e Quimiometria da Universidade de Amsterdam, pela sua hospitalidade durante o período dedicado para que, parte deste trabalho, fosse realizado.
- Ao Instituto de Química e aos seus funcionários
- À FAPESP pelo apoio financeiro (No. Proc. 98/11327-9).

Resumo

Neste trabalho foram aplicados métodos quimiométricos em dados de espectrometria de massas obtidos a partir de análise de compostos orgânicos voláteis (VOC). Diferentes experimentos foram realizados utilizando a amostragem direta através da simples vaporização da amostra a partir de soluções orgânicas e também utilizando a análise em fluxo junto à espectrometria de massas por introdução via membrana (MIMS) para análise de soluções aquosas.

Inicialmente dois experimentos foram realizados utilizando a técnica MIMS. No primeiro, a técnica FIA-MIMS foi utilizada para análise de amostras de água contaminadas com gasolina. No segundo experimento, a técnica MIMS foi utilizada para a análise de amostras de cerveja. Nos espectros de massas dos dois conjuntos de dados foi aplicada a análise exploratória, através da análise de componentes principais (PCA), que separou as amostras em grupos, de acordo com a variação da concentração dos compostos orgânicos voláteis detectados. Verificou-se o grande potencial desta técnica em análises ambientais ou de rotina para controle de qualidade.

Em uma segunda etapa, análises de amostras de misturas de isômeros de etilpiridina (2-etilpiridina, 3-etilpiridina e 4-etilpiridina) com espectros de massas muito similares, foram realizadas. Dados bidimensionais ou de primeira ordem foram obtidos a partir da análise seqüencial através de reação íon-molécula e dissociação induzida por colisão (CID) e dados tridimensionais ou de segunda ordem foram obtidos a partir da análise seqüencial por dissociação CID. Diferentes métodos de calibração multivariada foram aplicados e os resultados comparados. Os dados de reação íon-molécula apresentaram bons resultados através da calibração por mínimos quadrados parciais (PLS), com erros da ordem de 8 a 13% para as misturas contendo os três isômeros. Os dados de segunda ordem também apresentaram bons resultados, principalmente através do modelo de calibração por mínimos quadrados parciais multidimensionais (N-PLS), com erros da ordem de 11 a 15%. Os resultados a partir da análise seqüencial mostraram ser superiores aos obtidos pela simples ionização por elétrons (EI) a 70 eV, cujos erros foram entre 18 a 29%. Entretanto, verificou-se que a aplicação da calibração multivariada só foi possível depois de encontrado o melhor método de pré-processamento, porque espectros de massas apresentam ruído heteroscedástico, no qual varia de instrumento para instrumento. É possível dizer que a transformação logarítmica deve ser aplicada em casos onde o ruído heteroscedástico interfere no modelo de calibração.

Um último experimento foi realizado aplicando a técnica FIA-MIMS juntamente com a calibração multivariada para verificar a viabilidade de utilização destes métodos para a quantificação

de BTX (benzeno, tolueno e xilenos) em água. Bons resultados foram obtidos, comprovando que a espectrometria de massas pode ser uma ferramenta muito valiosa em análises de rotina e monitoramento em linha, no qual espectros de massas podem ser adquiridos rapidamente e quantificado pela calibração multivariada. Erros da ordem de 12 a 17% foram obtidos para amostras contendo misturas de BTX com concentrações variando de 300 a 1000 $\mu\text{g L}^{-1}$.

Abstract

In this work, chemometric methods have been applied to mass spectrometry data from volatile organic compounds (VOC). Different experiments were performed by direct sampling mass spectrometry analysis using simple vaporization of samples of organic solutions and also using flow analysis and membrane introduction mass spectrometry (MIMS) for the analysis of aqueous solutions.

Two experiments were performed using the MIMS technique. In the first, FIA-MIMS technique was used to analyze water samples contaminated by gasoline were analyzed. In the second experiment, MIMS technique was used to analyze beer samples. In the mass spectra from the both data set were applied the data exploration by principal component analysis (PCA), that separated the group of samples in agreement with the concentration variation of volatile organic compounds detected. It was verify the potential of this technique for environmental analysis or routine analysis for quality control.

In the second stage, an analysis of ethyl pyridine isomers mixture samples (2-ethylpyridine, 3-ethylpyridine e 4-ethylpyridine) that have very similar spectra was carried out. Two-way or first order data were obtained from sequential analysis by ion-molecule reaction and collision induced dissociation (CID) and three-way or second order data were obtained by sequential analysis dissociation by CID. Different multivariate calibration methods were applied and the results were compared. Ion-molecule reaction analysis showed good results by partial least squares (PLS) calibration and error between 8 and 13%, for the mixture of three isomers. The second order data also showed good results, mainly by the N-PLS model and error between 11 and 15%. Both results from the sequential analysis are excellent and showed to be better than the results from the simple ionization by 70 eV electron ionization (EI), which the error value are between 18 and 29%. However, it was found that the application of multivariate calibration only was possible after having found the optimum preprocessing method, because the mass spectra have heteroscedastic noise levels, which varies for each different instrument. It is possible to say that the log transform should be applied in cases where the heteroscedastic noise interferes with the calibration model.

In the last experiment, the FIA-MIMS technique and multivariate calibration where used to investigate the possibility of applying these methods for BTX (benzene, toluene e xylene) quantitation in water. Good results were obtained, confirming that mass spectrometry can be a valuable tool for

routine analysis and on-line monitoring, where the mass spectra can be obtained rapidly and quantified using multivariate calibration. Errors of between 12 and 17% were obtained for mixtures of BTX with a concentration range of 300 to 1000 $\mu\text{g L}^{-1}$.

CURRICULUM VITAE

Formação Acadêmica

Mestrado – Depto. de Química Analítica, Universidade Estadual de Campinas – UNICAMP

Título: *Determinação simultânea de ácidos fracos, em análise por injeção em fluxo, empregando calibração multivariada não linear, por treinamento de redes neurais artificiais.*

Concluído: 06/08/98 (data de defesa)

Graduação - Bacharelado (Complementação Tecnológica) e Ciências com Habilitação em Química, Universidade Estadual Paulista – UNESP – Araraquara – S.P.

Concluído: dezembro de 1993

Artigos Científicos

1. C.G. Zamprônio, R. Almeida, L.L. Rocha, R.J. Poppi, M.N. Eberlin, *Beer fingerprinting by membrane introduction mass spectrometry*, submetido.
2. C.G. Zamprônio, L.A.B. Moraes, M.N. Eberlin, R.J. Poppi, *Quantitation of isomeric ethyl pyridine mixtures by multivariate calibration applied to ion-molecule reaction/collision-induced dissociation triple-stage mass spectra*, submetido.
3. C.G. Zamprônio, S.P. Gurden, L.A.B. Moraes, M.N. Eberlin, A.K. Smilde, R.J. Poppi, *Direct sampling tandem mass spectrometry (MS/MS) and multiway calibration for isomer quantitation*, *Analyst*, 127 (2002) 1054-1060.
4. C.G. Zamprônio, J.J.R. Rohwedder, R.J. Poppi, *Artificial neural networks applied to potentiometric acid-base flow injection titrations*, *Chemometrics and Intelligent Laboratory Systems*, 62 (2002) 17-24.
5. R.M. Alberici, C.G. Zamprônio, R.J. Poppi, M.N. Eberlin, *Water solubilization of ethanol and BTEX from gasoline: on-line monitoring by membrane introduction mass spectrometry*, *Analyst*, 127 (2002) 230-234.
6. C.G. Zamprônio, L.A.B. Moraes, M.N. Eberlin, R.J. Poppi, *Multivariate curve resolution applied to MS/MS data obtained from isomeric mixtures*, *Analytica Chimica Acta*, 446 (2001) 495-502.
7. C.G. Zamprônio, R.J. Poppi, J.J.R. Rohwedder, *Development of a potentiometric flow cell with a stainless steel electrode for pH measurements. Determination of acid mixtures using flow injection analysis*, *Talanta*, 51 (2000) 1163-1169.
8. C.G. Zamprônio, M.R. Davolos, E.B. Stucchi, M. Jafelicci Jr., *Spherical particles of pure and manganese doped zinc oxide and zinc hydroxycarbonate*, *Materials Research Society Symposium Proceedings*, 372 (1995), 69-74.

Trabalhos apresentados em reuniões científicas

1. 11^o. Encontro Nacional de Química Analítica – UNICAMP, Campinas, 18 a 21/09/2001.
2. 24^a. Reunião Anual Sociedade Brasileira de Química, Poços de Caldas, 28 a 31/05/ 2001.
3. 24^a. Reunião Anual Sociedade Brasileira de Química, Poços de Caldas, 28 a 31/05/ 2001.
4. 7th International Conference on Chemometrics in Analytical Chemistry, Antuérpia, 16 a 20/10/2000.
5. 23^a. Reunião Anual Sociedade Brasileira de Química, Poços de Caldas, 23 a 26/05/2000.
6. X Encontro Nacional de Química Analítica, Santa Maria, 31/08 a 03/09/1999.

7. 1^o. Workshop Científico dos Pós-Graduandos da UNICAMP, Campinas, 01 a 03/09/1998.
8. 21^a. Reunião Anual Sociedade Brasileira de Química, Poços de Caldas, 24 a 27/05/1998.
9. VII International Conference on Flow Analysis, Águas de São Pedro, 25 a 28/08/1997.
10. IX Encontro Nacional de Química Analítica, São Carlos, 31/08 a 03/09/1997.
11. 20^a. Reunião Anual Sociedade Brasileira de Química, Poços de Caldas, 24 a 27/05/1997.
12. 16^a. Reunião Anual Sociedade Brasileira de Química, Caxambú, 25 a 29/05/1993.
13. 10^o. Encontro Regional de Química - 2^a. Jornada de iniciação científica e tecnológica em química. SBQ Regional Araraquara – Ribeirão Preto – São Carlos, Ribeirão Preto, 19 a 21/11/1992.
14. IV Congresso de Iniciação Científica da UNESP, Araçatuba, 27 a 29/08/1992.
15. 44^a. Reunião Anual da Sociedade Brasileira para o Progresso da Ciência, São Paulo, julho de 1991.

Experiência Acadêmica

1. Visita ao grupo de Process Analysis & Chemometrics da Universidade de Amsterdam (UVA) para pesquisa na área de análises quimiométricas de dados multidimensionais MS/MS.
2. Atuação como auxiliar didático na disciplina de Graduação QA-281 (Química Analítica Clássica Experimental), no Instituto de Química – UNICAMP – Campinas, no 1^o. semestre de 2000.
3. Iniciação científica: Depto. de Química Inorgânica – Instituto de Química – UNESP – Araraquara. *Precursores para novos materiais luminescentes*, de março de 91 a dezembro de 92.
4. Apresentação oral na sessão coordenada de Eletroquímica e Eletroanalítica I, 20^a Reunião Anual da SBQ, *Construção e avaliação de uma cela de fluxo potenciométrica empregando eletrodo de aço inox para a medida de pH*.
5. Apresentação oral na sessão coordenada de Química Analítica, 24^a. SBQ, *Determinação de isômeros a partir de espectros de massas tridimensionais através da calibração multivariada por N-PLS*.

Experiência Profissional

1. Atuação como química responsável no Instituto Campineiro de Análise de Solo e Adubo Ltda, de fevereiro de 1994 a novembro de 1995.
2. Estágio na Bristol Myers Squib Brasil S A, de março a novembro de 1993.

Cursos

1. Espectrometria de Absorção Atômica, Thermo Jarrell Ash Corporation – A concentrated program of study of principles and application, 02 de agosto de 1994.
2. Visual Basic 3.0, Centro de Computação da UNICAMP, 06 a 10 de outubro de 1997.
3. Matlab, Centro de Computação da UNICAMP, 22 a 26 de novembro de 2001.
4. Introduction to Multiway Analysis, Instituto de Química – UNICAMP, 18 a 20 de fevereiro de 2002.

Índice

Glossário	xvii
Considerações Preliminares	1
Capítulo 1 Espectrometria de massas	5
1. Introdução	5
2. Espectrômetro de massas	5
2.1. Ionização por elétrons	6
2.2. Analisador de massas quadrupolar	6
2.3. Detector	9
3. Espectrometria de massas seqüencial	9
3.1. Reação íon-molécula	10
3.2. Dissociação seqüencial	11
4. Técnica MIMS	11
Capítulo 2 Análise multivariada de dados	15
1. Introdução	15
2. Análise exploratória através da análise de componentes principais	15
3. Calibração multivariada	17
3.1. Calibração multivariada de dados de primeira ordem	18
3.1.1. PLS	19
3.2. Calibração multivariada de dados de segunda ordem	21
3.2.1. U-PLS	22
3.2.2. N-PLS	23
3.2.3. MCR	24
Capítulo 3 Análise exploratória de espectros de massas	27
1. Introdução	27
2. Análise de amostras de água contaminadas com gasolina, através da técnica FIA-MIMS.	27
2.1. Parte experimental	30
2.2. Resultados e discussão	32
2.2.1. Quantificação de BTEX e etanol a partir da análise por FIA-MIMS, através da calibração univariada	32
2.2.2. Análise exploratória a partir da análise por FIA-MIMS através da PCA	34
3. Análise de cervejas através da técnica MIMS	38
3.1. Parte experimental	40
3.2. Resultados e discussão	41
4. Conclusão	50
Capítulo 4 Pré-processamento do ruído heteroscedástico	53
1. Introdução	53
2. Parte experimental	55
3. Resultados e discussão	57

4. Conclusão	61
Capítulo 5 Calibração multivariada aplicada a dados de primeira ordem	63
1. Introdução	63
1.2. Reação íon-molécula e dissociação através da técnica CID	63
2. Parte experimental	65
3. Resultados e discussão	67
3.1. Primeiro experimento : EI a 70 eV	67
3.2. Segundo experimento: reação íon-molécula seguida de dissociação por CID	70
4. Conclusão	73
Capítulo 6 Calibração multivariada aplicada a dados de segunda ordem	75
1. Introdução	75
2. Parte experimental	76
2.1. Análise seqüencial de misturas de etilpiridina equantificação por U-PLS, N-PLS e MCR	76
2.2. Quantificação de misturas de BTX dissovídidos em água através do PLS e N-PLS	78
3. Resultados e discussão	80
3.1. Análise seqüencial de misturas de etilpiridina e quantificação por U-PLS, N-PLS e MCR	80
3.1.1. U-PLS e N-PLS	82
3.1.2. MCR	90
3.1.3. Métodos de calibração de segunda ordem	95
3.2. Quantificação de misturas de BTX dissolvidos em água através do PLS e N-PLS	95
4. Conclusão	100
Conclusões gerais e perspectivas futuras	103
Referências bibliográficas	107

Glossário

Estão definidas a seguir, apenas as siglas que aparecem com mais frequência no texto.

Sigla	Inglês	Português
VOC	Volatile Organic Compound	Composto Orgânico Volátil
FIA-MIMS	Flow Injection Analysis-Membrane Inlet Mass Spectrometry	Análise por Injeção em Fluxo-Espectrometria de Massas por Introdução via Membrana
CID	Collision Inducted Dissociation	Dissociação Induzida por Colisão
SIM	Selective-Ion Monitoring	Monitoramento de Íons Seleccionados
TIM	Total-Ion Monitoring	Monitoramento de Íons Totais
MS/MS	Tandem Mass Spectrometry	Espectrometria de Massas em Tandem
PCA	Principal Component Analysis	Análise de Componentes Principais
PLS	Partial Least Square	Mínimos Quadrados Parciais
N-PLS	Multilinear Partial Least Squares	Mínimos Quadrados Parciais Multilineares
U-PLS	Unfold- Partial Least Square	Mínimos Quadrados Parciais Desdobrados
MCR	Multivariate Curve Resolution	Resolução de Curva Multivariada

Considerações Preliminares

A espectrometria de massas é uma técnica que tem progredido extremamente rápido durante as últimas décadas. Esta técnica tem alcançado uma posição muito importante entre os métodos analíticos devido a algumas características como alta sensibilidade, baixo limite de detecção e diversidade em suas aplicações [1-4]. A evolução da espectrometria de massas na área de química analítica deve-se também ao desenvolvimento das técnicas de amostragem que ampliaram o número de suas aplicações.

Técnicas de amostragem direta como a espectrometria de massas por introdução via membrana ("membrane introduction mass spectrometry", MIMS) e técnicas MIMS hífenadas como, "trap-and-release", "Cryotrap" (CT-MIMS) e "headspace" [5-19] têm sido estudadas e utilizadas extensivamente na última década. A amostragem direta em espectrometria de massas significa que nenhum outro equipamento como cromatógrafo gasoso ou líquido estão acoplados. A grande vantagem destas técnicas é que permitem o monitoramento de compostos orgânicos voláteis com alta sensibilidade e rapidez, sem a preparação das amostras. Estas propriedades da técnica MIMS vêm contrastar com as técnicas GC/MS e LC/MS, em que é necessária a separação do analito através de uma coluna cromatográfica, que resulta em um aumento no tempo e complexidade da análise.

Em uma primeira etapa deste trabalho a técnica MIMS foi utilizada para a análise de amostras de água contaminadas com gasolina e também amostras de cerveja que foram analisadas diretamente sem nenhuma separação. Nestes dados foi aplicada a análise exploratória para a separação das amostras de acordo com a concentração dos compostos orgânicos voláteis detectados. A aplicação da análise multivariada é baseada em métodos matemáticos e estatísticos, ou mais comumente conhecidos como métodos quimiométricos, em que é possível utilizar todo o espectro de massas adquirido para descrever os diferentes grupos de amostras [20-22].

A amostragem direta, entretanto, apresenta como desvantagem a baixa seletividade [23], pois vários analitos podem passar através da membrana quase simultaneamente e, dependendo da amostra, pode ser obtido um espectro de massas bastante complexo. O maior problema neste caso é a presença de analitos com picos coincidentes. Algumas soluções vêm sendo propostas, como a análise por espectrometria de massas seqüencial ("tandem mass spectrometry", MS/MS) [23] que utiliza a técnica de dissociação induzida por colisão ("collision-induced dissociation", CID) e também

análise seqüencial por reação íon-molécula seguida por CID [23]. Estas duas soluções propostas têm por propriedade aumentar a seletividade da análise. A utilização destas técnicas pode tornar-se ainda mais poderosa se acompanhada pelo tratamento de dados através da análise multivariada. Para comprovar estas afirmações, foram realizados diferentes experimentos com amostras contendo compostos orgânicos voláteis. Foram utilizadas a amostragem direta sem membrana para amostras em meio orgânico e a amostragem direta através da técnica MIMS para amostras contendo compostos orgânicos voláteis solubilizados em água. A análise seqüencial foi aplicada utilizando um espectrômetro de massas pentaquadrupolar, obtendo-se conjuntos de dados bidimensionais (primeira ordem) e tridimensionais (segunda ordem).

Também um maior enfoque foi dado em relação à etapa de pré-processamento dos dados de espectrometria de massas, que é uma das etapas mais importantes quando se aplica a calibração multivariada para a quantificação das amostras.

Depois de pesquisas e consultas na literatura, juntamente com a colaboração do grupo de pesquisa do Prof. Age K. Smilde da Universidade de Amsterdam, foi possível interpretar melhor a distribuição de ruído presente nos dados de espectrometria de massas. Verificou-se a presença de ruído heteroscedástico, sendo necessário a transformação deste para ruído homoscedástico através da aplicação da função logarítmica. Através desta transformação foi possível aplicar a calibração multivariada e conseguir baixos erros de previsão de novas amostras, que não foram utilizadas no modelo de calibração.

Buscou-se, desta forma, um melhor entendimento da análise por espectrometria de massas utilizando todo o espectro de massas para a quantificação de misturas de compostos orgânicos voláteis e não somente um íon seletivo. Isto só foi possível com o auxílio da calibração multivariada, que tem por propriedade utilizar mais variáveis e conseqüentemente obter mais informação. Acredita-se que depois deste trabalho tem-se um caminho bastante promissor para análise de misturas complexas, como isômeros de posição ou outras aplicações que no futuro possam ser encontradas para a espectrometria de massas utilizando a amostragem direta juntamente com a quimiometria.

Este trabalho será apresentado em seis capítulos. No primeiro capítulo serão descritas a técnica de espectrometria de massas e especificamente o espectrômetro de massas utilizado neste trabalho. No segundo capítulo será discutida a análise multivariada, onde será descrito o método de análise de componentes principais ("principal component analysis", PCA) e os métodos de

calibração multivariada aplicados para a quantificação de amostras denominados de mínimos quadrados parciais ("partial least squares", PLS), mínimos quadrados parciais desdobrados ("unfold partial least squares", U-PLS), método dos mínimos quadrados parciais multilineares ("multilinear partial least squares", N-PLS) e o método de resolução de curvas multivariada ("multivariate curve resolution", MCR) com restrições. No capítulo três serão apresentados os resultados de duas aplicações de MIMS com amostras reais (água contaminada com gasolina e amostras de cerveja) a partir da separação destas amostras através da PCA. No capítulo quatro discutiremos o problema do ruído heteroscedástico presente em dados de espectrometria de massas e a necessidade do pré-processamento. Nos capítulos cinco e seis serão mostrados e comparados os resultados obtidos através da calibração multivariada de dados bidimensionais e tridimensionais a partir de amostras analisadas pela simples ionização e através da análise seqüencial.

Capítulo 1

Espectrometria de massas

1. Introdução

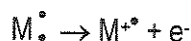
Amostras ambientais como água, solo e ar contaminados com compostos orgânicos voláteis e semivoláteis, ou amostras biológicas como soro biológico, urina, etc [5-19], são alguns exemplos comuns de aplicações da espectrometria de massas que vem crescendo muito nos últimos anos. A espectrometria de massas é um importante instrumento em geologia, como também no descobrimento de drogas. É utilizada freqüentemente por diversos ramos da ciência desde da arqueologia à zoologia, assim como em muitos campos da engenharia. Também nas últimas décadas foram desenvolvidas técnicas de desorção e ionização, que permitem a caracterização de moléculas de massas altas como proteínas e ácidos nucleicos. O estudo de substâncias biológicas e novos descobrimentos sobre a vida tem dominado parte da literatura de espectrometria de massas no início deste século 21, dentro de projetos como o genoma e proteoma [24].

A espectrometria de massas é um instrumento muito importante no estudo de substâncias, porque fornece informações detalhadas sobre a composição e estrutura das mesmas, a partir de uma quantidade de amostra muito menor que em outras técnicas analíticas. A única desvantagem encontrada na técnica é que a amostra é destruída depois de analisada. Entretanto, a quantidade de amostra necessária é muito pequena, conseqüentemente a perda não é significativa.

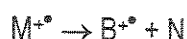
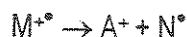
A espectrometria de massas tem por propriedade determinar uma substância através da separação e detecção de íons na fase gasosa, de acordo com a massa em função do número de cargas (m/z) destes íons [25-26]. Os dados são registrados em um espectro de massas que fornece a abundância relativa dos íons versus os valores de m/z . A unidade de massa utilizada em espectrômetros de massas é a unidade de massa atômica unificada (u) também conhecida por Dalton (Da).

2. Espectrômetro de massas

A espectrometria de massas utilizada na análise de compostos orgânicos está relacionada normalmente com a análise de íons positivos, derivados de precursores neutros. A simples ionização de uma molécula é mostrada abaixo:



O íon $M^{+\bullet}$ é conhecido como íon molecular que é positivo radicalar, o qual também pode se dissociar formando um novo produto ou ainda uma partícula neutra como:



Todos estes íons formados por diferentes processos podem ser selecionados, resultando em um espectro de massas bastante complexo dependendo da molécula ou mistura.

Existem diferentes tipos de espectrômetros de massas que são diferenciados de acordo com os métodos de ionização e os analisadores de massas utilizados. As formas mais comuns em relação aos métodos de ionização são: ionização por elétrons (EI), ionização química (CI), ionização por "electrospray" (ESI), ionização e desorção por laser em matriz assistida (MALDI) e bombardeamento por átomos rápidos (FAB) [25-26]. Neste trabalho, em todos os experimentos realizados foi aplicada a ionização por EI, um dos primeiros métodos de ionização utilizado em toda a história da espectrometria de massas.

2.1. Ionização por elétrons

Na fonte de ionização o vapor de amostra a pressão reduzida ($< 10^{-4}$ torr) colide com um fluxo de elétrons e é ionizado. Este fluxo de elétrons por sua vez é produzido a partir do aquecimento de um filamento de tungstênio ou rênio, através de um potencial aplicado. Na ionização por EI, a informação mais importante sobre a massa molecular (u) é o íon formado pela remoção de um elétron da molécula neutra (chamado de íon molecular). A energia de ionização normalmente aplicada é por volta de 70 eV [25,26].

A pressão na fonte de ionização é mantida baixa para prevenir colisões interpartículas, danos no filamento e descarga de corrente elétrica devido à alta diferença de potencial que é utilizada.

2.2. Analisador de massas quadrupolar

Os analisadores de massas são responsáveis pela seleção dos íons e existem vários tipos: setor magnético ("magnetic sector analyser"), analisador de massas quadrupolar ("quadrupole

analyser”), íon “traps” quadrupolar, analisador de tempo de voo (“time-of-flight analyser”, TOF), e analisador de ressonância ciclotrônica por transformada de Fourier (“Fourier transform-ion cyclotron resonance, FT-ICR”) [25-26].

Neste trabalho foi utilizado o analisador ou filtro de massas quadrupolar de íons, que consiste em quatro barras cilíndricas paralelas de aço inox. Nestas barras, uma voltagem d.c. de componente U , e uma radiofrequência (r.f.) de componente $V \cos \omega t$ são aplicadas entre as barras adjacentes (barras no eixo x ou barras no eixo y). Os cilindros opostos são conectados eletricamente, Figura 1.

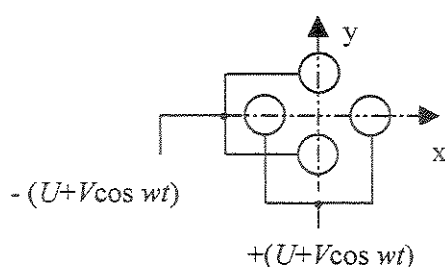


Figura 1 Esquema de um analisador quadrupolar [27].

Os íons formados na fonte de ionização são transferidos para os quadrupolos (região entre os cilindros), com pequena aceleração, tipicamente de 10 a 20 eV, e oscilam nas direções x e y em decorrência do campo elétrico formado.

Duas constantes a e q , podem ser definidas nas equações 1 e 2:

$$a = \frac{8zU}{mr_0^2 \omega^2} \quad (1)$$

$$q = \frac{4zV}{mr_0^2 \omega^2} \quad (2)$$

Nas equações r_0 é a distância do centro do quadrupolo até as barras, z é a carga do elétron, m é a massa do íon, U é a voltagem d.c. aplicada, V é a voltagem r.f. aplicada e ω é a frequência da voltagem r.f. Para certos valores de a e q as oscilações de alguns íons são estáveis, mas para outros as oscilações são instáveis. O diagrama de estabilidade, também conhecido como diagrama de Mathieu, Figura 2, mostra os valores de a contra q , em que estas condições são aplicadas.

Íons de massa m_1 , que estão dentro da região de oscilação estável, conseguem manter-se em uma trajetória estável dentro do quadrupolo até alcançar o detector. Íons de massa m_2 apresentam uma trajetória instável e são perdidos, chocando-se com as barras do quadrupolo. Espectros de massas são obtidos variando U e V , enquanto a razão U/V é mantida constante.

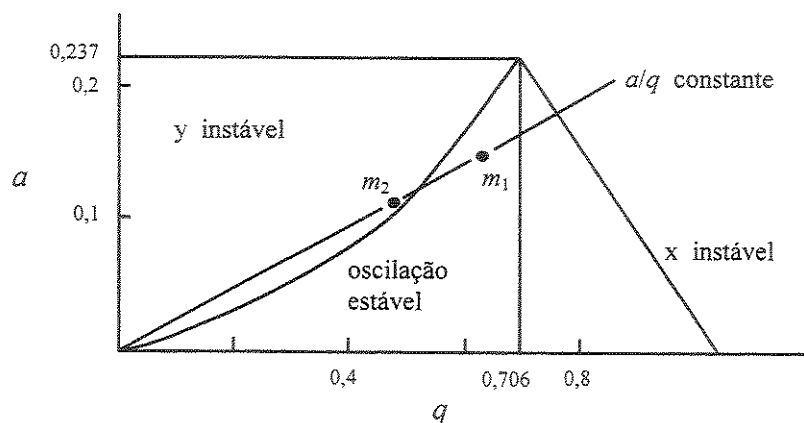


Figura 2. Diagrama de estabilidade para um analisador quadrupolar [27].

Teoricamente, a resolução de massas no quadrupolo pode ser aumentada trabalhando com valores de a/q próximos do ápice do triângulo de estabilidade, ou seja valores de a e q o mais próximo possível de 0,237 e 0,706 respectivamente. Na prática, porém, a resolução atingida depende da velocidade inicial do íon nas direções x e y e da direção e posição, em que o íon entra no filtro quadrupolar. Assim, o quadrupolo é essencialmente um filtro de resolução unitária. Na Figura 3 tem-se a representação de um espectrômetro de massas monoquadrupolar e o caminho percorrido por cada íon selecionado.

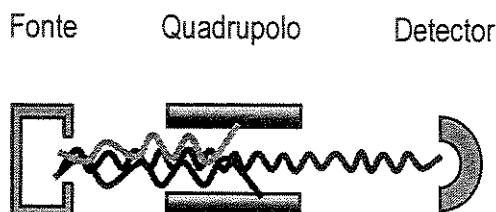


Figura 3 Esquema de um espectrômetro de massas monoquadrupolar

2.3. Detector

O detector consiste de uma série de eletrodos (conhecido como dinodos) arranjados próximos um do outro, onde é aplicada voltagem progressivamente mais alta em cada eletrodo [25-26]. Cada íon de m/z diferente selecionado é acelerado nos analisadores de massas e atingem o detector um em cada tempo. Quando os íons atingem o primeiro dinodo, alguns elétrons são liberados, os quais atingem o segundo dinodo. Este processo segue até o último dinodo ser atingido, gerando cada vez mais elétrons a cada impacto. No equipamento de espectrometria de massas utilizado é empregado um modelo de multiplicador de elétrons mais avançado em que possui dinodos contínuos. Neste sistema os elétrons podem seguir muitos caminhos, tornando indeterminado o número de estágios de multiplicação. No entanto, a amplificação é finita, saturando após 10^9 elétrons em um tempo da ordem de 10^{-8} segundos. Esta saturação pode ser evitada, controlando o gradiente de potencial aplicado [27].

Depois do sinal elétrico ser amplificado, é convertido em uma resposta digital que pode ser armazenada. Vantagens como baixo ruído na linha base, grande amplificação da corrente elétrica gerada a partir do primeiro dinodo e tempo de resposta em nanossegundos, são características deste tipo de detector. Entretanto, a desvantagem é que devido a grande amplificação da corrente elétrica, decorrente do arranjo de dinodos, resulta-se em uma grande amplificação do ruído de linha (60 Hz) na posição do pico, conseqüentemente gerando um tipo de ruído denominado de heteroscedástico e que será discutido no capítulo 4.

3. Espectrometria de massas seqüencial

No espectrômetro de massas monoquadrupolar, como o próprio nome diz, tem-se um quadrupolo apenas que seleciona cada íon que será detectado. Um espectrômetro de massas pentaquadrupolar permite a análise seqüencial de massas ou também conhecida como análise por espectrometria de massas em Tandem [28], no qual íons de m/z específico, formados a partir da ionização por elétrons na fonte de ionização, podem ser selecionados e dissociados em novos íons ou, então, sofrer reação com um gás reagente. Desta forma, a análise seqüencial pode fornecer informações importantes em comparação à espectrometria de massas padrão, pois resulta em um aumento da seletividade [23]. Isto porque analitos que possuem espectros de massas similares,

quando analisados pela simples ionização por elétrons, podem apresentar espectros de massas significativamente diferentes entre si, a partir da análise seqüencial, proporcionando assim um aumento de informações estruturais, conseqüentemente um aumento da seletividade para os analitos de interesse. Entretanto, a desvantagem da análise seqüencial é uma certa perda de sensibilidade, devido às várias etapas percorridas pelo íon desde fonte de ionização até o detector. Esta perda de sensibilidade pode ser compensada utilizando detectores multiplicadores de elétrons ou métodos de pré-concentração de amostra.

O espectrômetro de massas Extrel (Pittsburgh, PA) pentaquadrupolar, Figura 4, utilizado neste trabalho, consiste de três quadrupolos analisadores de massas (Q_1 , Q_3 , Q_5) e dois quadrupolos de reação ou dissociação (q_2 , q_4) [29]. Estes quadrupolos podem ser utilizados para a manipulação dos íons até que a informação estrutural da molécula seja totalmente elucidada.

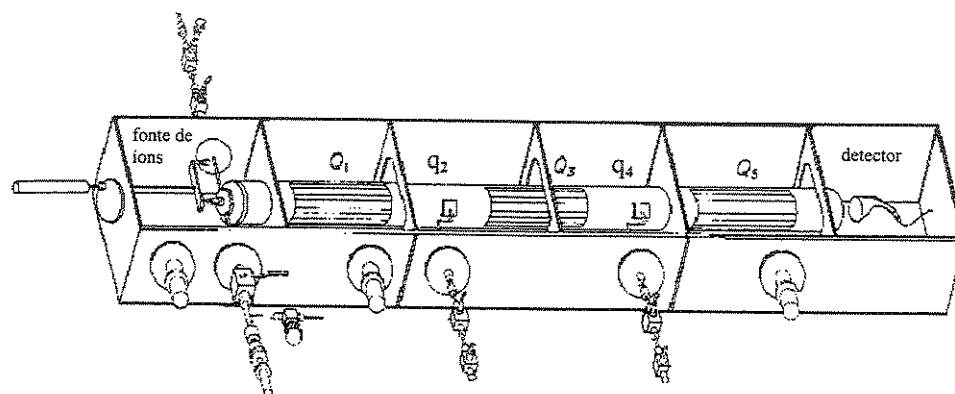


Figura 4. Espectrômetro de massas pentaquadrupolar com três analisadores de massas (Q_1 , Q_3 e Q_5) e dois quadrupolos de reação ou dissociação (q_2 e q_4).

3.1. Reação íon-molécula

Em um experimento de múltiplo estágio, em que reações seqüenciais são realizadas, íons obtidos na fonte de ionização por EI, a partir do analito de interesse, podem ser selecionados em Q_1 e transferidos para o quadrupolo q_2 , onde uma substância neutra pode colidir e reagir sob condições controladas. Os produtos desta reação podem ser selecionados em Q_3 e sofrer nova reação em q_4 com outro gás reagente, cujos produtos serão selecionados em Q_5 e detectados um a um pelo detector. Na Figura 5 está representado cada um destes passos nos cinco quadrupolos e os círculos preenchidos significam que os quadrupolos Q_1 , Q_3 e Q_5 estão sendo utilizados para selecionar os

ions. O resultado desta análise pode fornecer conjunto de dados bidimensionais, tridimensionais e até tetradimensionais, isto se as três etapas forem registradas.

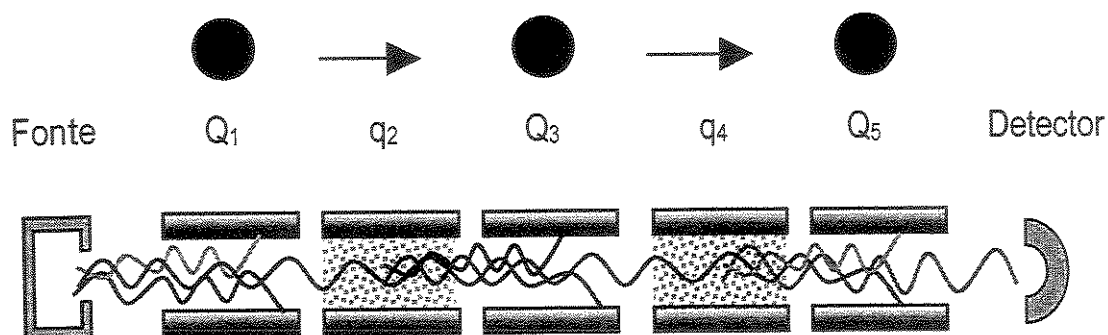


Figura 5. Esquema de um espectrômetro de massas pentaquadrupolar, no qual são representados todos os passos de uma análise seqüencial por reação íon-molécula.

3.2. Dissociação seqüencial

Outra vantagem no campo da espectrometria de massas é a possibilidade de análise seqüencial, através do processo de dissociação induzida por colisão ("collision-induced dissociation", CID). Em um experimento de múltiplo estágio, em que dissociações seqüenciais são realizadas, os íons são gerados na fonte de ionização, selecionados no primeiro analisador de massas Q_1 e dissociados em q_2 através da dissociação por CID, utilizando um gás inerte (por exemplo, argônio), gerando novos íons. Os produtos da dissociação podem ser selecionados em Q_3 e dissociados novamente em q_4 por CID, cujos produtos são selecionados um a um em Q_5 e detectados como representado na Figura 5. Cada etapa deste processo gera um novo espectro de massas que pode ser registrado. A versatilidade do espectrômetro de massas pentaquadrupolar é muito interessante, possibilitando diferentes varreduras que permitem a obtenção de conjuntos de dados bidimensionais, tridimensionais e mesmo tetradimensionais através de uma aquisição de dados simples e rápida.

4. Técnica MIMS

A espectrometria de massas por introdução via membrana é uma técnica analítica, na qual uma membrana semipermeável faz interface entre a amostra e o espectrômetro de massas. A técnica MIMS convencional é freqüentemente realizada com membranas de silicone. Este tipo de material mostra excelente performance para medidas de gases e compostos orgânicos voláteis de caráter hidrofóbico.

Da química à bioquímica, a espectrometria de massas por introdução via membrana ("membrane introduction mass spectrometry", MIMS) é uma técnica que está sendo aplicada em muitas áreas para análise de compostos orgânicos voláteis (VOC) e semivoláteis (SVOC). Monitoramento ambiental de compostos orgânicos voláteis em água, ar e solo, análise de sangue, tratamento de resíduos biológicos, bioreatores e reações químicas são alguns exemplos das aplicações que têm sido exploradas usando a técnica MIMS [5-19]. A grande vantagem da técnica MIMS em relação à maioria das técnicas analíticas é que pode eliminar algumas etapas preliminares de extração ou pré-concentração, sendo que a própria membrana pode ser utilizada na pré-concentração do analito. A junção da técnica MIMS com a técnica de análise por injeção em fluxo (FIA) [30] apresenta excelente precisão e exatidão, rapidez, simplicidade e economia de amostras. Na análise pela técnica FIA-MIMS todas as etapas de amostragem, ionização, separação e detecção de íons, aquisição e armazenamento de dados podem ser automaticamente controladas, minimizando assim o erro experimental.

As moléculas de polaridade similar à membrana são adsorvidas na superfície do polímero, ocorrendo um enriquecimento acentuado do analito. As moléculas difundem-se através das paredes e evaporam para o espectrômetro de massas, onde são ionizadas. Este processo de três etapas é denominado de pervaporação e é influenciado por três parâmetros: dimensão da membrana, pressão de vapor do analito e permeabilidade (produto da solubilidade e difusibilidade). A difusibilidade é dependente do gradiente de concentração e da dimensão da molécula e não varia significativamente entre compostos de peso molar similar. O processo é relativamente rápido (poucos minutos), comparado à cromatografia convencional, pois o tempo de análise depende basicamente do tempo de transporte do analito através da membrana.

Muitas aplicações encontradas na literatura têm utilizado a membrana de silicone [(poli(dimetil siloxano))] na forma planar ou tubular. A membrana de silicone tem a grande vantagem de permeabilidade para compostos orgânicos hidrofóbicos. Desta forma, as membranas de silicone são ideais para a amostragem de compostos orgânicos com moléculas pequenas e presentes em baixas concentrações em fluxos aquosos.

A membrana é denominada de semi-seletiva, devido às diferenças nas interações químicas e físicas entre o analito e a membrana, que resultam em variações nos tempos de difusão dos componentes das misturas. Os compostos com maior permeabilidade são normalmente os que possuem baixa polaridade e baixa massa molar, geralmente características inerentes dos compostos

orgânicos voláteis. Entretanto, os tempos de difusão podem ser muito próximos resultando na passagem simultânea dos compostos através da membrana.

Limites de detecção da ordem de $\mu\text{g L}^{-1}$ são facilmente alcançados para compostos orgânicos voláteis e esta propriedade juntamente com boa precisão e rapidez, fazem desta técnica uma excelente escolha no controle ambiental e para algumas aplicações que exigem monitoramento em linha e em campo [31]. Entretanto, novas tecnologias com relação à técnica MIMS têm sido desenvolvidas para obtenção de limites de detecção mais baixos. Para este fim etapas de pré-concentração e aumento de temperatura estão sendo utilizados, resultando em limites de detecção de ng L^{-1} e pg L^{-1} . Técnicas como "trap-and-release", "Cryotrap" (CT-MIMS) e "laser desorption" são exemplos de métodos que utilizam a pré-concentração das amostras [32].

Capítulo 2

Análise multivariada de dados

1. Introdução

A química analítica é uma ciência cujas medidas são imprecisas e por isso os químicos analíticos devem ter sempre a disposição técnicas estatísticas simples ou mais avançadas (métodos quimiométricos) para solucionar problemas [33].

A quimiometria é uma disciplina da química que utiliza matemática, estatística e lógica formal com o objetivo de planejar ou selecionar procedimentos experimentais e garantir o máximo de informações químicas relevantes a partir dos dados analisados, obtendo assim, conhecimento a partir do sistema químico [34,35].

Neste capítulo serão discutidos os métodos quimiométricos aplicados neste trabalho e que se mostraram mais adequados na análise multivariada dos dados.

2. Análise exploratória através da análise de componentes principais

A análise dos componentes principais ("principal component analysis", PCA) [20-22] é a base fundamental da maioria dos métodos de análise multivariada. A análise multivariada investiga a correlação entre variáveis e explora a estrutura de um grande conjunto de dados, representado por uma matriz X ($I \times J$), onde I é o número de amostras analisadas e J é o número de variáveis para cada amostra. A PCA organiza os dados condensando-os dentro de um reduzido número de variáveis, que são mais facilmente compreendidos. Para a análise exploratória ou classificação de amostras baseada em similaridades ou determinação de uma variável desconhecida disposta em um vetor y ($I \times 1$), muitas variáveis instrumentais podem ser utilizadas simultaneamente como: valores de absorção para determinados comprimentos de onda e intensidade de pico para determinados valores de m/z .

A PCA é um método de análise multivariada que decompõe uma matriz de dados X de posto ("rank") R , em uma soma de R matrizes de posto igual a 1, onde posto expressa o número de vetores linearmente independentes de uma matriz de dados. Quimicamente falando, o posto da matriz de dados de uma amostra deve ser igual ao número de espécies detectadas. Essas novas

matrizes de posto 1 são produtos de vetores chamados escores ("scores") $\mathbf{t}_r$ e pesos ("loadings") $\mathbf{p}_r$. Os escores e pesos podem ser calculados par a par por um processo iterativo, como na equação 3.

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^\top + \mathbf{t}_2\mathbf{p}_2^\top + \dots + \mathbf{t}_r\mathbf{p}_r^\top + \dots + \mathbf{t}_R\mathbf{p}_R^\top \quad (3)$$

Em muitos casos, o posto matemático de $\mathbf{X}$ pode ser alto, devido à presença de ruído, porém o posto químico ('ou pseudoposto') de $\mathbf{X}$ é baixo, pois existem poucos componentes químicos presentes. Isto significa que R é escolhido para reter apenas a informação importante em $\mathbf{X}$ (por exemplo, $R = 2$ ou 3) e o ruído é descrito separadamente na matriz $\mathbf{E}$, como na equação 4.

$$\begin{aligned} \mathbf{X} &= \mathbf{t}_1\mathbf{p}_1^\top + \mathbf{t}_2\mathbf{p}_2^\top + \dots + \mathbf{t}_R\mathbf{p}_R^\top + \mathbf{E} \\ &= \mathbf{TP}^\top + \mathbf{E} \end{aligned} \quad (4)$$

Em síntese, a PCA é um método que tem por finalidade básica a redução de dados a partir de combinações lineares das variáveis originais. Na Figura 6 é possível visualizar esta mesma decomposição dos dados na forma de matrizes e também visualizar geometricamente o plano definido pelos vetores, para apenas duas amostras.

Na representação geométrica da decomposição dos dados por PCA, as amostras representadas no espaço das variáveis $\mathbf{X}$ (espaço tridimensional amarelo) são projetadas em um novo espaço (plano cinza) de dimensões reduzidas. Este novo espaço é definido pelas componentes principais (vetores verde), cujas direções em relação ao espaço das variáveis $\mathbf{X}$ são definidas pelos valores pesos, de forma que estas componentes mantenham a ortogonalidade entre si. As posições das amostras no novo espaço são definidas pelos escores (azul) e o modelo busca minimizar os resíduos (rosa) entre as amostras e o novo espaço (plano cinza).

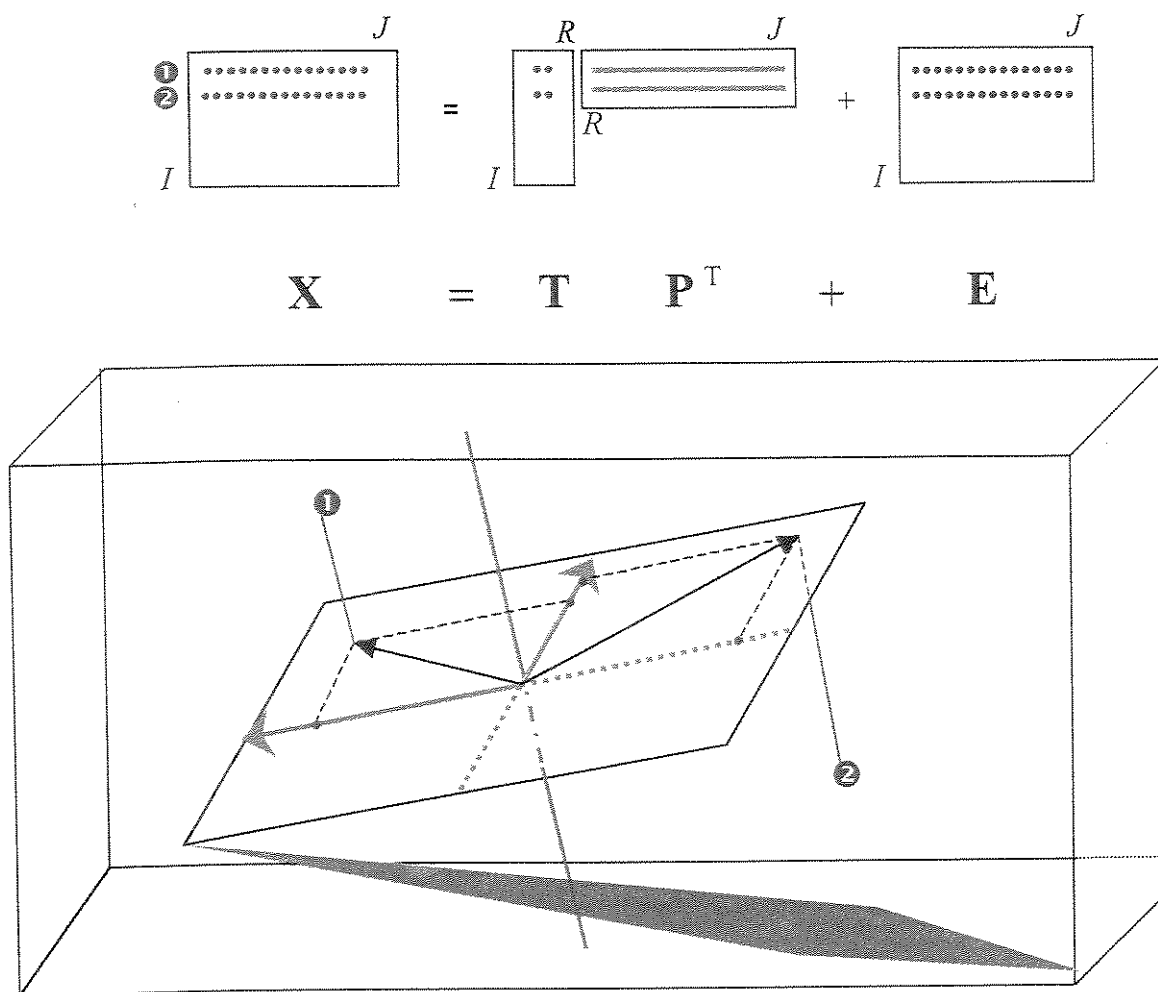


Figura 6. Representação da decomposição dos dados na forma de matriz e também através de representação geométrica. ① e ② são duas amostras representadas por "J" variáveis.

3. Calibração multivariada

Na calibração multivariada, um conjunto de amostras $\mathbf{X}$ ($I \times J$) de concentração $\mathbf{y}$ ($I \times 1$) conhecida para apenas um analito na amostra ou de concentração $\mathbf{Y}$ ($I \times M$), no caso de mais de um analito, é usado para a construção de um modelo de regressão. A partir deste modelo de regressão, pode ser encontrada a composição de amostras desconhecidas. Antes, porém, de discutir os tipos de modelos ou métodos de calibração multivariada utilizados neste trabalho é importante primeiro descrever quais são os tipos de dados analíticos obtidos.

Os instrumentos analíticos geram diferentes tipos de dados $\mathbf{X}$, que podem ser classificados em dados de ordem zero, de primeira ordem e de segunda ordem [33]. Na espectrometria de

massas, por exemplo, pode-se obter os três tipos diferentes de dados. Dados de ordem zero em espectrometria de massas são obtidos a partir da seleção e monitoramento de um pico específico para cada analito ("selective-ion monitoring", SIM). Neste caso são utilizados os valores máximos de intensidades para cada pico selecionado, que corresponde a um valor de concentração referente a uma dada substância, para a construção do modelo univariado e quantificação das amostras. Entretanto, para a análise de misturas complexas contendo muitos analitos ou analitos com espectros de massas muito semelhantes (isômeros), o método univariado SIM pode ser muito limitado. Isto porque, encontrar um pico seletivo para cada analito de interesse pode não ser possível. Para casos particulares como estes é necessário aplicar a análise multivariada, em que se utilizam dados de primeira ordem, ou seja, todo o espectro de massas pode ser utilizado ("total-ion monitoring", TIM), resultando assim em maior informação em relação aos analitos presentes na mistura. Uma outra alternativa é os chamados dados de segunda ordem ou ainda de ordem N , que aumenta ainda mais a seletividade da análise e no caso da espectrometria de massas, são gerados a partir da análise seqüencial dos íons.

Dentre os métodos de calibração multivariada que existem, o mais comumente aplicado em dados de análise química, é o método dos mínimos quadrados parciais ("partial least squares", PLS) [20-22]. Dois trabalhos recentes na literatura descrevem a análise por amostragem direta, em espectrometria de massas, de misturas contendo analitos com espectros similares [36,37]. Foi utilizado o PLS para a calibração multivariada de dados de primeira ordem e análise de fatores paralelos ("parallel factor analysis", PARAFAC) para dados de segunda ordem. Estes trabalhos demonstram claramente que a calibração multivariada a partir de espectros de massas fornece resultados superiores em comparação a aproximação univariada.

3.1. Calibração multivariada de dados de primeira ordem

Instrumentos que geram dados de primeira ordem incluem todos os tipos de espectroscopia, cromatografia, eletroquímica e a espectrometria de massas. Estes instrumentos são capazes de gerar múltiplas medidas a um determinado tempo ou para uma determinada amostra. Estas medidas podem ser ordenadas como um vetor e o conjunto de vetores decorrentes de cada amostra analisada gera uma matriz de dados X ($I \times J$), onde I é igual ao número de amostras analisadas e J é igual ao número de variáveis em cada vetor de dados.

3.1.1. PLS

O PLS é um método de calibração multivariada bem estabelecido dentro da química analítica [20-22] e, como todo método de regressão, tem o objetivo de encontrar uma relação entre um grupo de dados X e a resposta y . A maior diferença, entretanto, entre o PLS e a regressão linear múltipla (multiple linear regression, MLR) é que o PLS é capaz de dar uma previsão estável mesmo quando X contém variáveis altamente correlacionadas (como em dados espectrométricos). Diferentemente do método univariado, o PLS pode dar bons valores de previsão mesmo se interferentes estão presentes nas amostras. A única exigência é que os interferentes presentes nas amostras a serem analisadas estejam também presentes nas amostras utilizadas para a construção do modelo de calibração.

O algoritmo para a construção do modelo pode ser descrito em duas etapas. Inicialmente é construído um modelo de posto 1, como mostrado na PCA, para cada componente a partir de X e y , simultaneamente. Estes modelos são subtraídos de X e y e um novo grupo de componentes é encontrado a partir dos resíduos. O cálculo das componentes é a principal parte do algoritmo e é também o principal problema na determinação de um vetor peso, w , que maximiza a função de covariância entre os escores de X e y . Note que o vetor peso neste caso é denominado w , e não p , como na PCA, pois no PLS os valores pesos que determinam a direção das componentes principais não garantem mais a ortogonalidade destas componentes. Para calcular cada componente no modelo PLS a partir de X temos:

$$\hat{x}_{ij} = t_{r,i} w_{r,j} \quad (5)$$

Cada variável $\hat{x}_{ij}$ é estimada pelo produto de um escalar $t_{r,i}$ que corresponde ao valor de escore da amostra i para a componente r e $w_{r,j}$ é o valor de peso da variável estimada $\hat{x}_{ij}$ para a componente r . Assim o objetivo do algoritmo do PLS é encontrar componentes bilineares (t e w) em X , que tenham uma máxima covariância com uma resposta y , minimizando o resíduo entre x_{ij} e $t_{r,i} w_{r,j}$ como descrito pela expressão abaixo:

$$\max_{w_r} \left[\text{cov}(t_r, y) \mid \min \left(\sum_{i=1}^I \sum_{j=1}^J (x_{ij} - t_{r,i} w_{r,j})^2 \right) \right]$$

onde $\mathbf{t}_r$, $(I \times 1)$ é o vetor de escores para cada amostra I em relação a componente r , e $\mathbf{w}_r$, $(J \times 1)$ é o vetor de pesos da componente r no modelo, para cada variável J .

O uso simultâneo da informação em $\mathbf{X}$ e $\mathbf{y}$ faz o método mais complexo que os outros métodos de regressão e dois vetores pesos são necessários para garantir a ortogonalidade das componentes.

As variáveis dependentes e independentes são centradas na média e a matriz $\mathbf{X}_0$ e o vetor $\mathbf{y}_0$ são encontrados. Para cada componente, $r = 1 \dots R$, seja incluída no modelo de regressão, os seguintes passos são realizados:

- (a) O vetor peso $\mathbf{w}_r$ é calculado maximizando a covariância entre $\mathbf{X}_{r-1}$ and $\mathbf{y}_{r-1}$ que é dado por $\mathbf{w}_r^T \mathbf{w}_r = 1$.
- (b) Os escores, $\mathbf{t}_r$, são estimados a partir da projeção de $\mathbf{X}_{r-1}$ em $\mathbf{w}_r$.
- (c) O vetor peso $\mathbf{p}_r$ é determinado pela regressão de $\mathbf{X}_{r-1}$ em $\mathbf{t}_r$ e similarmente $\mathbf{q}_r$ pela regressão de $\mathbf{y}_{r-1}$ em $\mathbf{t}_r$.
- (d) A partir de $(\mathbf{X}_{r-1} - \mathbf{t}_r \mathbf{p}_r^T)$ e $(\mathbf{y}_{r-1} - \mathbf{t}_r \mathbf{q}_r^T)$ novas matrizes $\mathbf{X}_r$ e $\mathbf{y}_r$ são formadas.

O número ótimo de componentes para incluir no modelo pode ser encontrado pela validação estatística e/ou por observação do gráfico dos componentes encontradas pelo modelo. O modelo PLS pode ser escrito na forma de equação de regressão generalizada:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad (6)$$

onde $\mathbf{e}$ é o erro de previsão. O coeficiente de regressão, $\mathbf{b}$, pode ser calculado por:

$$\mathbf{b} = \mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1} \mathbf{q} \quad (7)$$

onde $\mathbf{W}$ é a matriz de peso que maximiza a covariância entre $\mathbf{X}_{t-1}$ and $\mathbf{y}_{t-1}$ e $\mathbf{P}$ é uma matriz de peso determinada pela regressão de $\mathbf{X}_{t-1}$ em $\mathbf{t}_t$ e similarmente $\mathbf{q}_t$ pela regressão de $\mathbf{y}_{t-1}$ em $\mathbf{t}_t$. O modelo PLS é calculado usando um grupo de amostra de treinamento. Depois que o modelo é desenvolvido pode ser utilizado para fazer previsão de novas amostras de concentração desconhecida.

Quando um conjunto de amostras contém diferentes respostas para R diferentes analitos, o mais comum é construir modelos de calibração separados para cada analito.

3.2. Calibração multivariada de dados de segunda ordem

Devido à crescente sofisticação das técnicas instrumentais nos últimos anos, o número de problemas de regressão dentro da química, decorrentes de dados de segunda ordem, tem crescido muito. Exemplos de dados de segunda ordem são os gerados a partir das chamadas técnicas hífenadas como GC/MS e GC/FT-IR e técnicas como MS/MS e 2DNMR [33]. Apesar da complexidade instrumental, a vantagem do uso deste tipo de técnicas é o aumento da seletividade analítica.

Entre estes instrumentos existem duas classes de dados: dados bilineares e dados não bilineares. A técnica GC/MS, por exemplo, gera dados bilineares, nos quais cada espectro adquirido em diferentes tempos para uma mesma amostra está correlacionado e cada analito presente na amostra tem posto igual a 1. O posto da matriz de dados de uma amostra, como discutido antes, deve ser igual ao número de espécies espectrometricamente ativas ou detectadas.

Para dados resultantes de análise 2D NMR e MS/MS, denominados de não bilineares, o posto de cada analito não é igual a 1. Em MS/MS, por exemplo, cada fragmento ou íon selecionado na primeira dimensão gera um novo espectro de massas. Estes espectros de massas resultantes na segunda dimensão não apresentam nenhuma correlação entre si. Assim o posto de uma matriz de dados de MS/MS de um analito puro será igual ao número de espectros diferentes gerados a partir dos fragmentos iniciais. Se uma espécie produz 10 espectros diferentes o posto será igual a 10.

Um espectro de massas tridimensional obtido por MS/MS pode ser representado em uma matriz de dados, $\mathbf{X}$ ($J \times K$), onde K representa cada íon precursor de m/z específico e que é dissociado produzindo um espectro de J valores de m/z . Um conjunto de amostras com dados bidimensionais produz um arranjo tridimensional representado por um tensor de dados, $\underline{\mathbf{X}}$ ($I \times J \times K$), onde I é o número de amostras. As concentrações dos M diferentes analitos destas amostras podem ser representadas em uma matriz $\mathbf{Y}$ ($I \times R$). Esta situação é representada na Figura 7.

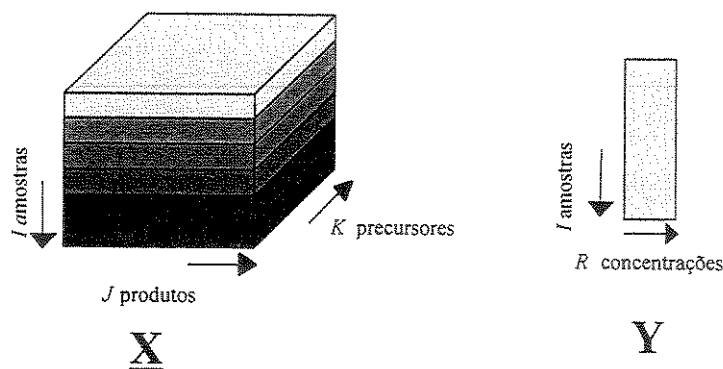


Figura 7 Arranjo tridimensional $\underline{\mathbf{X}}$ ($I \times J \times K$) e a matriz de concentração, $\mathbf{Y}$ ($I \times R$).

Vários métodos de calibração multivariada são descritos na literatura para dados mais complexos como os de segunda ordem. Neste trabalho foram aplicados todos os métodos de calibração multivariada descritos na literatura e que podem ser adequados para dados de segunda ordem não bilineares. Entre eles o método dos mínimos quadrados parciais desdobrados ("unfold partial least squares", U-PLS) [38], o método dos mínimos quadrados parciais multilineares ("multilinear partial least squares", N-PLS) [39-42], o método de resolução de curvas multivariada (multivariate curve resolution, MCR) com restrições [43,44], o método de aniquilação de posto não bilinear ("non-bilinear rank annihilation", NBRA) [45] e bilinearização residual ("residual bilinearization", RBL) [46]. Entretanto, somente o U-PLS, N-PLS e o MCR mostram ser mais adequados, fornecendo bons resultados de previsão. Por isso, para os dados de segunda ordem serão apresentadas a teoria e discussão apenas para o U-PLS, N-PLS e MCR.

3.2.1. U-PLS

Os algoritmos do U-PLS e do N-PLS, que serão discutidos em seguida, são uma extensão do algoritmo original do PLS e foram desenvolvidos para dados multidimensionais, $\underline{\mathbf{X}}$ ($I \times J \times K$).

Quando o U-PLS é utilizado em dados multidimensionais, $\underline{X}$ é primeiramente rearranjado ou desdobrado ('matricized' ou 'unfolded') para produzir um arranjo bidimensional, Figura 8. Assim, os dados tridimensionais, $\underline{X} (I \times J \times K)$, são desdobrados resultando em um arranjo bidimensional, $X (I \times JK)$.

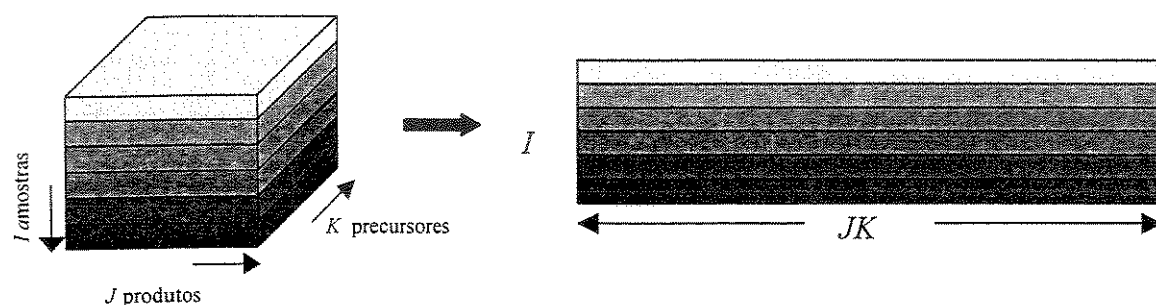


Figura 8. Representação do desdobramento realizado para cada uma das matrizes de dados I com K linhas e J colunas, resultando em uma matriz de dados única com I linhas e JK colunas.

Desta forma é possível utilizar o mesmo algoritmo padrão do PLS bidimensional, encontrando componentes bilineares em X (escores e pesos) que apresentem a máxima covariância com a resposta, y , como na expressão abaixo:

$$\max_{w_r} \left[\text{cov}(t_r, y) \left| \min \left(\sum_{i=1}^I \sum_{jk=1}^{JK} (x_{ijk} - t_{r,i} w_{r,jk})^2 \right) \right] \right]$$

onde $t_r (I \times 1)$ é o vetor escore do PLS da componente r , para cada amostra i e $w_r (JK \times 1)$ é o vetor peso do PLS para a componente r no modelo de calibração. Note que o vetor peso w_r tem dimensão JK , ou seja, o produto das duas dimensões J e K do tensor desdobrado $\underline{X}$.

Da mesma forma que no PLS, coeficientes de regressão b_{U-PLS} , que relacionam X e y são encontrados e utilizados na previsão de amostras desconhecidas.

3.2.2. N-PLS

Assim como o U-PLS, o N-PLS é uma extensão do algoritmo do PLS bidimensional e trabalha com dados multidimensionais, $\underline{X}$, que é um arranjo de dados de ordem maior que dois. Como no PLS bidimensional, o objetivo é encontrar componentes em $\underline{X}$ que tenham a máxima

covariância com $\mathbf{y}$. A diferença entre os dois modelos de calibração é que para o N-PLS são utilizadas componentes multilineares. No caso específico de dados tridimensionais, $\underline{\mathbf{X}}$, são utilizadas componentes trilineares e podem ser representadas pela expressão abaixo:

$$\max_{\mathbf{w}_r^J, \mathbf{w}_r^K} \left[\text{cov}(\mathbf{t}_r, \mathbf{y}) \right] \min \left(\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (x_{ijk} - t_{r,i} \mathbf{w}_{r,j}^J \mathbf{w}_{r,k}^K)^2 \right)$$

onde pode ser verificado que dois vetores peso são usados, desta vez, para cada componente no modelo: $\mathbf{w}_r^J$ ($J \times 1$) e $\mathbf{w}_r^K$ ($K \times 1$). Assim, em contraste com o PLS e U-PLS bidimensionais, cada dimensão do arranjo tridimensional é descrita por seus vetores escores e pesos. Na Figura 9, tem-se os dois modelos, onde é mostrado em (a) a decomposição do modelo bidimensional e em (b) a decomposição para o modelo tridimensional.

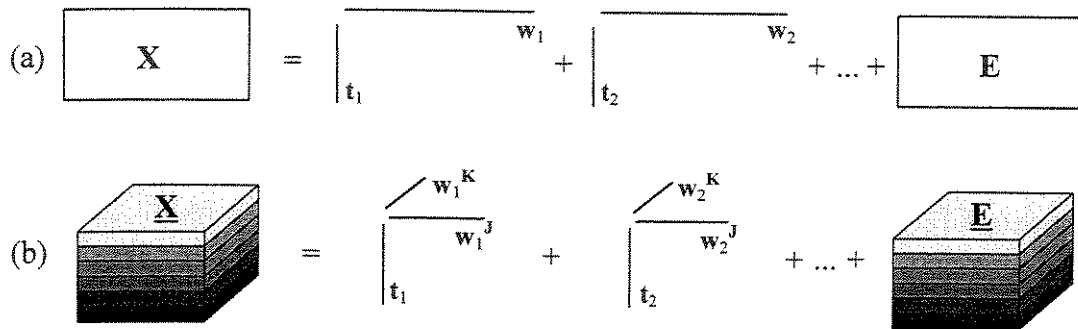


Figura 9. (a) Decomposição de $\mathbf{X}$ para PLS e U-PLS e (b) Decomposição de $\underline{\mathbf{X}}$ para N-PLS.

3.2.3. MCR

O método de resolução de curva multivariada ("multivariate curve resolution", MCR), é um outro tipo de calibração multivariada para tratamento de dados de segunda ordem. Da mesma forma que no U-PLS, inicialmente o tensor de dados $\underline{\mathbf{X}}$ é desdobrado, Figura 8, obtendo-se uma matriz de dados, neste caso representada por $\mathbf{X}$.

Freqüentemente, parte da variância nos dados experimentais não é causada pela mudança na concentração das espécies químicas de interesse, mas por outros fatores que incluem erros experimentais, interferências e flutuações do sinal por causa do detector.

Para eliminar estes fatores o MCR segue o mesmo princípio de decomposição da matriz de dados aplicada em PCA. Entretanto, esta decomposição segue uma outra interpretação, no qual o sinal analítico obedece a uma relação semelhante à lei de Beer, ou seja, se os espectros têm uma relação linear com a concentração, pode ser representado de acordo com a equação 7.

$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad (8)$$

onde $\mathbf{X}$ é a matriz de dados contendo os espectros de massas desdobrados, $\mathbf{C}$ é a matriz de perfil de concentração das amostras, $\mathbf{S}$ é a matriz que contém os espectros dos analitos puros com concentração unitária dentro das misturas e $\mathbf{E}$ é a matriz de resíduo que não é explicado pelas espécies químicas em $\mathbf{C}$ e $\mathbf{S}$. As dimensões destas quatro matrizes definidas são $\mathbf{X}$ ($I \times J$), $\mathbf{C}$ ($I \times R$), $\mathbf{S}$ ($J \times R$) e $\mathbf{E}$ ($I \times J$), onde I é igual ao número de amostras ou espectros analisados, J é o número de variáveis em cada espectro e R é o número de espécies químicas nas misturas. Para que o MCR possa ser aplicado, além de necessariamente ter que haver uma relação linear entre espécies e concentração, o posto da matriz de dados deve ser igual ao número de espécies analisadas nas misturas.

O algoritmo MCR utiliza o método de decomposição de valores singulares ("singular value decomposition", SVD) [34], assim como, o PCA tem a finalidade de filtrar os dados eliminando a parte do sinal com ruído ou que não tenha informação analítica.

O MCR é baseado no método dos mínimos quadrados alternados ("alternating least squares", ALS) [47], onde o processo de otimização é realizado iterativamente, resolvendo as equações a seguir:

$$\begin{aligned} \mathbf{C} &= \mathbf{X}^* \mathbf{S} (\mathbf{S}^T \mathbf{S})^{-1} \\ &= \mathbf{X}^* \mathbf{S}^{+T} \end{aligned} \quad (9)$$

$$\begin{aligned} \mathbf{S}^T &= (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{X}^* \\ &= \mathbf{C}^{+} \mathbf{X}^* \end{aligned} \quad (10)$$

onde a matriz $\mathbf{C}^{+}$ é a pseudoinversa da matriz $\mathbf{C}$ e a matriz $\mathbf{S}^{+}$ é a pseudoinversa da matriz $\mathbf{S}$. A cada iteração de otimização, nova matriz $\mathbf{C}$ de concentração e $\mathbf{S}$ de espectro é estimada e obtida sucessivamente. Quando é realizada uma estimativa inicial da matriz $\mathbf{S}$ a equação 7 é resolvida para

C usando a equação 9. Da mesma forma, quando é realizada uma estimativa inicial da matriz **C** a equação 7 é resolvida para **S** usando a equação 10.

Pode-se dizer que o MCR é muito similar a análise de componentes principais (PCA), pois também decompõe a matriz de dados em vetores, eliminando uma matriz de resíduos que não é explicada pelas espécies químicas puras em **C** e **S**, no qual espera-se que seja próximo ao erro experimental.

O MCR recupera os espectros dos analitos puros e a partir destes espectros é possível prever a concentração de amostras desconhecidas [48], como mostrado nas equações 11, 12 e 13.

$$\mathbf{C}_{\text{teste}} = \mathbf{X}_{\text{teste}} \mathbf{S}_{\text{MCR}} (\mathbf{S}_{\text{MCR}}^T \mathbf{S}_{\text{MCR}})^{-1} \quad (11)$$

$$\mathbf{b} = (\mathbf{C}_{\text{MCR}}^T \mathbf{C}_{\text{MCR}})^{-1} \mathbf{C}_{\text{MCR}}^T \mathbf{C}_{\text{calibração}} \quad (12)$$

$$\mathbf{C}_{\text{previsto}} = \mathbf{C}_{\text{teste}} * \mathbf{b} \quad (13)$$

A equação 11 é decorrente da equação 7, onde $\mathbf{C}_{\text{teste}}$ é a concentração das amostras testes obtidas a partir dos espectros das amostras testes $\mathbf{X}_{\text{teste}}$ e dos espectros puros $\mathbf{S}_{\text{MCR}}$ recuperados pelo MCR. Nesta equação é utilizada a pseudoinversa da matriz $\mathbf{S}_{\text{MCR}}$, pois esta não é uma matriz quadrada e sua inversa não existe.

Os valores de concentração das amostras teste encontrados pela equação 11 estão em uma escala diferente dos valores reais de concentração, assim deve-se encontrar um fator de regressão **b** através dos mínimos quadrados de acordo com a equação 12. O $\mathbf{C}_{\text{MCR}}$ é a concentração das amostras de calibração recuperada pelo MCR e $\mathbf{C}_{\text{calibração}}$ são os valores de concentrações reais das amostras de calibração. Na equação 13 o $\mathbf{C}_{\text{previsto}}$ é a concentração prevista a partir de $\mathbf{C}_{\text{teste}}$ corrigida pelo fator de regressão, **b**.

A principal vantagem deste método é que é facilmente adaptado a conjuntos de dados de complexidades diferentes como estruturas trilineares ou não trilineares, resultando em ótimas soluções por mínimos quadrados.

Capítulo 3

Análise exploratória de espectros de massas

1. Introdução

A separação de amostras analisadas pela técnica FIA-MIMS a partir da análise dos componentes principais (PCA), garante a identificação de compostos orgânicos voláteis a partir das diferenças entre espectros de massas ou da difusibilidade destes compostos através da membrana.

Neste capítulo, dois experimentos diferentes foram realizados aplicando a técnica FIA-MIMS, MIMS e a análise por PCA. No primeiro experimento foram analisadas amostras de água contaminadas com gasolina e identificados quais são os compostos orgânicos voláteis que mais solubilizam em água. No segundo experimento foram analisadas amostras de cerveja e identificados alguns dos compostos orgânicos voláteis responsáveis pela qualidade deste produto.

Nos dois casos, nenhum processo de separação ou derivatização foi utilizado e as amostras foram injetadas diretamente no sistema. A membrana semipermeável de silicone utilizada mostrou excelente desempenho para análise de compostos orgânicos voláteis presentes nos dois tipos de amostras analisadas.

2. Análise de amostras de água contaminadas por gasolina através da técnica FIA-MIMS

As águas subterrâneas contaminadas por combustíveis são um dos sérios problemas ambientais que tem sido destacado com frequência. Derramamento de óleo e vazamentos subterrâneos a partir de tanques de estocagem de gasolina têm contaminado o solo e conseqüentemente atingem os recursos hídricos subterrâneos. A maior preocupação é que alguns constituintes orgânicos tóxicos podem se solubilizar em água, devido à adição de solventes orgânicos polares que são completamente miscíveis ou altamente solúveis em água. A presença de compostos como o etanol, metanol e metil *tercio*-butil éter (MTBE) [49-51] em uma mistura de hidrocarbonetos e água, resultará no efeito de co-solvência que aumentará a concentração de compostos orgânicos na solução aquosa, contaminando os recursos naturais de água subterrânea. Dentre estes compostos orgânicos estão os hidrocarbonetos aromáticos, benzeno, tolueno e os isômeros etilbenzeno, *orto*-, *meta*- e *para*-xilenos [49-51].

O BTEX apresenta um alto potencial de contaminação devido a sua concentração relativamente alta em combustível e grande solubilidade em água na presença de etanol. Na Tabela 1 encontram-se os valores máximos de contaminação aceitos para consumo humano, encontrados na PORTARIA N.º 1469, de 29 de dezembro de 2000, do Ministério da Saúde.

Tabela 1. Valores máximos permitidos de BTEX para o consumo humano

Compostos orgânicos voláteis	VMP* ($\mu\text{g L}^{-1}$)
Benzeno	5
Tolueno	170
Etilbenzeno	200
Xilenos	300

*VMP: valor máximo permitido

As gasolinas brasileiras (comum, aditivada e premium) são uma mistura de 75 a 80% de gasolina e uma proporção de álcool etílico (etanol) que varia, dependendo das conjunturas econômicas, de $20 \pm 1\%$ a $24 \pm 1\%$ [52]. Um importante efeito de co-solvência é produzido a partir do etanol, na solubilização do BTEX em águas subterrâneas em contato com a gasolina [53-57] e conseqüentemente os valores de contaminação podem ser bem superiores aos valores máximos permitidos pelo Ministério da Saúde.

No monitoramento de contaminação de águas subterrâneas por combustíveis, vários contaminantes químicos estão presentes e devem ser quantificados em concentrações baixas (traços). Entretanto, a matriz água pode conter constituintes que tornam a análise bastante complexa, podendo apresentar um potencial de interferência considerável [58]. A técnica de análise comumente utilizada é a cromatografia gasosa e vem sendo empregada com sucesso até agora. Entretanto, a espectrometria de massas por introdução via membrana (MIMS), como já discutido anteriormente, é uma técnica que apresenta algumas vantagens em relação à cromatografia gasosa e uma delas é a total compatibilidade com a técnica FIA [59]. Desta forma, o monitoramento em tempo real de compostos orgânicos voláteis ou semivoláteis pode ser realizado e aplicado diretamente no campo, pois espectrômetros de massas que viabilizem esta análise já estão sendo desenvolvidos [31].

Em MIMS, a introdução do analito no espectrômetro de massas ocorre via transporte seletivo através da membrana de silicone. A hidrofobicidade da membrana de silicone e sua alta permeabilidade para compostos apolares ou menos polares permitem uma eficiente e seletiva permeação de compostos orgânicos voláteis a partir de águas contaminadas, que reduz consideravelmente o efeito da matriz. A extração e pré-concentração do analito na membrana são realizadas rápida e simultaneamente.

Combustíveis como gasolinas são amostras complexas que contêm principalmente hidrocarbonetos (Tabelas 2 e 3) [50,58], cuja composição depende do tipo de petróleo e método de produção.

Tabela 2. Composição típica de gasolinas livres de aditivos.

constituintes	porcentagem (V/V %)
alcanos (C3-C10 n-parafinas e C4-C13 isoparafinas) e cicloalcanos (C5-C13 cicloparafinas)	56
alcenos (C2-C12)	10
hidrocarbonetos aromáticos (C6-C12)	34
hidrocarbonetos policíclicos aromáticos (PAHs)	<1

Na Tabela 3 estão os valores de solubilidade dos hidrocarbonetos aromáticos BTEX em água e a porcentagem destes compostos encontrados em gasolina.

Tabela 3. Solubilidade em água e porcentagem de BTEX em gasolina comum.

BTEX	solubilidade em água (mg L ⁻¹)	porcentagem em gasolina (V/V %)
Benzeno	1740-1860	2-5
Tolueno	500-627	6-7
Etilbenzeno	131-208	5
Xilenos	167-213	6-7

Nas gasolinas aditivadas, além do etanol, são adicionados pelas distribuidoras outros aditivos como detergentes dispersantes, com a função de manter limpo todo o sistema por onde passa a gasolina. Nas gasolinas premium também é adicionado etanol, mas a diferença para a

gasolina comum e aditivada é que o índice de octanagem é mais alto. A octanagem é uma medida da resistência da gasolina à detonação ("batida de pino") [52].

Neste trabalho foi realizado inicialmente um monitoramento SIM e as concentrações de BTEX e etanol em águas contaminadas por gasolina foram determinadas através da técnica FIA-MIMS. Estas mesmas amostras foram também analisadas através do monitoramento TIM e a análise multivariada por PCA foi aplicada nos espectros de massas obtidos. Desta forma foi possível avaliar a habilidade desta técnica para monitorar a variação de composição de BTEX e etanol em águas contaminadas por gasolinas brasileiras.

2.1. Parte experimental

Foram analisadas por MIMS amostras de águas contaminadas por gasolina, no qual foram utilizadas para a contaminação gasolinas comuns, aditivadas e premium. No total, 19 alíquotas de gasolina foram recolhidas aleatoriamente na cidade de Campinas, SP, a partir de vários postos de distribuição com diferentes bandeiras.

Cada amostra foi preparada adicionando 2,5 mL da gasolina em 50 mL de água deionizada em erlenmeyer de 250 mL. Cada recipiente foi fechado com um septo de borracha e um tubo de polietileno, mergulhado abaixo do nível da gasolina, foi adaptado para coletar a solução aquosa. Este sistema contendo as duas fases foi deixado em repouso por 24 h a 22 ± 2 °C, resultando na solubilização dos compostos orgânicos voláteis (BTEX, etanol e outros compostos orgânicos).

As concentrações de BTEX e etanol na fase aquosa foram medidas através do sistema mostrado na Figura 10, em que foi adaptado um injetor manual de teflon, para a injeção de 500 μ L de cada amostra em linha. Neste sistema tem-se um espectrômetro de massas com um único quadrupolo ABB Extrel (Pittsburgh, PA) que utiliza ionização eletrônica (EI) a 70 eV, uma sonda de MIMS padrão com membrana de silicone 125 μ m de espessura [57,60,61] (Silastic 500-3 da Dow Corning Co.), Figura 11, e uma bomba peristáltica multicanal Ismatec.

Foi utilizado o monitoramento de íons selecionados (SIM) e cada amostra injetada foi bombeada diretamente no sistema, através da sonda de membrana com uma vazão de 2 mL/min. Foram monitorados simultaneamente os íons: etanol (m/z 45); benzeno (m/z 78), tolueno (m/z 92), etilbenzeno e xilenos (m/z 106). Note que para o etilbenzeno e os xilenos, o íon molecular é o de m/z 106 e, neste caso, não podem ser separados.

Soluções aquosas padrão de BTX entre 5 e 40 mg L⁻¹ foram preparadas a partir de diluições de soluções estoques de 100 mg L⁻¹ de BTX solubilizados em metanol. Também soluções de etanol entre 0,75 e 2,0% (V/V) foram preparadas e monitoradas para a quantificação dos compostos orgânicos voláteis, construindo-se desta forma quatro curvas de calibração univariadas

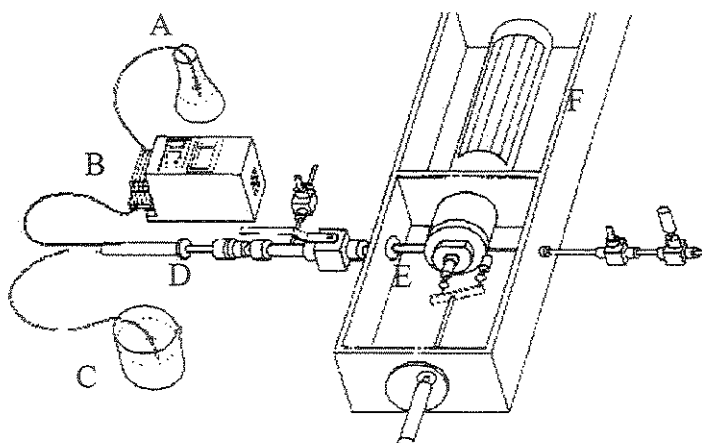


Figura 10. (A) Frasco de amostragem, (B) Bomba peristáltica, (C) Frasco de descarte, (D) Sonda de membrana, (E) Fonte de ionização e (F) Quadrupolo

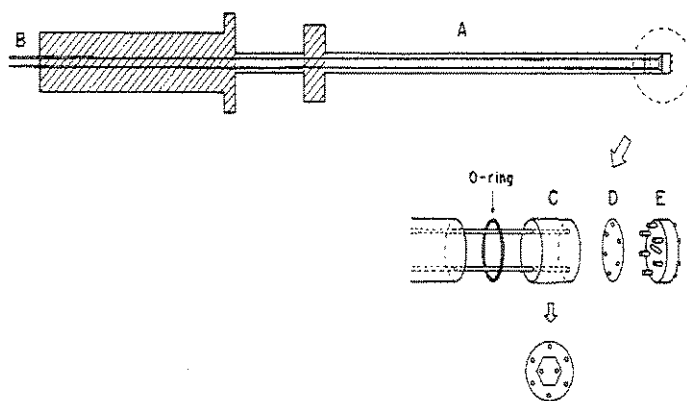


Figura 11. (A) Corpo da sonda, (B) Tubos de circulação de amostra, (C) Cilindro de fixação da membrana, (D) Membrana e (E) Suporte vazado para a fixação da membrana [60].

Estas mesmas amostras foram analisadas através do monitoramento de íons totais (TIM) e um volume de 500 µL de amostra de água contaminada com BTEX e etanol foi injetado no fluxo de água, através de um injetador manual e bombeado através do sistema com vazão de 2 ml/min.

Espectros de massas foram varridos entre m/z 40 e 150, com um incremento de massas de 0,1 unidades de massas, gerando assim uma matriz de dados X (19×1101).

Os espectros de massas foram utilizados para a separação das amostras através da análise exploratória por PCA. A rotina PCA utilizada é do pacote PLS_toolbox [62], escrito em ambiente MATLAB 5.3 [63].

2.2. Resultados e discussão

2.2.1. Quantificação de BTEX e etanol a partir da análise por FIA-MIMS, através da calibração univariada

A técnica FIA-MIMS foi utilizada para a análise e quantificação de BTEX e etanol em água depois de prolongado contato gasolina/água, a partir de soluções padrão de BTX e etanol. Na Figura 12, são mostrados os perfis obtidos a partir do monitoramento SIM dos íons de m/z 45, 78, 92 e 106 para cada uma das soluções padrão e as curvas analíticas obtidas a partir do máximo de intensidade do sinal contra concentração mostraram uma boa linearidade ($R = 0,996 - 0,998$). Assim, o BTEX e o etanol foram simultaneamente determinados em 19 amostras de águas contaminadas por gasolina. A quantificação de etilbenzeno e xilenos, separadamente, não foi possível, pois estes isômeros não apresentam íons seletivos que possam ser monitorados independentemente. Desta forma o íon de m/z 106 foi monitorado e indica a concentração total dos quatro isômeros.

Na Figura 12, existe um indício de que o benzeno é o composto que permeia mais rapidamente pela membrana, depois o tolueno, os xilenos e por último o etanol. Este comportamento está de acordo com a propriedade apolar da membrana, que dificulta a passagem de substâncias mais polares, assim como o coeficiente de difusão da molécula pela membrana, ou seja, moléculas menores permeiam mais facilmente pela membrana. Considerando os espectros de massas dos compostos orgânicos puros, analisados nas misturas, e corrigindo os diagramas de acordo com as intensidades dos sinais, ainda assim a sequência de permeação dos compostos no diagrama não se altera.

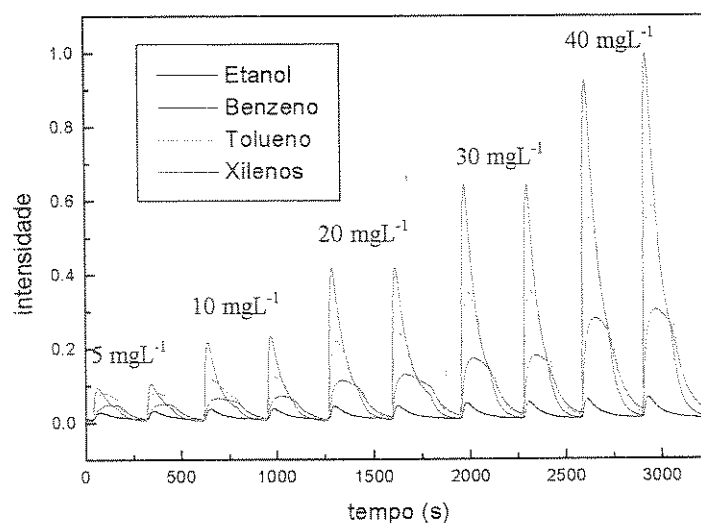


Figura 12. Monitoramento SIM das soluções padrões preparadas para a construção da curva de calibração

Na Tabela 4 estão as faixas de concentração encontradas para o etanol e o BTEX e verifica-se que variam dentro de uma faixa relativamente ampla refletindo, assim, a diversidade em composição quanto a estes compostos nas gasolinas.

Verificou-se que as concentrações em relação aos compostos monitorados nas amostras de água não apresentaram nenhum padrão de contaminação em relação ao tipo de gasolina utilizada para as contaminações (comum, aditivada e premium). Isto porque a solubilização destes compostos em água depende apenas da presença de etanol, que exerce o efeito de co-solência para os mesmos. Este nível de contaminação, no entanto, poderia variar de região para região e período do ano, pois a composição da gasolina pode mudar, dependendo do fornecedor e mudanças nas condições de operação na refinaria.

Tabela 4. Faixas de concentração de BTEX e etanol nas 19 amostras de água contaminadas por gasolina.

contaminantes	faixas de concentração
Benzeno	7,5 – 20,4 mg L ⁻¹
Tolueno	0,9 – 24,9 mg L ⁻¹
Etilbenzeno e Xilenos	4,8 – 19,1 mg L ⁻¹
Etanol	1,2 – 1,5 %

Para o etanol a faixa de concentração entre 1,2 e 1,5% encontrada nas águas contaminadas por gasolinas, corresponde a uma concentração de etanol entre 25,2 e 31,5% nas gasolinas. Estes valores representam uma concentração de etanol, para a maioria das gasolinas, acima do permitido por lei. Isto é um indício que algum outro composto, que gera o mesmo íon de m/z 45, está sendo monitorado junto ao etanol ou ainda uma possível adulteração das gasolinas através da adição de etanol [64].

2.2.2. Análise exploratória a partir da análise por FIA-MIMS através da PCA

As mesmas amostras utilizadas no experimento anterior foram analisadas através do monitoramento TIM e do sistema FIA-MIMS utilizando EI a 70 eV, para que uma análise qualitativa ou exploratória fosse aplicada nos espectros de massas obtidos. Espectros de massas entre m/z 40 e 150 foram monitorados e na Figura 13 tem-se um exemplo de espectro de massas obtido a partir de uma das amostras.

É possível observar todos os compostos orgânicos solubilizados em água e que permearam seletivamente pela membrana. A seletividade, conforme já discutido, é dependente da menor polaridade e maior difusibilidade dos compostos orgânicos voláteis através da membrana de silicone hidrofóbica. No espectro de massas é possível verificar os íons decorrentes dos compostos orgânicos voláteis da gasolina como o BTEX, cujos valores de m/z são 51, 65, 78, 91, 92, 105 e 106 e hidrocarbonetos cíclicos ou insaturados de m/z 41 e 55. O íon de m/z 91, que é bastante intenso, pode ser decorrente do íon tropílio formado tanto da fragmentação do íon de m/z 92 do tolueno, assim como do íon de m/z 106 do etilbenzeno e xilenos. Note que apesar do etanol estar presente em alta concentração nas amostras de águas contaminadas, pois é totalmente miscível, a intensidade de pico é pouco intensa (m/z 45). Isto porque o etanol é mais polar e permeia menos eficientemente através da membrana. Os íons de m/z 41 e 55 que também são bastante intensos são provavelmente decorrentes da perda de um radical CH_3 ou $\text{CH}_3\text{-CH}_2$ de mistura de isômeros de alcenos ou alcanos cíclicos de composições C_5H_{10} (m/z 70). O íon de m/z 44, que também aparece bastante intenso, é proveniente do CO_2 presente na solução aquosa e que permeia pela membrana. Todos os outros íons de menor intensidade presentes no espectro de massas são decorrentes da fragmentação dos compostos orgânicos voláteis da gasolina, mas de menor importância com relação à contaminação de águas.

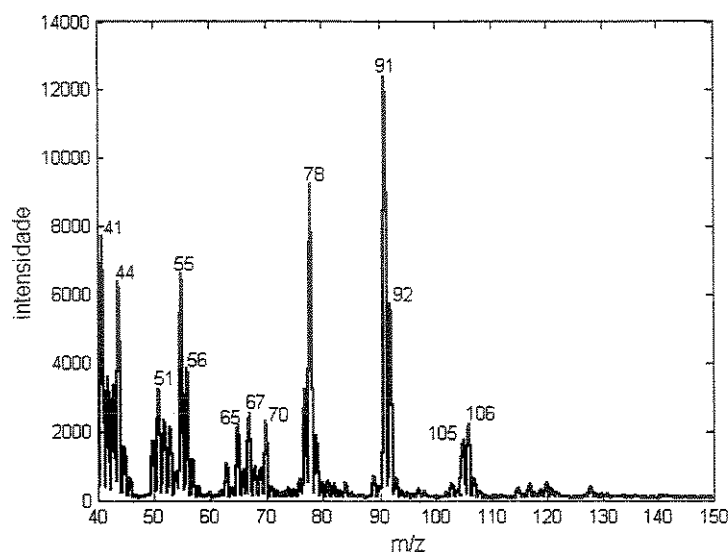


Figura 13. Espectro de massas de uma das amostras de água contaminada com gasolina, no qual estão indicados os íons de maior intensidade.

O método de análise exploratória multivariada por PCA foi empregado para verificar a variação de composição de compostos orgânicos voláteis, nas águas contaminadas por três tipos diferentes de gasolina, (nc) comum, (na) aditivada e (np) premium, onde n é o número da amostra e varia de 1 a 19. Os espectros de massas foram centrados na média e a PCA foi aplicada. Na Tabela 5 estão as porcentagens de variância explicada para os seis primeiros componentes principais e pode ser verificado que PC1 mais PC2 são capazes de explicar aproximadamente 90% da variância dos dados.

Tabela 5. Porcentagem de variância explicada pelos 6 primeiros componentes principais.

componente principal	% variância explicada para cada CP	% variância total explicada
1	75,92	75,92
2	13,27	89,19
3	5,62	94,80
4	2,19	96,99
5	0,96	97,95
6	0,74	98,69

Essa variância definida pelos componentes principais está totalmente correlacionada com a solubilidade dos compostos orgânicos: benzeno (m/z 50, 51 e 78), tolueno (m/z 65, 91 e 92), etilbenzeno mais xilenos (m/z 51, 65, 91, 105 e 106) e de dois compostos de m/z 41 e 55. Isto pode ser confirmado a partir dos gráficos dos pesos para os CP1 e CP2, mostrados na Figura 14, onde é possível verificar que estes íons são os mais importantes para explicar o modelo e separar as amostras.

A Figura 15 corresponde ao gráfico de escores, que representa o novo espaço de variáveis definido pelo primeiro componente principal (PC1) versus o segundo componente principal (PC2), onde as amostras de 1 a 19 estão projetadas. As 19 amostras estão distribuídas nos 4 quadrantes do gráfico (I, II, III e IV), conforme as semelhanças ou correlações descritas pelas suas respectivas variáveis.

O CP1 que explica aproximadamente 76% da variância dos dados está relacionado principalmente com a variação dos íons de m/z 41, 55 e 78, Figura 14(a), e as amostras que apresentam maior abundância destes íons estão nos quadrantes II e IV do gráfico dos escores, Figura 15. O CP2, que explica aproximadamente 13% da variância dos dados, está relacionado principalmente com os íons de m/z 78, 91, 92 e 106, Figura 14(b), ou seja, são amostras com alto teor principalmente de tolueno, etilbenzeno e xilenos e são encontradas nos quadrantes III e IV da Figura 15. Desta forma pela intersecção destes resultados pode-se dizer que as amostras menos contaminadas estão no quadrante I e as amostras mais contaminadas estão no quadrante IV.

A idéia inicial deste trabalho era separar em grupos as amostras de água contaminadas com gasolina, com base nos diferentes tipos de gasolinas (comum, aditivada e premium) encontradas no mercado e utilizadas para a contaminação. Entretanto, verificou-se que não existe nenhuma correlação entre a distribuição das amostras no gráfico de escores e o tipo de gasolina utilizada para a contaminação. Verifica-se uma grande dispersão das amostras na Figura 15, mesmo para aquelas pertencentes ao mesmo tipo (comum, aditivada e premium). Isto porque, a concentração de BTEX e outros compostos que solubilizam na água está relacionada com o processo nas destilarias e a presença de co-solventes como o etanol, adicionado como aditivo em todos os tipos de gasolinas brasileiras [65].

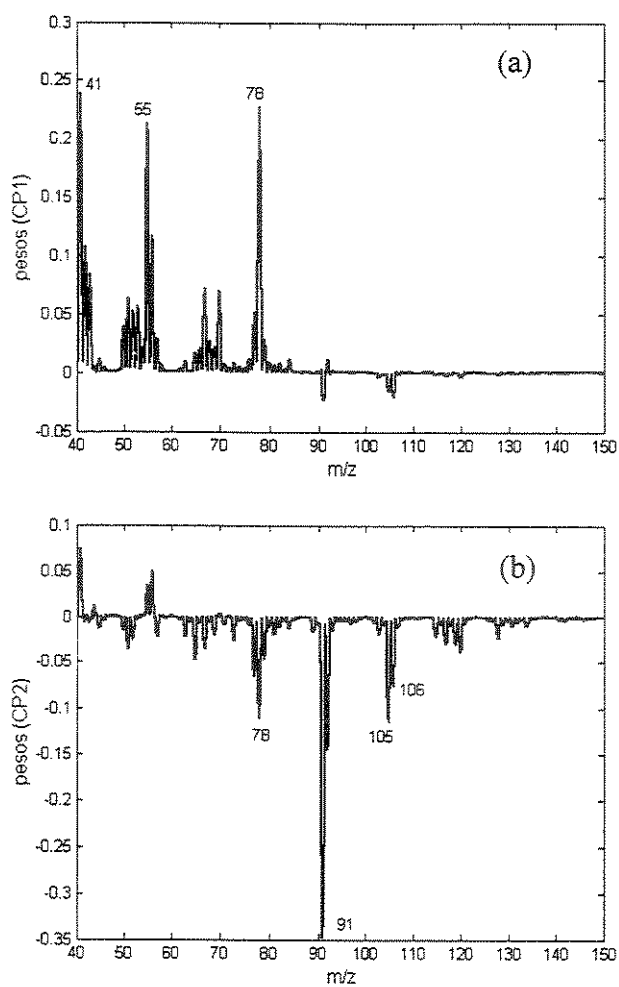


Figura 14. Gráficos dos pesos das variáveis. (a) CP1 (76%); (b) CP2 (13%).

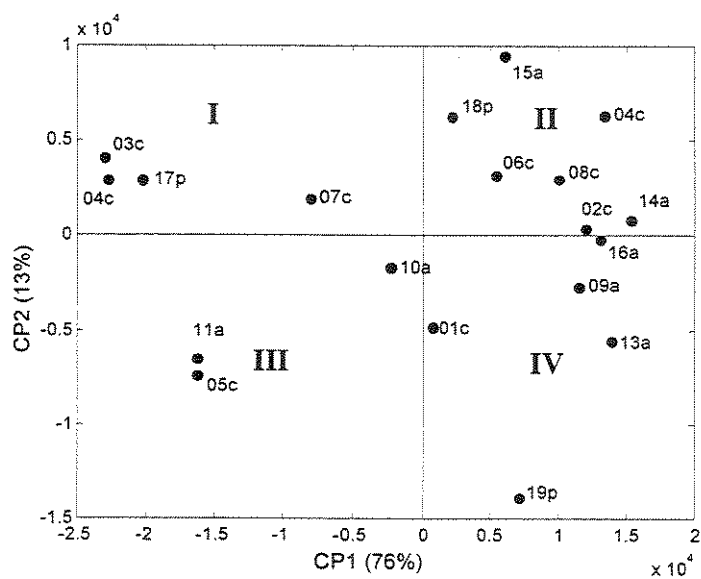


Figura 15. Gráfico dos escores de CP1 versus CP2: c comum, a aditivada; p premium.

A análise por FIA-MIMS verificou uma grande variação em relação à solubilização destes compostos em água a partir das amostras de gasolina. Esta técnica analítica juntamente com a análise quimiométrica foi capaz de detectar rapidamente quais amostras apresentaram maior contaminação (amostras no quadrante IV) e menor contaminação (amostras no quadrante I). Também foi possível identificar quais os compostos orgânicos voláteis que apresentaram maior índice de contaminação por gasolina nas amostras de água. Foi confirmado o alto poder de contaminação dos hidrocarbonetos aromáticos BTEX em águas. Os compostos de m/z 41 e 55 também apresentaram alto poder de contaminação, porém são menos tóxicos que os aromáticos.

Este tipo de análise poderia ser de grande utilidade para detecção rápida, em campo, de contaminação de águas por combustíveis, a partir de vazamentos ou acidentes industriais, indicando com precisão, quais regiões estão mais contaminadas ou menos contaminadas. Também poderia ser muito eficiente e rápido para a identificação de adulterações de combustíveis através da adição de outros tipos de compostos orgânicos voláteis que poderiam ser detectados pela técnica.

3. Análise de cervejas através da técnica MIMS.

A cerveja é uma solução aquosa de composição complexa que contém principalmente CO₂, etanol, vários sais inorgânicos e mais de 800 compostos orgânicos (Tabela 6) [66]. Todos estes compostos são decorrentes de quatro ingredientes essenciais: água, malte (cevada germinada), lúpulo (flor europeia que confere o amargor característico do produto) e leveduras (responsáveis pela fermentação do produto). A diversidade das técnicas de mistura e de fermentação permite obter uma infinidade de cervejas diferentes.

Um grande número de compostos orgânicos voláteis decorrentes da fermentação do açúcar (maltose) pelas leveduras é responsável pelo sabor característico que são específicos das cervejas, entre eles tem-se: álcoois, compostos nitrogenados, ácidos orgânicos, aldeídos, ésteres e cetonas.

A estabilidade do sabor é um fator de qualidade importante das cervejas. Comparado com outros produtos, cervejas apresentam alta estabilidade no sabor e sua degradação a partir de processos oxidativos, produz um sabor característico causado principalmente por uma série de compostos carbonílicos. Muitos esforços têm sido realizados para o controle e melhora no processo e aumento da qualidade do sabor e estabilidade [67-74]. Entretanto, a composição química das cervejas não é totalmente conhecida e muito menos os compostos resultantes da degradação pela

luz, temperatura, agitação e oxigênio. A solução é evitar a exposição das cervejas a estes agentes que induzem as mudanças químicas antes do produto chegar nas mãos do consumidor.

Tabela 6. Composição química típica de cervejas.

Compostos	Concentração	Número de compostos	Fonte
água	90 – 94 %	-	-
etanol	2.5 - 5 %	-	levedo e malte
CO ₂	2.1 - 2.9 %	-	levedo e malte
outros álcoois	0.1 - 0.3 %	15	levedo e malte
carboidratos	1.0 – 5.0 %	≈ 100	malte
sais inorgânicos	0.5 – 0.8 %	26	água e malte
compostos de nitrogênio	0.26 – 0.45 %	≈ 100	levedo e malte
ácidos orgânicos	0.17 – 0.22 %	≈ 200	levedo e malte
aldeídos	30 - 40 mg L ⁻¹	≈ 50	levedo e lúpulo
ésteres e cetonas	0.03 – 0.06 %	≈ 150	levedo e malte
compostos de enxofre	2.0 - 3 mg L ⁻¹	41	levedo e malte
derivados do lúpulo	30 - 60 mg L ⁻¹	>100	lúpulo
vitamina B	5.0 - 11 mg L ⁻¹	13	levedo e malte

A análise sensorial é muito utilizada pelas indústrias de alimento e bebida, entretanto, esta pode ser insuficiente, especialmente quando se trata de problemas de classificação de amostras. A análise multivariada é uma ferramenta que pode ajudar a lidar com a complexidade e sutilezas confrontadas por pesquisadores ao tentar analisar, caracterizar e classificar análises sensoriais como: sabor, aroma, textura e aparência. Pode-se dizer que estas propriedades são multivariadas, ou seja, é uma mistura de sensações que devem ser avaliadas para que possa ser classificado um certo tipo de alimento ou bebida. Entretanto, estas propriedades são subjetivas e podem levar a dúvidas e induzir a erros de avaliação, quando se trata de um grande volume de amostras. Métodos instrumentais baseados em respostas de medidas instrumentais são objetivos e podem ser uma alternativa para verificar a qualidade de um alimento ou bebida [75]. Entre as técnicas mais

utilizadas para este tipo de amostra temos espectrometria de massas, cromatografia gasosa, eletroforese, cromatografia líquida.

Considerando alguns dos pontos discutidos anteriormente sobre as vantagens da espectrometria de massas em relação à cromatografia gasosa, neste trabalho foi realizada análise de amostras de cerveja através da injeção através de membrana no equipamento de espectrometria massas, sem nenhuma preparação como derivatização ou separação cromatográfica. As técnicas MIMS e MIMS/MS foram testadas para a análise da composição química de cervejas e será mostrada a eficiência destas técnicas como um instrumento para a caracterização rápida de cervejas e também a aplicação da análise de componentes principais, como uma alternativa para o controle de qualidade das mesmas.

3.1. Parte experimental

O espectrômetro de massas monoquadrupolar ABB Extrel (Pittsburgh, PA) com ionização por EI a 70 eV e uma sonda de MIMS padrão com membrana de silicone 125 μm de espessura (Silastic 500-3 da Dow Corning Co.) foram utilizados para a análise de cerveja (Figuras 10 e 11). No total, 30 amostras de cervejas em lata foram analisadas, das quais, 13 nacionais e 17 importadas, disponíveis no mercado. Na Tabela 7 estão descritas estas amostras com as respectivas abreviações e prazo de validade. Cada amostra foi diretamente bombeada através da sonda de membrana com vazão de 2 mL/min, por 2 minutos, utilizando uma bomba peristáltica multicanal Ismatec. Os espectros foram obtidos através do monitoramento no modo TIM registrando valores de m/z entre 48 e 150, gerando uma matriz de dados X (30×1021).

Para a separação das amostras foi utilizada a análise exploratória através da análise dos componentes principais (PCA) a partir dos espectros de massas obtidos. A rotina PCA é um pacote do PLS_toolbox [62] escrito em ambiente MATLAB 5.3 [63].

Tabela 7. Amostras de cerveja analisadas por MIMS

amostras	cerveja	abreviações	validade
1	Antartica	Ant	08/03/2001
2	Bavaria	Bav	22/03/2001
3	Bohemia	Boh	04/01/2001
4	Brahma	Bra	01/03/2001
5	Brahma light	Brl	
6	Kaiser	Kai	23/01/2001
7	Kaiser Bock	Kab	24/03/2001
8	Cintra	Cin	08/02/2001
9	Crystal	Cry	06/12/2000
10	Caracu	Car	16/12/2001
11	Schincariol	Scn	10/12/2000
12	Schincariol preta	Sci	22/02/2001
13	Skol	Sko	30/03/2001
14	Calsberg	Cal	27/10/2000
15	Warsteiner	War	30/03/2001
16	Veltins	Vel	07/12/2000
17	Gold cap	Goc	31/10/2000
18	Guinness	Gui	29/11/2000
19	Heineken	Hei	24/08/2000
20	Horse	Hor	31/11/2000
21	Corona	Cor	04/02/2001
22	Budweiser	Bud	12/01/2001
23	Boss	Bos	31/09/2000
24	Red Lion	Lio	31/10/2000
25	Pheonix	Pho	31/09/2000
26	Double Dutch	DD	31/09/2000
27	Gold	Gol	31/11/2000
28	Urquell	Urq	28/01/2001
29	Scream'n beaver	Scr	31/10/2000
30	Summer	Sum	09/09/2000

3.2. Resultados e discussão.

Como descrito anteriormente, mais de 800 compostos químicos são identificados nas cervejas e muitos estão presentes em baixas concentrações, dos quais somente aqueles que

possuem um sabor mais acentuado podem influenciar no sabor das mesmas. A maioria destes compostos está presente na matéria prima que constituem as cervejas e outros são produzidos durante a fermentação. Entretanto, existem alguns que são produzidos durante o armazenamento antes de ser consumido, entre eles estão alguns álcoois, ésteres e aldeídos. Eles ocorrem devido a algumas reações enzimáticas que ainda podem ocorrer depois do processo de envasamento.

Muitos destes compostos orgânicos voláteis das amostras de cerveja permearam eficientemente através da membrana de silicone e produziram uma série de íons característicos, depois de ionizados por EI a 70 eV e foram detectados pela espectrometria de massas. A análise de cerveja a partir da técnica MIMS foi bastante rápida (5 min/amostra) e mostrou alta sensibilidade em relação a alguns destes compostos que apresentaram melhor solubilidade e difusibilidade pela membrana, pois a maioria dos compostos orgânicos voláteis presentes na cerveja como álcoois, ésteres, aldeídos, cetonas, etc, são de caráter polar e apresentam massa molar relativamente alta [66].

Nas Figuras 16 e 17 são apresentados dois espectros de massas característicos das amostras de cervejas analisadas e que pertencem a dois grupos extremos dentro do conjunto de dados. Na Figura 16 é mostrado o espectro de massas de uma das amostras (Heineken) que apresentou sabor e aroma desagradáveis, indicando sua degradação durante o seu armazenamento. Esta amostra estava com o prazo de validade vencido, pois sua validade foi até o dia 24/08/2000 e as análises foram realizadas no dia 26/09/2000. Picos intensos para os íons de m/z 55, 70, 88, 91 e 104 estão presentes. O espectro de massas da Figura 17 exemplifica uma significativa mudança nas abundâncias relativas dos compostos orgânicos voláteis nas amostras de cervejas que não apresentavam nenhum indício de degradação (Skol), principalmente os íons de m/z 70 e 104, que apresentam uma grande variação de intensidade em relação ao íon de m/z 91. Note que os espectros de massas foram varridos de m/z 48 a 150, eliminando os picos do etanol (m/z 45 e 46) que está presente em uma concentração de $\approx 5\%$. O objetivo deste trabalho foi detectar apenas os compostos orgânicos voláteis que na sua grande maioria estão em baixa concentração e que determinam o sabor e aroma das cervejas.

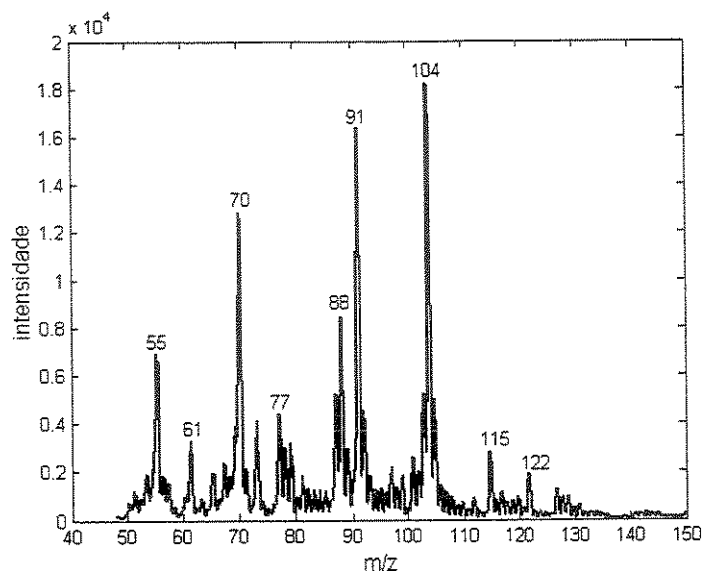


Figura 16. Espectro de massas da cerveja Heineken

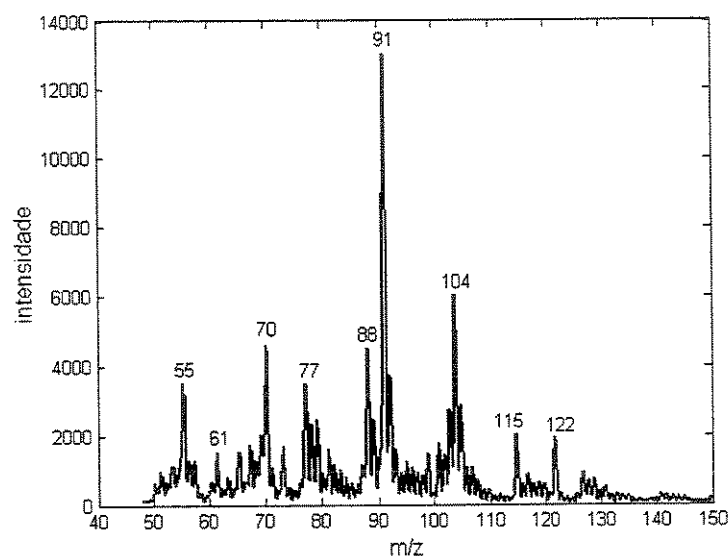


Figura 17. Espectro de massas da cerveja Skol

Apesar da complexidade da composição química das cervejas (Tabela 6) é possível, a partir dos espectros de massas obtidos, atribuir alguns dos íons detectados a constituintes das cervejas. Os íons de m/z 122 e 91 poderiam ser atribuídos ao composto orgânico 2-feniletanol, como mostrado na Figura 18 (a). Este composto é um constituinte de cervejas bem conhecido, está presente na faixa de 8 a 35 ppm, e é considerado um marcador de controle de qualidade [66].

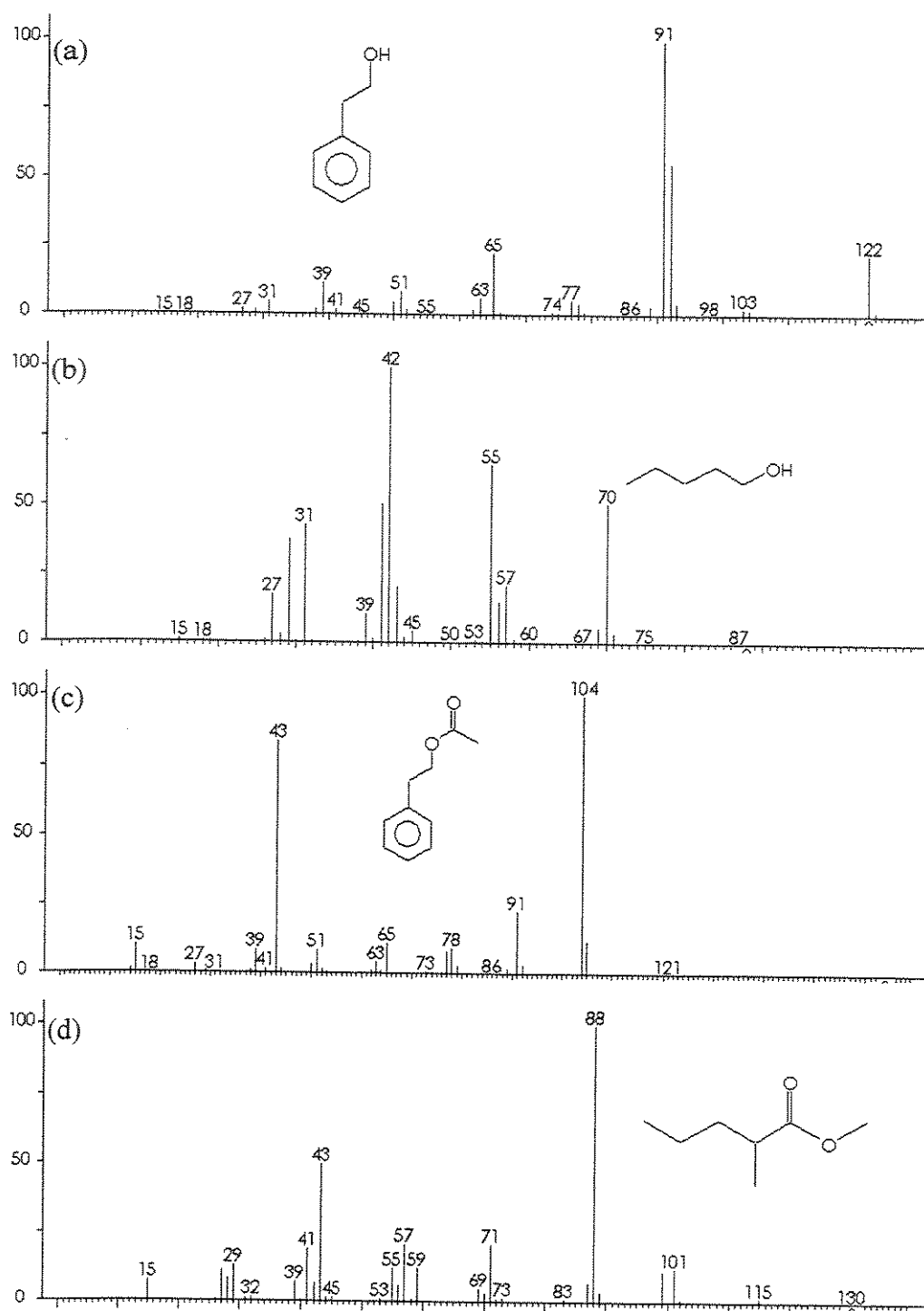


Figura 18. Espectros de massas obtidos por EI a 70 eV para alguns dos compostos orgânicos voláteis mais comuns em cerveja. (a) 2-feniletanol; (b) n-pentanol; (c) acetato de feniletila; (d) 2-metil pentanoato de metila

Similarmente os íons de m/z 70 e 55 podem ser atribuídos a vários álcoois de massa molar baixa como n-pentanol, mostrado na Figura 18 (b). O íon de m/z 104 pode ser decorrente de uma série de ésteres do 2-feniletanol como o acetato de feniletila, mostrado na Figura 18 (c). O m/z 88 poderia ser atribuído ao metil éster decorrente do 2-metil pentanoato de metila, Figura 18 (d).

Todos estes compostos estão presentes em baixa concentração na maioria das cervejas e são prováveis candidatos para certos íons presentes nos espectros de massas obtidos por MIMS. Estes compostos podem indicar as diferentes origens para cada espectro de massas, a partir das diferentes marcas de cervejas analisadas.

Depois de realizada esta breve avaliação nos dados, os espectros de massas foram centrados na média e utilizou-se a análise de componentes principais (PCA) para uma melhor separação das amostras, baseada nas diferenças de intensidade de picos. Na Tabela 8 estão as variâncias explicadas para os 7 primeiros componentes principais e verifica-se que os componentes principais 1, 2 e 3 explicam mais de 93% dos dados.

Tabela 8. Porcentagem de variância explicada pelos 7 primeiros componentes principais.

componente principal	% variância explicada para cada CP	% variância total explicada
1	73,02	73,02
2	11,32	84,34
3	9,38	93,72
4	1,56	95,28
5	1,29	96,57
6	0,85	97,42
7	0,58	98,00

Foi possível verificar através da análise por PCA, que os íons de m/z 55, 70, 88, 91 e 104 são as variáveis de maior peso para descrever o conjunto de amostras. Isto pode ser conferido nos gráficos dos pesos das variáveis a partir dos componentes principais CP1, CP2 e CP3, mostrados na Figura 19, ou seja, são principalmente esses íons que diferenciam as cervejas.

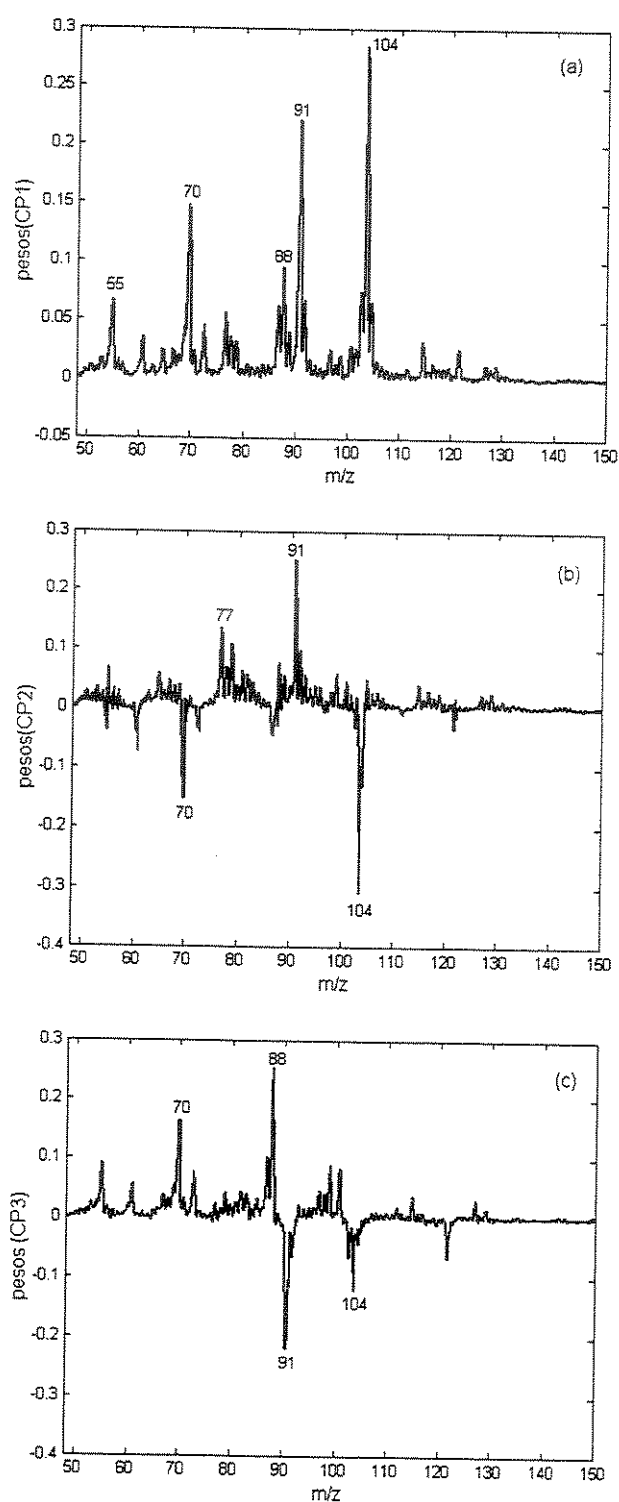


Figura 19. Gráfico dos pesos das variáveis para (a) CP1 (73%), (b) CP2 (11%) e (c) CP3 (9%).

No gráfico de escores da CP1 versus a CP2, mostrado na Figura 20, as cervejas estão separadas principalmente de acordo com a abundância dos íons de m/z 70, 91 e 104. No quadrante

I estão as amostras com valores de intensidade de pico baixo para os íons de m/z 104 e 70, mas alto para o íon de m/z 91, como mostrado na Figura 17. Nos quadrantes, II e III, estão as amostras com os picos dos íons de m/z 104 e 70 um pouco mais altos, mas que ainda o íon de m/z 91 é o mais intenso. No último quadrante, IV, as amostras de cerveja apresentam os íons de m/z 104 e 70 bem mais intensos, como mostrado na Figura 16.

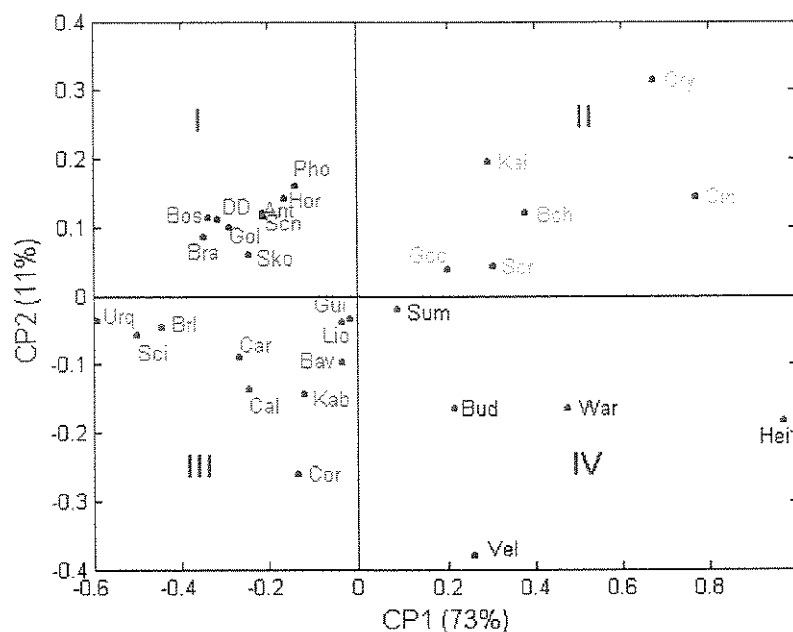


Figura 20. Gráfico dos escores da CP1 (73%) contra a CP2 (11%)

Entre as amostras que apresentam intensidade alta do íon de m/z 104 em relação ao íon de m/z 91, (quadrante IV), estão aquelas em que não foi adicionado antioxidante (vitamina C), segundo informações na embalagem, no processo de fabricação, como: “Hei”, “War”, “Bud” e “Vel”. A amostra “Hei”, mais afastada das outras amostras e que apresenta o íon de m/z 104 bem intenso, estava com o prazo de validade vencido, como pode ser verificado na Tabela 7, e com odor e sabor desagradáveis.

Com base nestes resultados experimentais pode-se dizer que o aumento das intensidades dos íons de m/z 104 e 70 em relação ao íon de m/z 91, pode ser um forte indício da degradação da cerveja, pois foi este comportamento que caracterizou a amostra que apresentou visíveis sinais de degradação.

O íon de m/z 88 que tem alto valor de peso na CP3 apresentou valores elevados de escores somente para duas amostras “Ant” e “Bra”, todas as outras amostras apresentaram valores

parecidos. Estas duas amostras de cerveja aparentemente não mostraram nenhuma anomalia em relação ao sabor ou odor. Portanto, o composto orgânico que gera o íon de m/z 88 deve ser um subproduto do processo de fermentação que não causa nenhum efeito aparente na qualidade da cerveja.

Para melhorar a seletividade e caracterização das cervejas, espectros de massas através da técnica MIMS/MS foram adquiridos. Três íons precursoros foram selecionados, os íons de m/z 122, 104 e 88 e os espectros de massas MS/MS obtidos estão mostrados na Figura 21.

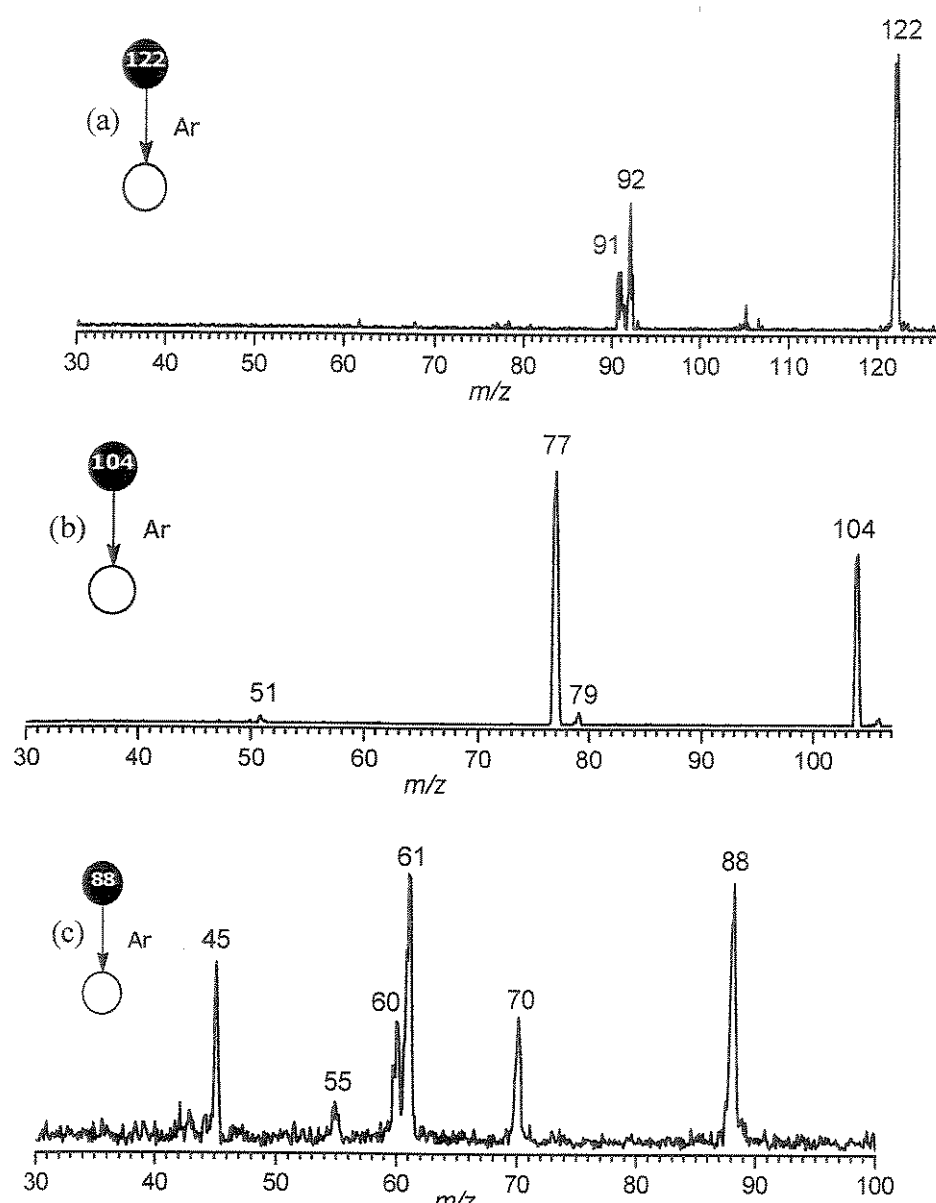


Figura 21. Espectros de massas MS/MS obtidos a 15 eV por CID a partir de: (a) m/z 122; (b) m/z 104; (c) m/z 88

Estes espectros foram obtidos pela análise sequencial utilizando a dissociação por CID no espectrômetro de massas pentaquadrupolar descrito no capítulo 1. Os íons de interesse foram selecionados em Q_1 , dissociados a 15 eV através da colisão com argônio em q_2 e seus produtos foram selecionados em Q_3 , enquanto q_4 e Q_5 foram mantidos com uma radio frequência que permite a passagem dos produtos.

Entre os íons selecionados está o de m/z 104, um dos mais importantes na separação das cervejas por PCA, e também os íons de m/z 88 e 122, que como indicam os espectros de MS/MS da Figura 21 (a) e (c), são responsáveis por gerar os outros dois íons de m/z 91 e 70 que também foram determinados como íons importantes para a separação dos diferentes grupos de cerveja analisadas.

O 2-feniletanol de m/z 122 está presente nas cervejas e como pode ser verificado na Figura 21 (a) gera o íon de m/z 91, a partir da dissociação por CID. Entretanto este íon pode ser gerado também a partir de vários outros compostos orgânicos voláteis como, ácidos orgânicos aromáticos ou fenólicos, assim como álcoois fenólicos ou aldeídos encontrados em baixa concentração em cervejas e que permearam pela membrana [66].

O íon de m/z 104 produz principalmente o íon de m/z 77, decorrente do anel aromático da molécula, como mostrado na Figura 21 (b). O íon de m/z 104, como verificado a partir da análise dos componentes principais, está provavelmente relacionado com algum produto de degradação das cervejas. É possível que este íon seja decorrente da fragmentação de algum éster produzido a partir do 2-feniletanol durante a degradação da cerveja, como por exemplo o acetato de feniletila, ou também de ésteres produzidos durante o processo de fermentação a partir de certos tipos de leveduras utilizadas e que influenciam na qualidade das cervejas [66].

O íon de m/z 88 produz a partir da dissociação por CID o íon de m/z 70, 61, 60, 55 e 45, como mostrado na Figura 21(c). Este íon pode estar sendo gerado a partir de dois compostos, o metil éster e ou acetato de etila. Os dois compostos podem estar gerando o íon de m/z 70, que como verificado pela análise por PCA é intenso na cerveja que apresentou indícios de degradação. O íon de m/z 88 decorrente do metil éster, não tem nenhuma relação com a degradação das cervejas pois, em amostras que este íon apresentou alta intensidade ("Ant" e "Bra") não foi constatado nenhum problema com relação ao sabor e odor. O íon de m/z 88 decorrente do acetato de etila é um dos ésteres produzidos a partir do mosto no processo de fabricação das cervejas e que não é desejado no produto final [66] e provavelmente está gerando o m/z 70 que foi importante para a separação das cervejas.

Desta forma pode-se dizer que embora as variáveis que causam a degradação da cerveja sejam bem conhecidas, como luz, temperatura, agitação, oxigênio e matéria prima, os efeitos desta degradação são muito mais complexos. Entretanto, através da análise de MIMS convencional e a análise exploratória pela análise dos componentes principais, foi possível separar as amostras de acordo com o tempo de validade, adição ou não de antioxidante a partir das diferentes marcas de cerveja analisadas. Isto porque, todas as cervejas que apresentaram alta intensidade do íon de m/z 104, no gráfico de escores, estavam com o prazo de validade vencido ou quase vencido e em alguns casos não continham antioxidante em sua composição. As cervejas brasileiras normalmente contêm antioxidantes para evitar a degradação, decorrente das altas temperaturas ambientais, até que o produto chegue nas mãos do consumidor. Normalmente, nas cervejas européias estes antioxidantes não são adicionados.

Mudanças de composição em relação aos compostos detectados pela espectrometria de massas podem também estar relacionadas não só com o processo de armazenamento, como também com o próprio processo de fabricação (fermentação e maturação) ou matéria prima utilizada (malte, lúpulo e água). Devido a todos estes fatores, o sabor das cervejas se mantém estável, por períodos completamente diferentes entre um tipo de cerveja para outro, assim como de uma marca para outra e um estudo mais detalhado poderia ser feito para determinar com precisão, quais são os compostos responsáveis por estas variações. Entretanto, não foi este o objetivo deste trabalho e as amostras foram apenas separadas de acordo com os espectros de massas obtidos.

Pode-se afirmar a partir destes experimentos, que a técnica MIMS tem um futuro promissor como método de análise para o controle de qualidade. Esta técnica em conjunto com análise por PCA pode ser utilizada para verificar rapidamente mudanças na composição de cervejas durante o processo de fabricação ou de armazenamento. Isto porque, estas variações na composição podem ser reconhecidas facilmente através da análise exploratória por PCA, o que sem a ajuda da quimiometria dificilmente seriam encontradas.

4. Conclusão

Nos experimentos realizados com amostras de gasolina e amostras de cerveja, foram obtidos espectros de massas bastante complexos devido à composição química das amostras nos dois casos. O monitoramento SIM juntamente com a técnica FIA-MIMS mostrou ser eficiente e

rápido para a quantificação de compostos orgânicos voláteis, presentes em águas contaminadas por gasolina. Também o monitoramento TIM juntamente com a análise exploratória por PCA, foi eficiente para verificar quais os constituintes da gasolina que mais solubilizam em água potável, depois de um contato prolongado (gasolina/água). Foi verificado que compostos como BTEX, etanol e outros compostos que geram íons de m/z 41 e 55 estão presentes em altas concentrações nas águas contaminadas e variam bastante em composição. Entretanto, esta variação em composição não tem nenhuma relação com o tipo de gasolina (comum, aditivada e premium) analisada.

A espectrometria de massas tem mostrado ser uma técnica analítica bastante promissora em análise de rotina através da técnica MIMS. Neste trabalho a técnica MIMS foi testada pela primeira vez como um método para a análise da composição química de cervejas. Foi possível verificar a eficácia da técnica MIMS (e MIMS/MS) como um instrumento rápido para a caracterização de espectros de cerveja. Dos resultados obtidos por MIMS para a análise de cerveja juntamente com a análise por PCA, pode-se dizer que é uma alternativa mais rápida no controle de qualidade, no qual alterações em cervejas durante o processo de fabricação ou ainda no armazenamento do produto podem ser detectadas.

Capítulo 4

Pré-processamento do ruído heteroscedástico

1. Introdução

Uma das etapas mais importantes na aplicação de métodos de calibração multivariada é a etapa de pré-processamento dos dados. Em quase todas as técnicas analíticas é impossível utilizar a calibração multivariada sem o pré-processamento dos dados, que envolve a aplicação de um filtro para eliminar ou minimizar a influência do ruído no desenvolvimento do modelo de calibração. Estes filtros são baseados em técnicas matemáticas que variam conforme o tipo de dado utilizado e a escolha adequada deste filtro é de extrema importância no pré-processamento. Simples artifícios matemáticos são muito utilizados no processo de filtragem de dados. Um exemplo muito comum é a subdivisão do sinal analítico em pequenos conjuntos de pontos ("janelas") e a partir destes encontrar um novo conjunto contendo as médias destes subconjuntos. O resultado a partir dos valores médios é uma matriz de dados com menos pontos e com menor flutuação de sinal. Também outras técnicas matemáticas mais complicadas como, transformada de Fourier [76] e suavização por derivatização [77], podem ser aplicadas para eliminar ruído ou desvio de linha bases.

Como pode ser verificado, existem muitos métodos de pré-processamento e deve-se selecionar o mais adequado para os dados em questão. Antes, porém, de responder esta pergunta deve-se primeiro saber qual o tipo de ruído que está presente nos dados e como é a sua distribuição.

Existem basicamente dois tipos de ruído, o homoscedástico e heteroscedástico [78-80]. O ruído homoscedástico é representado na Figura 22 e foi obtido a partir de um espectro simulado. Como pode ser verificado o sinal analítico apresenta um nível de ruído constante em toda sua extensão. Este tipo de ruído é menos problemático quando utilizados métodos de calibração multivariada como PLS, N-PLS, MCR etc, pois estes métodos quimiométricos têm como propriedade eliminar parte do sinal que apresenta comportamento sistemático, ou seja, ruído constante.

O ruído heteroscedástico, em contraste ao homoscedástico, não apresenta um comportamento sistemático na extensão do sinal analítico e por isso é um tipo de ruído mais difícil de ser eliminado pelos métodos quimiométricos existentes. Como ruído heteroscedástico enquadram-se todos os tipos de ruído que não apresentam uma variação constante na superfície do sinal. Na Figura 23 é mostrado um caso particular de ruído heteroscedástico e como pode ser observado, o

ruído é proporcional à intensidade do sinal, ou seja, quanto maior a intensidade de pico maior o ruído presente [22-24].

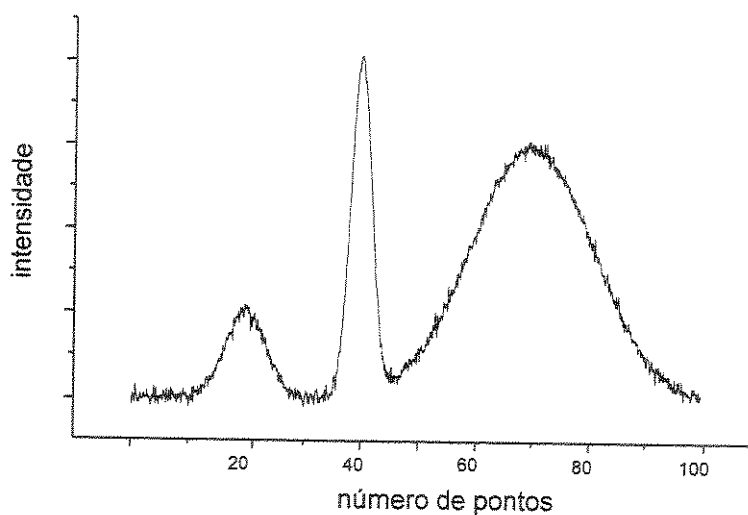


Figura 22. Representação do ruído homoscedástico em um sinal analítico simulado

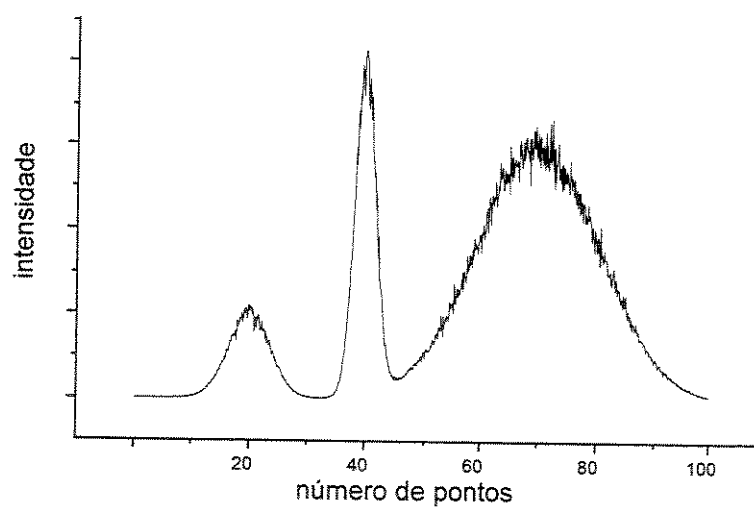


Figura 23. Representação do ruído heteroscedástico em um sinal analítico simulado

Como já discutido no capítulo 1, a espectrometria de massas apresenta ruído heteroscedástico, devido ao funcionamento do detector e a etapa de pré-processamento, neste caso, é de extrema importância no processo de quantificação de amostras. Diferentemente da maioria das técnicas analíticas como espectrometria no ultravioleta, visível ou infravermelho, em que cada sinal

analítico é definido por um grande número de pontos, nos espectros de massas cada pico é definido por um conjunto reduzido de pontos, o que dificulta a aplicação dos filtros citados anteriormente.

Em alguns trabalhos encontrados na literatura [78,79] é discutida a aplicação da transformação logarítmica, como solução do problema de ruído heteroscedástico, nos casos onde o ruído é proporcional ao sinal analítico. Este tipo de estratégia matemática é viável em dados de espectrometria de massas, pois é aplicada a função logarítmica em cada ponto do espectro. Entretanto, a função logarítmica é uma transformação não linear que implica algumas vantagens e desvantagens que serão discutidas nos próximos capítulos.

2. Parte experimental

Espectros de massas foram obtidos para investigar a reprodutibilidade e distribuição do ruído no sinal analítico. Uma série de dez replicatas foi analisada a partir do composto padrão perfluorotributilamina (FC43), normalmente utilizado na calibração de equipamentos de espectrometria de massa. Isto porque, este composto apresenta picos de intensidade baixa, média e alta, fornecendo assim uma boa distribuição de intensidade de picos para serem tomadas como referência.

Cada amostra foi preparada em ampolas que foram acopladas ao equipamento de espectrometria de massas pentaquadrupolar Extrel (Pittsburgh, PA). Este consiste de três quadrupolos analisadores de massas (Q_1 , Q_3 , Q_5) e de dois quadrupolos de reação ou dissociação (q_2 e q_4), já descritos anteriormente no capítulo 1. Através do alto vácuo, o vapor de amostra foi transferido para dentro do espectrômetro de massas prosseguindo então a análise.

Foram obtidos espectros de massas bidimensionais (MS), Figura 24, e tridimensionais (MS/MS), Figura 25.

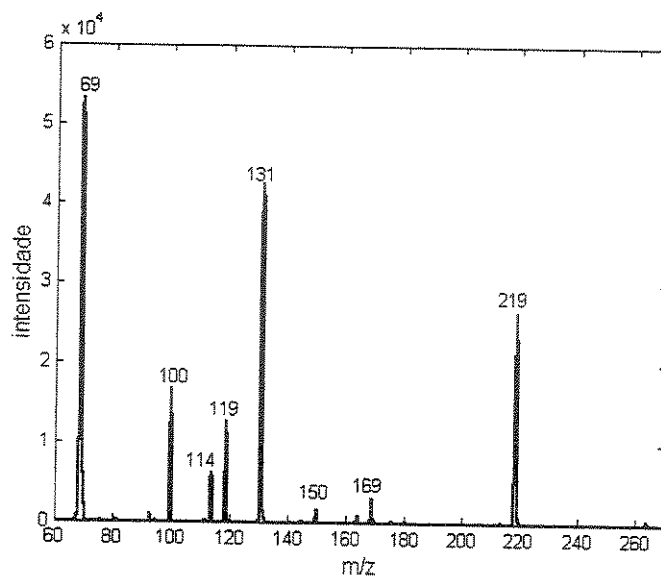


Figura 24. Espectro de massas bidimensional obtido a partir do FC43.

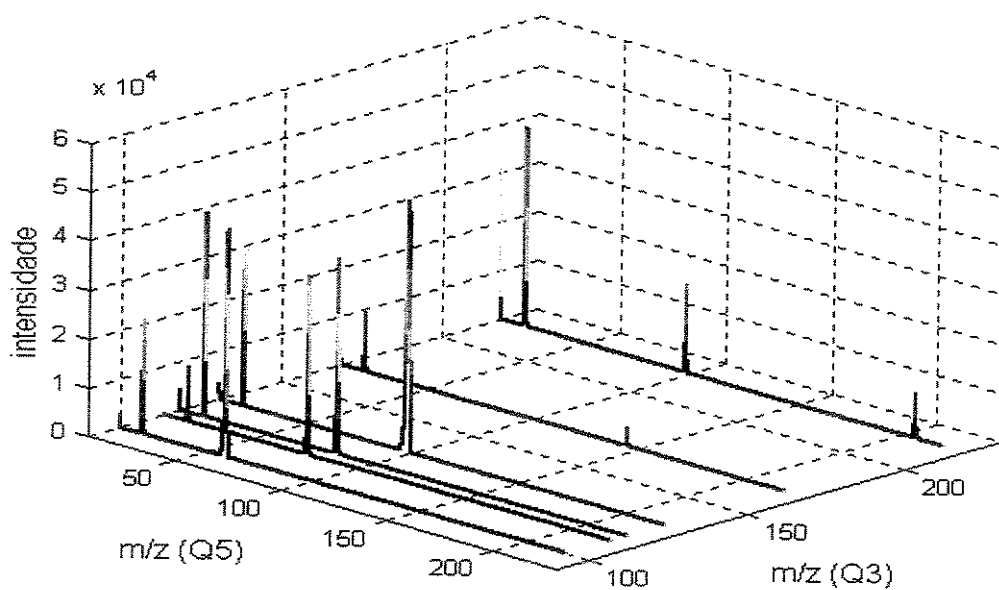


Figura 25. Espectro de massas tridimensional obtido a partir do FC43.

Os dados bidimensionais foram obtidos a partir da simples ionização por elétrons a 70 eV, utilizando o quadrupolo Q₅ na seleção dos íons entre m/z 60 a 270. Os dados tridimensionais foram obtidos pela análise seqüencial, utilizando o quadrupolo Q₃ na seleção dos íons precursores gerados através da ionização por elétrons a 70 eV. O quadrupolo q₄ foi utilizado na dissociação por CID com

argônio a 15 eV. Os íons produzidos foram selecionados em Q₅ em uma faixa de íons de m/z 60 a 230. A pressão no equipamento durante a aquisição dos dados tridimensionais foi mantida aproximadamente em 9×10^{-5} Torr.

Cada espectro de massas obtido é a média de cinco espectros registrados sucessivamente e cada um dos cinco espectros é obtido em uma varredura, em que cada ponto do espectro é obtido após várias leituras do detector no tempo de integração selecionado (tempo de integração em dois ciclos 33,3 ms). Este procedimento foi utilizado para reduzir o ruído nos espectros de massas adquiridos.

Os espectros tridimensionais foram gravados na forma binária pelo equipamento, para reduzir espaço em disco, e assim foi necessário transformá-los para o formato ASCII, para que os dados obtidos pudessem ser tratados utilizando rotinas em MATLAB.

3. Resultados e discussão

A espectrometria de massas é normalmente utilizada para trabalhos qualitativos ou quantificação usando monitoramento de íons selecionados (SIM). Os espectros de massas, quando analisados qualitativamente são submetidos a um pré-processamento mínimo, por exemplo, uma simples normalização do espectro de massas pelo pico máximo. Para utilizar os espectros de massas como uma medida quantitativa, outros cuidados devem ser tomados com relação à otimização ou pré-processamento dos dados. Estes cuidados com relação ao pré-processamento devem ser maiores ainda quando misturas com um grande número de analitos, ou com espectros muito similares são analisadas.

A grande desvantagem da espectrometria de massas é a baixa reprodutibilidade do sinal analítico, decorrente de diferentes fatores fundamentais e ambientais como: variações termodinâmicas, variações na pressão e temperatura e vibrações mecânicas [26]. Estes fatores resultam em um sinal contaminado com um nível significativo de ruído que pode ser medido e separado.

Um primeiro tipo de ruído é aquele em que toda a extensão do espectro de massas é influenciada por uma constante e pode variar de acordo com a equação abaixo.

$$X = (1 + e)\tilde{X} \quad (14)$$

onde $\mathbf{X}$ é o sinal medido e $\tilde{\mathbf{X}}$ é o sinal verdadeiro para uma amostra, ambos tendo dimensões $I \times J$. O termo ruído escalar, e , tem um efeito multiplicativo em todo o espectro e varia de medida para medida. Este tipo de ruído pode ser tratado aplicando algum tipo de normalização como: normalização por um padrão interno, pelo pico máximo ou a soma quadrada [22,23].

Um segundo tipo de ruído é aquele presente em cada variável, que pode ser expressa por:

$$x_{ij} = \tilde{x}_{ij} + e_{ij} \quad (15)$$

onde x_{ij} e $\tilde{x}_{ij}$ são os sinais medidos e o sinal esperado no ponto ij do espectro de massas, onde i representa o número de amostras e j representa cada variável. O termo ruído, e_{ij} , pode ter um desvio padrão constante (ruído homoscedástico) ou diferente a cada ponto do espectro (ruído heteroscedástico). No caso onde o nível de ruído é dependente da intensidade de pico, por exemplo, $e_{ij} = \pm \alpha \tilde{x}_{ij}$, é conhecido como ruído heteroscedástico proporcional. Entretanto, os modelos de calibração multivariada geralmente atribuem pesos iguais para os resíduos em cada variável e desta forma é preciso transformar o ruído para uma forma aproximadamente uniforme em toda a extensão do espectro, antes de aplicar os métodos de calibração.

Como observado por Kvalheim e colaboradores [78], assim como Brakstad e outros [79,80] é difícil separar estes dois tipos de ruídos, resultando em uma situação complexa. Por exemplo, picos com uma alta intensidade têm maior influência durante a normalização do espectro pelo pico máximo, por exemplo, devido a heteroscedasticidade. Isto, porque o ruído em picos com alta intensidade é transferido para picos menores, neste tipo de normalização, resultando na distorção dos dados [78]. Por esta razão é necessário utilizar a transformação logarítmica para remover o ruído heteroscedástico antes da normalização espectral. No caso específico de ruído heteroscedástico proporcional, a equação 15 pode ser escrita como:

$$\begin{aligned} x_{ij} &= \tilde{x}_{ij} \pm \alpha \tilde{x}_{ij} \\ &= (1 \pm \alpha) \tilde{x}_{ij} \end{aligned} \quad (16)$$

Aplicando o logaritmo neperiano ($\ln$) na equação 16 temos:

$$\begin{aligned}\ln(x_{ij}) &= \ln((1 \pm \alpha)\tilde{x}_{ij}) \\ &= \ln(1 \pm \alpha) + \ln(\tilde{x}_{ij})\end{aligned}\tag{17}$$

Considerando que $\ln(1 \pm \alpha)$ é uma constante, pode-se dizer desta forma que o ruído proporcional ao sinal (ruído heteroscedástico), foi transformado em ruído homoscedástico pela função logarítmica e será facilmente removido pela calibração multivariada.

Visualizar esta transformação em dados de espectrometria de massas não é possível, pois como já discutido acima, os espectros de massas apresentam picos definidos por apenas 10 pontos. Entretanto, uma outra forma de verificar o comportamento do ruído é através do desvio padrão entre replicatas, no qual um desvio padrão constante indica presença de ruído homoscedástico e desvio padrão proporcional ao sinal indica a presença de ruído heteroscedástico.

Depois de obtidos os espectros de massas, bidimensionais e tridimensionais das 10 replicatas de FC43, foi calculado o desvio padrão do sinal entre as replicatas, nos dois conjuntos de dados e os espectros de massas médios decorrentes das 10 replicatas. A partir destes valores foram obtidos dois gráficos, um para cada tipo de espectro adquirido (bidimensional e tridimensional). Verificou-se o mesmo comportamento para os dois tipos de espectros com relação ao desvio padrão e a intensidade do sinal médio. Na Figura 26 é mostrado o resultado para os espectros de massas tridimensionais que por apresentar mais picos com intensidades diferentes, traduz melhor o comportamento do ruído heteroscedástico nos espectros de massas.

Verifica-se claramente que o ruído heteroscedástico está presente, pois, o desvio padrão entre as replicatas aumenta com a intensidade do sinal, ou seja, o ruído é aproximadamente proporcional à intensidade do sinal, uma situação, no qual, a transformação logarítmica pode ser apropriada [78,79].

Na Figura 27 tem-se o desvio padrão do sinal analítico versus a intensidade do sinal médio, para as 10 replicatas de FC43 depois de aplicada a transformação logarítmica. Como pode ser verificado, o intervalo do sinal com valores de intensidade de pico entre 0,5 e $3,5 \times 10^4$ do gráfico da Figura 26 caem na faixa entre 8 e 10,5 no gráfico da Figura 27. Esta região mostra que o desvio padrão praticamente não muda com a intensidade de pico médio entre as replicatas.

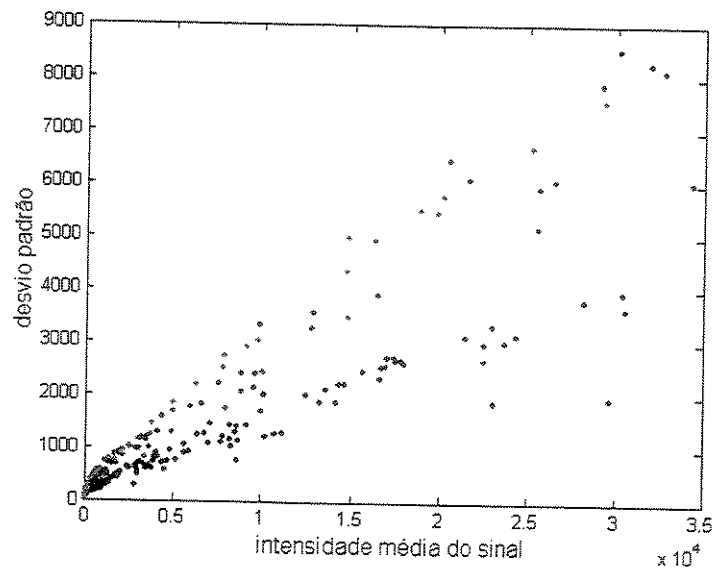


Figura 26. Desvio padrão versus intensidade média do sinal para os espectros de massas das 10 replicatas do FC43.

Verifica-se desta forma que a influência do ruído heteroscedástico desaparece nas posições dos picos com intensidades mais altas, ou seja, o desvio padrão diminui drasticamente em sinais com alta intensidade de pico, transformando-o em ruído com um comportamento quase que perfeitamente homoscedásticos nesta região. Esta correção só é viável se for verificada a presença intensa de ruído heteroscedástico em toda a extensão do sinal.

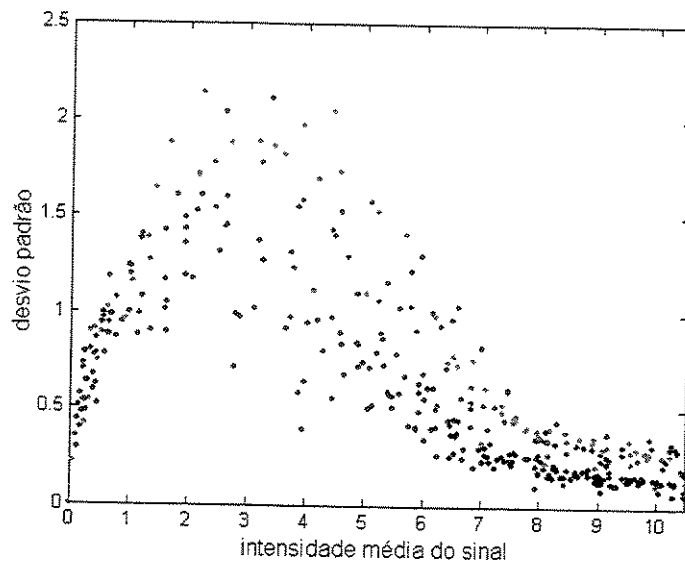


Figura 27. Desvio padrão versus intensidade média do sinal para os espectros de massas das 10 replicatas do FC43, depois de aplicada a transformação logarítmica.

Como desvantagem da transformação logarítmica é que esta destrói a linearidade do sinal analítico com a concentração, podendo prejudicar a quantificação das amostras. Desta forma pode-se dizer que o tipo de pré-processamento utilizado vai depender do tipo ou distribuição do ruído e do nível de ruído presente nos dados. Esta escolha naturalmente vai variar de instrumento para instrumento e de conjunto de dados para conjunto de dados.

4. Conclusão

Foi identificada a presença de um tipo de ruído heteroscedástico que apresenta um nível de ruído proporcional à intensidade do sinal analítico. Este ruído heteroscedástico deve ser transformado para um nível de ruído constante para que possa ser identificado pelos modelos de calibração multivariada e conseqüentemente eliminado na forma de resíduo a partir da decomposição da matriz de dados. Para esta transformação, foi utilizada a função logarítmica aplicada em toda a matriz de dados. Entretanto, este artifício matemático implica a perda de linearidade do sinal analítico com a concentração e deve ser utilizado com muito cuidado dependendo do tipo de dado e modelo de calibração multivariada aplicado.

Capítulo 5

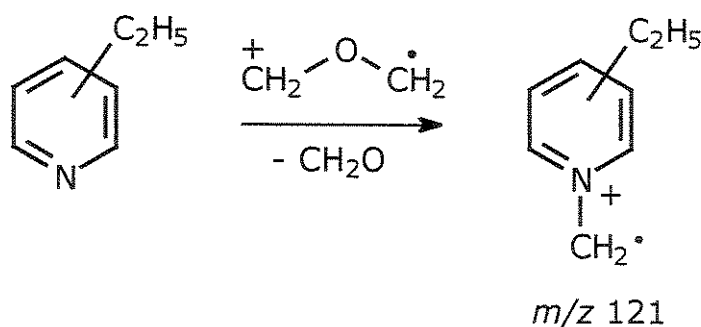
Calibração multivariada aplicada a dados de primeira ordem

1. Introdução

1.2. Reação íon-molécula e dissociação através da técnica CID

A reação íon-molécula é muito utilizada na investigação de estrutura de íons ou moléculas neutras e no desenvolvimento de métodos para distinguir isômeros na fase gasosa [81]. Este estudo é possível através da análise seqüencial em espectrometria de massas, que permite estudar a reatividade e estrutura de reagentes e produtos através da seleção de íons em múltiplos estágios [82]. A química de cátions distônicos em fase gasosa tem sido estudada extensivamente e suas reações utilizadas na análise de estrutura e distinção de isômeros [83].

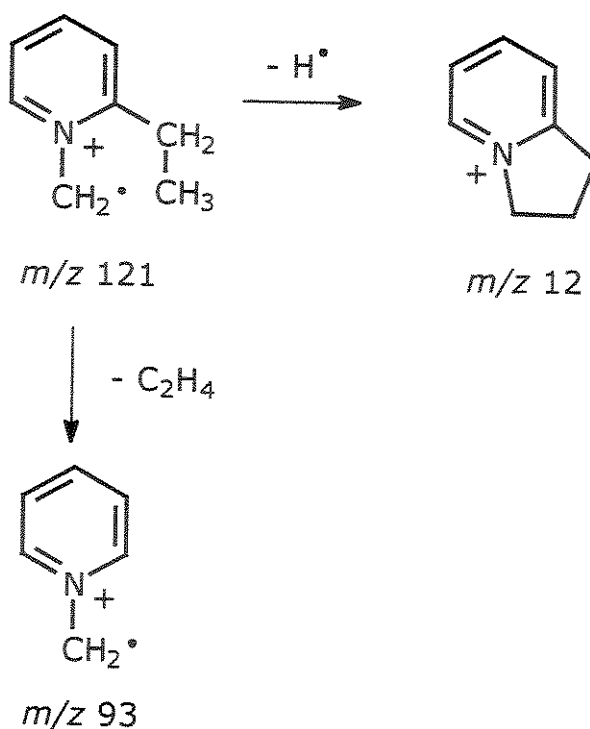
A partir de trabalhos anteriores, pode-se verificar que íons distônicos como $^+\text{CH}_2\text{--O--CH}_2^{\bullet}$ reagem facilmente com os isômeros de etilpiridina através da transferência de CH_2^{+} (Esquema1) [84]. A dissociação induzida por colisão (CID) de cada íon isomérico distônico *N*-metileno-etilpiridínio (m/z 121) formado, mostra um comportamento de dissociação característico.



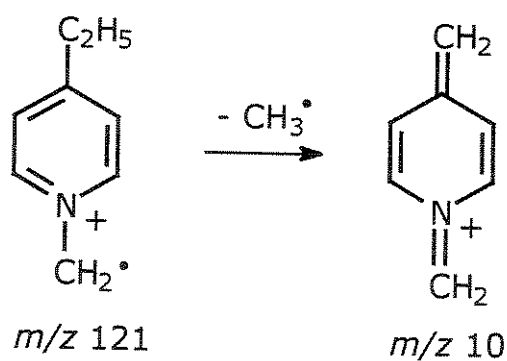
Esquema 1

O íon de m/z 121 formado a partir do isômero 2-etilpiridina produz dois fragmentos característicos: um de m/z 120 pela perda de um átomo de hidrogênio, um interessante efeito orto [84] e outro m/z 93 por perda de uma molécula de eteno (Esquema 2).

A dissociação do íon de m/z 121 a partir do isômero 4-etilpiridina, produz quase que exclusivamente um único fragmento de m/z 106 por perda de um radical metila, conhecido como efeito *para* (Esquema 3).



Esquema 2



Esquema 3

O efeito *orto* e *para* mostrado acima não ocorre na substituição *meta* e a perda do átomo de hidrogênio e de um radical metila não são favorecidas para o íon *N*-metileno-etilpiridínio de m/z 121 a partir do isômero 3-etilpiridina. Isômeros com substituição *meta* são mais estáveis e produzem a partir da dissociação por CID a 10 eV íons de m/z 120, 106, 93, 92 e 79 muito pouco intensos.

A reação íon-molécula dos isômeros de etilpiridina com $^+CH_2-O-CH_2^+$ seguida por CID permite uma maior diferenciação entre as misturas isoméricas de etilpiridina. A quantificação destas

misturas, entretanto, só é possível se aplicado algum método quimiométrico de calibração multivariada.

Neste capítulo é investigada a possibilidade de quantificação de misturas de etilpiridinas, a partir de dados de primeira ordem (espectros de massas bidimensionais) utilizando a calibração multivariada.

As etilpiridinas, mais especificamente a 3-etilpiridina é um dos compostos orgânicos voláteis presentes em cigarro e café. Amostras de ar foram analisadas a partir de casas onde residiam pessoas fumantes e foi encontrada uma concentração média de $0,018 \mu\text{g L}^{-1}$ [85].

Dois conjuntos de amostras foram analisados: o primeiro através da simples ionização por elétrons a 70 eV e o segundo através da espectrometria de massas por reação ion-molécula seguida pela dissociação induzida por colisão (CID). O grau de exatidão dos dois processos de análise foi comparado utilizando modelos de calibração multivariada de primeira ordem, PLS. A calibração multivariada foi empregada, pois ao contrário do método univariado, é capaz de quantificar misturas complexas, onde nem todos os analitos apresentam um íon seletivo, como no caso das etilpiridinas.

2. Parte experimental

As misturas dos isômeros de etilpiridina foram preparadas em ampolas, cujo volume total de cada amostra foi igual a $100 \mu\text{L}$, e acopladas ao equipamento de espectrometria de massas. Através do alto vácuo, gerou-se uma atmosfera de amostragem direta, em que o vapor de amostra foi transferido para dentro do espectrômetro de massas. Os experimentos foram realizados no espectrômetro de massas Extrel (Pittsburgh, PA) pentaquadrupolar, descrito no capítulo 1.

O planejamento experimental aproximado dos dois conjuntos de amostras analisadas é apresentado na Figura 28. Os valores de concentrações utilizados foram calculados em termos de frações molares, a partir dos volumes adicionados e as densidades de cada isômero nas misturas, sendo que para o primeiro conjunto o intervalo de fração molar é de 0,1 e do segundo conjunto o intervalo é de 0,125. Deve-se salientar que não é adequado comparar os resultados obtidos a partir de dois conjuntos de amostras que não apresentam as mesmas concentrações, porém, pode-se fazer uma boa avaliação, quando as faixas de concentrações entre os experimentos são muito próximas.

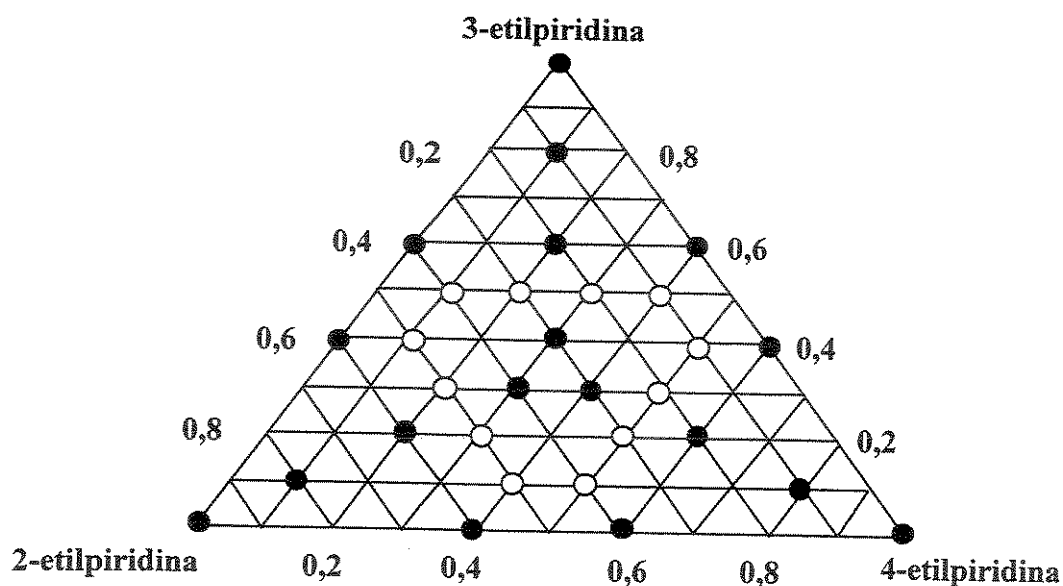


Figure 28. Planejamento experimental com as respectivas frações molares dos isômeros de etilpiridina nas misturas: (●) amostras no modelo de calibração e (○) amostras para testar o modelo.

No primeiro conjunto de amostras, as ampolas com as misturas foram acopladas na fonte de ionização e a simples ionização por EI foi aplicada a 70 eV, obtendo-se espectros de massas bidimensionais. Todos os quadrupolos, exceto o Q₅, foram mantidos sob radiofrequência, permitindo que todos os íons produzidos na fonte de ionização atingssem Q₅ para serem selecionados e detectados. Os íons entre m/z 20 e 120, com incremento de massas de 0,1 m/z , foram registrados, resultando em uma matriz X (30 × 1001).

No segundo experimento, foi utilizada a análise seqüencial, conforme a Figura 29, na qual foi aplicada a reação íon-molécula seguida da dissociação por CID.

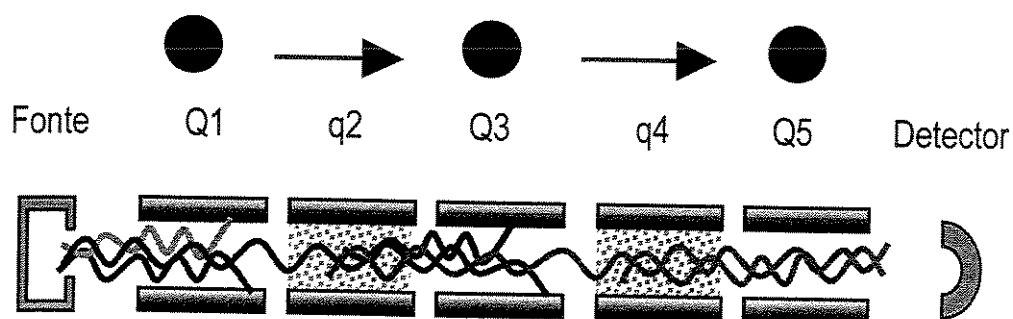


Figura 29. Esquema de um espectrômetro de massas pentaquadrupolar em que a análise seqüencial é aplicada.

O reagente óxido de etileno foi acoplado na fonte de ionização, produzindo o íon distônico ${}^+\text{CH}_2\text{--O--CH}_2^+$ que foi selecionado em Q_1 . As ampolas contendo as misturas dos isômeros de etilpiridina foram acopladas em Q_2 , possibilitando a reação íon-molécula entre as amostras e o íon ${}^+\text{CH}_2\text{--O--CH}_2^+$ com uma energia de colisão aproximadamente de 1 eV. Os produtos da reação em Q_2 , contendo mistura dos isômeros do íon de m/z 121, foram selecionados em Q_3 e dissociados a 10 eV em Q_4 , através da dissociação CID utilizando argônio e em Q_5 os produtos da dissociação foram selecionados e detectados.

Desta forma espectros de massas bidimensionais foram adquiridos registrando apenas o produto final do processo, com intervalo de m/z entre 40 e 130 e incremento de massas de 0,1, resultando em uma matriz X (30×901).

Cada conjunto tem um total de 30 amostras analisadas, das quais 15 foram utilizadas para construir o modelo de calibração através do PLS e as 15 amostras restantes foram utilizadas para testar o modelo de calibração construído. A rotina PLS utilizada é do pacote PLS_toolbox [62] escrito em ambiente MATLAB 5.3 [63].

3. Resultados e Discussão

3.1. Primeiro experimento: EI a 70 eV

Na Figura 30 são mostrados os espectros de massas obtidos por simples ionização por EI a 70 eV para os três isômeros puros das etilpiridinas. A ionização por elétrons é bastante eficiente e produz muitos fragmentos. Antes de aplicar o modelo de calibração multivariada PLS, os espectros de massas foram processados utilizando a transformação logarítmica, como discutido anteriormente no capítulo 4 para a transformação do ruído heteroscedástico em homoscedástico. Em seguida os espectros de massas foram normalizados pelo pico máximo e centrados na média. Assim como os espectros, as concentrações também foram centradas na média.

Um modelo de calibração foi construído separadamente para cada um dos três isômeros através do PLS, utilizando 15 amostras de calibração. O número de componentes utilizadas para a 2-etilpiridina, 3-etilpiridina e 4-etilpiridina foram 5, 5 e 6 respectivamente, encontradas a partir da validação cruzada ("leave-one-out cross validation"). O número de componentes relativamente alto é devido à complexidade do espectro de massas obtido por EI, Figura 30. As concentrações das 15 amostras restantes foram previstas a partir dos modelos de calibrações construídos.

Os erros de previsão para os dois experimentos (EI a 70 eV e reação íon-molécula/CID) foram calculados através dos valores de porcentagem de previsão de erro padrão (percentage standard error of prediction, %SEP) como na equação 18 e os valores obtidos encontram-se na Tabela 9 no final deste capítulo.

$$\%SEP = \frac{100}{\bar{y}} \cdot \sqrt{\frac{\sum_{n=1}^N (y_n - \hat{y}_n)^2}{N}} \quad (18)$$

onde y_n e $\hat{y}_n$ são respectivamente os valores de concentração real e o previsto para N amostras teste e $\bar{y}$ é a concentração média das amostras.

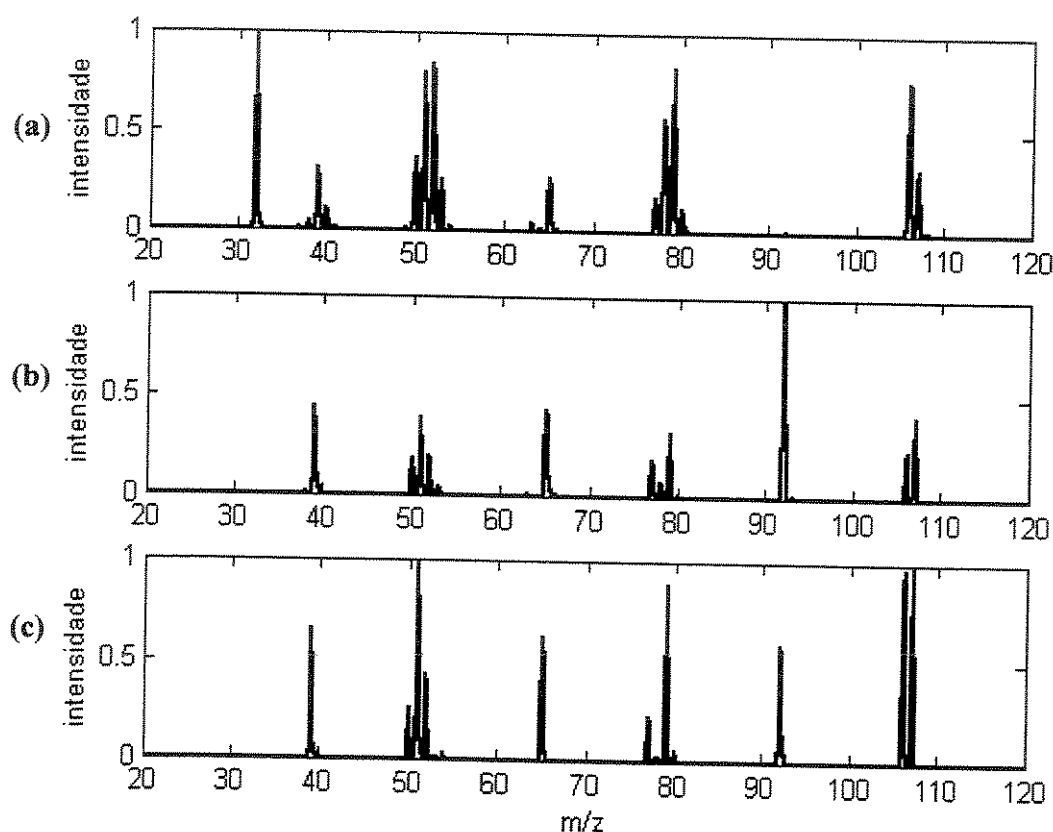


Figura 30. Espectros de massas adquiridos por EI a 70 eV: (a) 2-etilpiridina (b) 3-etilpiridina e (c) 4-etilpiridina.

Na Figura 31 encontram-se os valores de concentrações esperadas contra concentrações previstas para os três isômeros nas misturas e os respectivos valores de coeficientes de correlação.

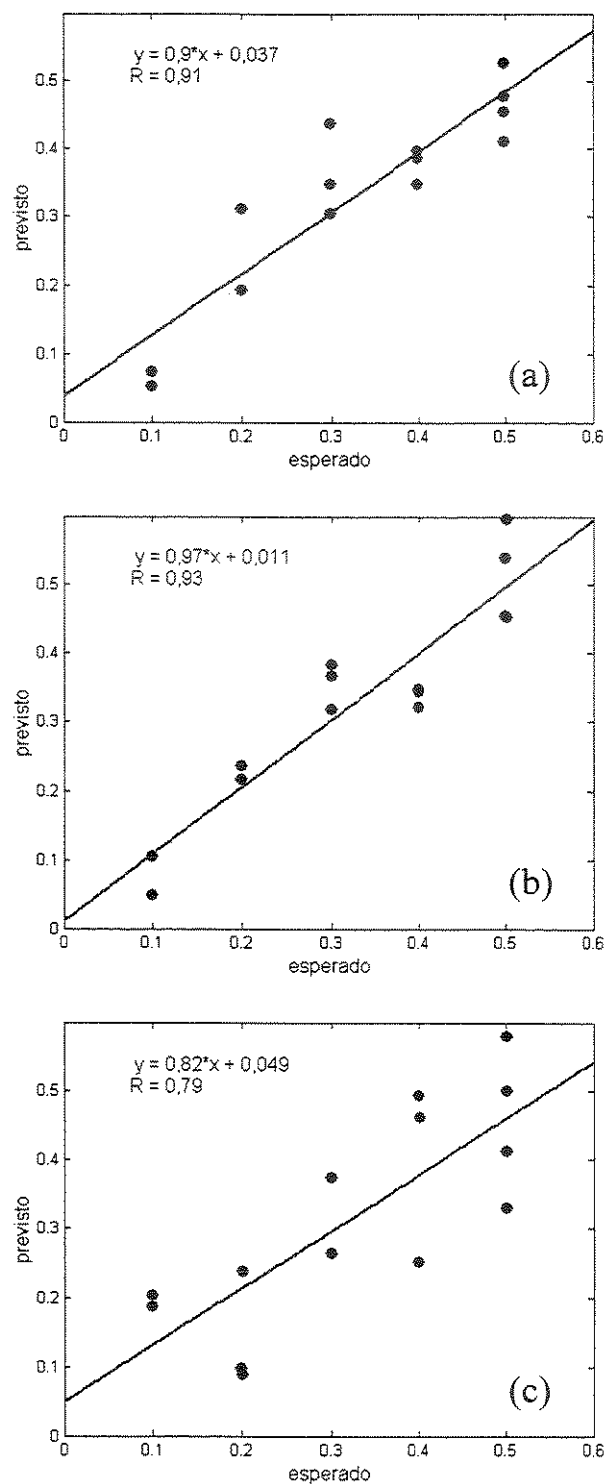


Figura 31. Valor real contra o valor previsto para: (a) 2-etilpiridina, (b) 3-etilpiridina e (c) 4-etilpiridina.

Como pode ser verificado o modelo de calibração não conseguiu um bom ajuste nos dados, provavelmente isto se deve a grande sobreposição de picos nos espectros de massas. Os

resultados na Tabela 9 mostram que os valores de %SEP (18,0; 16,5; 28,9 para os três isômeros respectivamente) não são satisfatórios, pois os erros apresentam valores mais altos, particularmente para o 4-etilpiridina. Desta forma, fica claro que a quantificação a partir de espectros de massas coletados pela simples ionização por EI a 70 eV é bastante complexa pois, mesmo para a calibração multivariada, os espectros de massas são difíceis de serem resolvidos.

3.2. Segundo experimento: Reação íon-molécula seguida de dissociação por CID

Na Figura 32 encontram-se os espectros de massas resultantes da análise seqüencial de triplo estágio dos isômeros de etilpiridina puros.

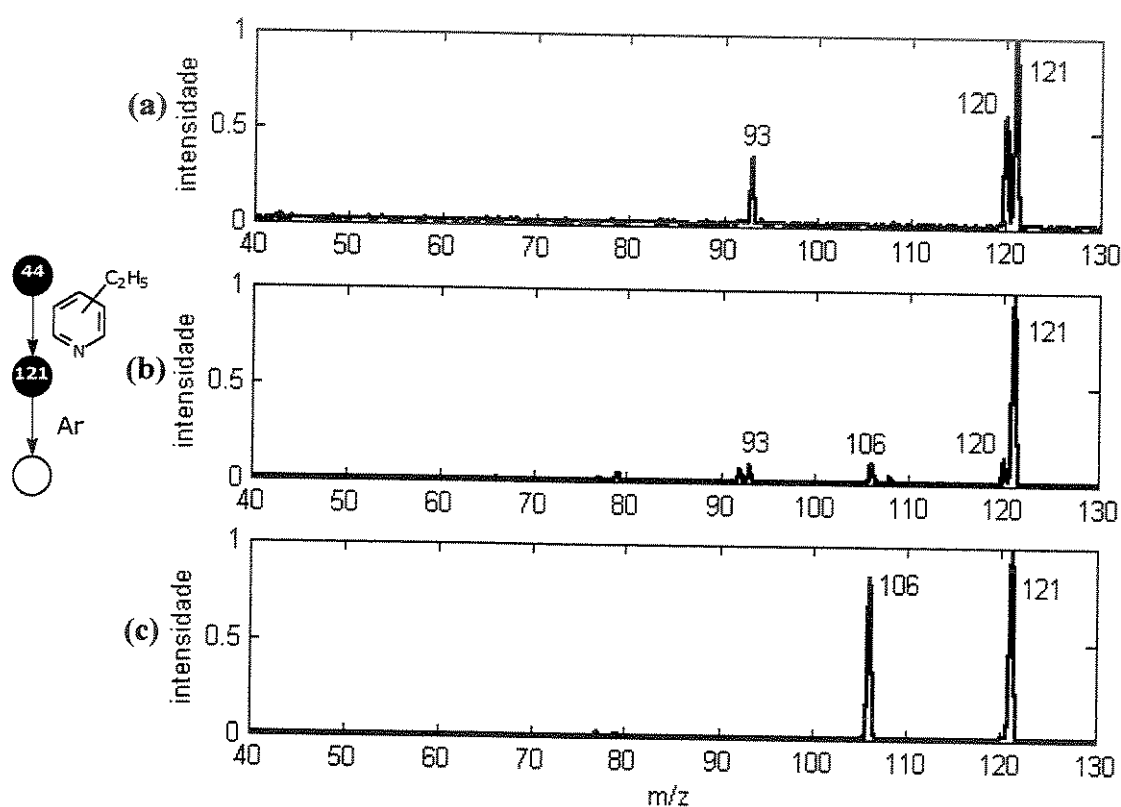


Figura 32. Espectros de massas resultantes da análise seqüencial de triplo-estágio: (a) 2-etilpiridina, (b) 3-etilpiridina e (c) 4-etilpiridina.

Como já mostrado anteriormente, o radical *N*-metileno dos isômeros *orto* e *para* do íon precursor de m/z 121 formado por reação íon-molécula em q_2 , produz um interessante efeito *orto* e *para* (Esquemas 2 e 3) que permite diferenciar as dissociações. Os produtos da dissociação do 2-

etilpiridina (Figura 32a) são especificamente dois fragmentos o m/z 93 e 120, já o 4-etilpiridina (Figura 32c) produz exclusivamente o fragmento de m/z 106.

Pode-se dizer que o íon de m/z 121 decorrente do 3-etilpiridina (Figura 32b) também se dissocia distintivamente, pois não apresenta fragmentos particularmente abundantes. Os íons de m/z 120, 106, 93, 92 e 79 para o 3-etilpiridina são de abundância relativamente baixa, mas que causam interferência quando presentes em misturas com isômeros *orto* e *para*. Esta interferência faz com que a quantificação de misturas contendo os três isômeros torne-se mais complexa e a simples calibração univariada não pode ser aplicada. A utilização da calibração multivariada é uma saída e espectros de massas obtidos através da análise sequencial de misturas contendo os três isômeros, é capaz de realizar uma quantificação mais eficiente. A Figura 33 mostra um espectro de massas típico para uma mistura dos três isômeros. A grande vantagem da calibração multivariada é que esta considera não apenas a abundância de um íon selecionado para cada isômero, mas sim a relação combinada de vários íons presentes nos espectros de massas adquiridos.

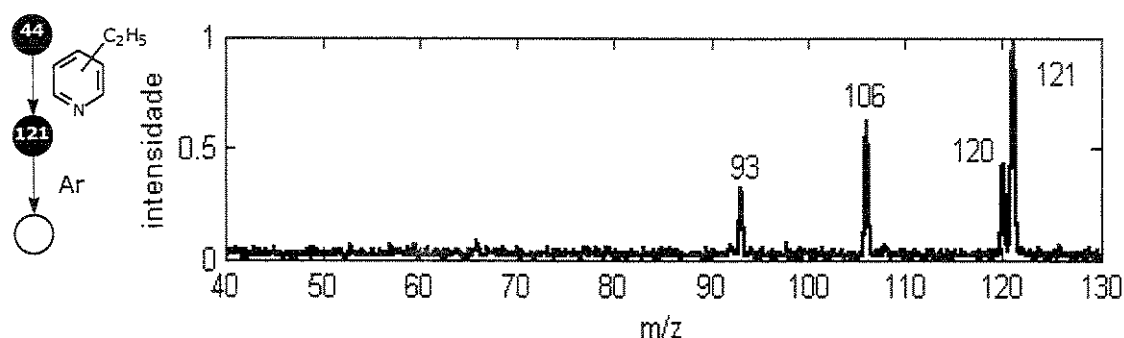


Figura 33. Espectro de massas de uma mistura de isômeros de etilpiridina, obtido a partir da análise sequencial de triplo estágio (frações molares 0,125 : 0,750 : 0,125 de 2-etilpiridina, 3-etilpiridina e 4-etilpiridina respectivamente).

Antes de construir o modelo de calibração através do PLS, os espectros de massas foram pré-processados como no caso anterior, utilizando a transformação logarítmica e na sequência foram normalizados pelo pico máximo e tanto os espectros como as concentrações foram centradas na média. Na construção do modelo de calibração foram utilizadas 4, 3 e 4 componentes para o 2-etilpiridina, 3-etilpiridina e 4-etilpiridina respectivamente, encontrados a partir da validação cruzada. Depois de construídos os três modelos as 15 amostras restantes foram quantificadas e os valores de %SEP para cada um dos isômeros foram calculados (Tabela 9). Na Figura 34 estão os gráficos de valores esperados contra os valores previstos para os três isômeros analisados nas misturas.

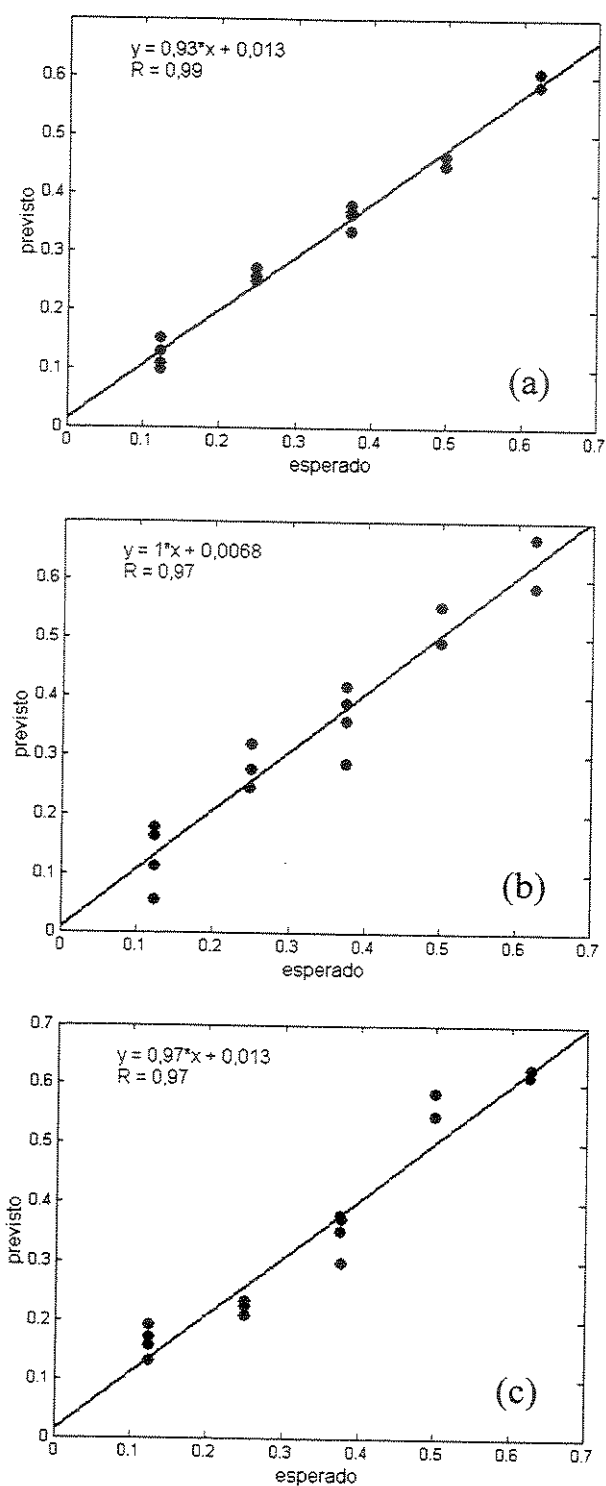


Figura 34. Valor real contra o valor previsto para: (a) 2-etilpiridina, (b) 3-etilpiridina, (c) 4-etilpiridina.

Tabela 9. %SEP a partir de dados de espectrometria de massas adquiridos.

	%SEP		
	2-etilpiridina	3-etilpiridina	4-etilpiridina
El a 70 eV	18,0	16,5	28,9
reação íon-molécula/CID	7,6	13,8	12,5

Os resultados mostram claramente que as quantificações das misturas através da análise seqüencial por reação íon-molécula seguida por CID são muito superiores que aqueles obtidos por simples ionização por EI a 70 eV, obtendo-se erros de previsão na faixa de 8 a 14%. Verifica-se também que um número menor de componentes foi utilizado para descrever o modelo de calibração, confirmando assim que este conjunto de amostra é mais simples que o anterior para ser resolvido pelo PLS.

4. Conclusão

A quantificação de misturas de isômeros de etilpiridina utilizando o modelo de calibração multivariada de primeira ordem PLS é consideravelmente melhor para os dados obtidos pela análise seqüencial reação íon-molécula/CID, do que os obtidos por simples ionização por EI a 70 eV. Os resultados exemplificam a superioridade dos métodos quimiométricos combinados à análise seqüencial através da reação íon-molécula/CID que utilizam as características de reatividade química e comportamento dissociativo para uma rápida e mais exata quantificação de misturas complexas como de isômeros de posição. Esta mesma metodologia pode ser aplicada para diferentes tipos de amostras que apresentem a mesma complexidade com relação a isômeros de posição como é o caso das etilpiridinas.

Capítulo 6

Calibração multivariada aplicada a dados de segunda ordem

1. Introdução

Com o desenvolvimento da instrumentação analítica e a maior ocorrência de dados de segunda ordem, novos métodos de calibração multivariada para tratamento de dados multidimensionais estão sendo desenvolvidos.

Neste capítulo será discutida a aplicação de alguns destes métodos, como N-PLS, U-PLS e MCR descritos no capítulo 2. Alguns outros métodos de calibração de segunda ordem foram também testados, embora estes não mostraram ser adequados por razões que serão discutidas posteriormente.

Dois tipos diferentes de experimentos foram realizados através da espectrometria de massas seqüencial no equipamento pentaquadrupolar. No primeiro experimento foi utilizada a amostragem direta através da simples evaporação das amostras preparadas em ampolas. Amostras contendo 2-etilpiridina, 3-etilpiridina e 4-etilpiridina, foram analisadas através da análise seqüencial, no qual a ionização por elétrons foi seguida de dissociação por CID, gerando dados tridimensionais.

A partir destes espectros, diferentes métodos de pré-processamento foram aplicados para comprovar a importância da função logarítmica como ferramenta para a correção do ruído heteroscedástico. As calibrações multivariadas por U-PLS e N-PLS foram utilizadas para comparar os resultados a partir dos diferentes métodos de pré-processamento aplicados. Estes mesmos dados também foram calibrados pelo método de calibração multivariada MCR e todos estes resultados serão discutidos destacando as vantagens e desvantagens de cada um dos modelos.

No segundo experimento também foi utilizada a amostragem direta, porém aplicando a técnica MIMS para a análise de soluções aquosas contendo compostos orgânicos solubilizados. Misturas contendo benzeno, tolueno e p-xileno (BTX), solubilizados em água na presença de metanol foram analisadas pela ionização por elétrons (EI a 70 eV) gerando dados bidimensionais ou de primeira ordem. Estas soluções também foram analisadas pela análise seqüencial por EI a 70 eV seguida da dissociação por colisão induzida (CID) gerando dados tridimensionais ou de segunda ordem. Os dados de primeira ordem foram quantitativamente analisados, utilizando a calibração multivariada por PLS e os dados de segunda ordem foram quantitativamente analisados, utilizando a calibração multivariada por N-PLS.

2. Parte experimental

2.1. Análise sequencial de misturas de etilpiridina e quantificação por U-PLS, N-PLS e MCR.

Cada amostra foi preparada em ampolas que, por sua vez, foram acopladas ao equipamento de espectrometria de massas. Através do alto vácuo, gerou-se uma atmosfera de amostragem direta que introduziu o vapor de amostra para dentro do espectrômetro de massas pentaquadrupolar Extrel (Pittsburgh, PA) descrito anteriormente no capítulo 1. A pressão no equipamento durante a aquisição dos dados foi mantida aproximadamente em 9×10^{-5} Torr.

Foi preparado um conjunto de 24 amostras contendo misturas dos isômeros 2-etilpiridina, 3-etilpiridina e 4-etilpiridina em diferentes frações molares com um volume total de 100 μL . Na Figura 35 é mostrado o planejamento experimental das misturas e na Tabela 10 encontram-se os valores de frações molares calculados a partir dos volumes adicionados e as densidades de cada isômero nas misturas.

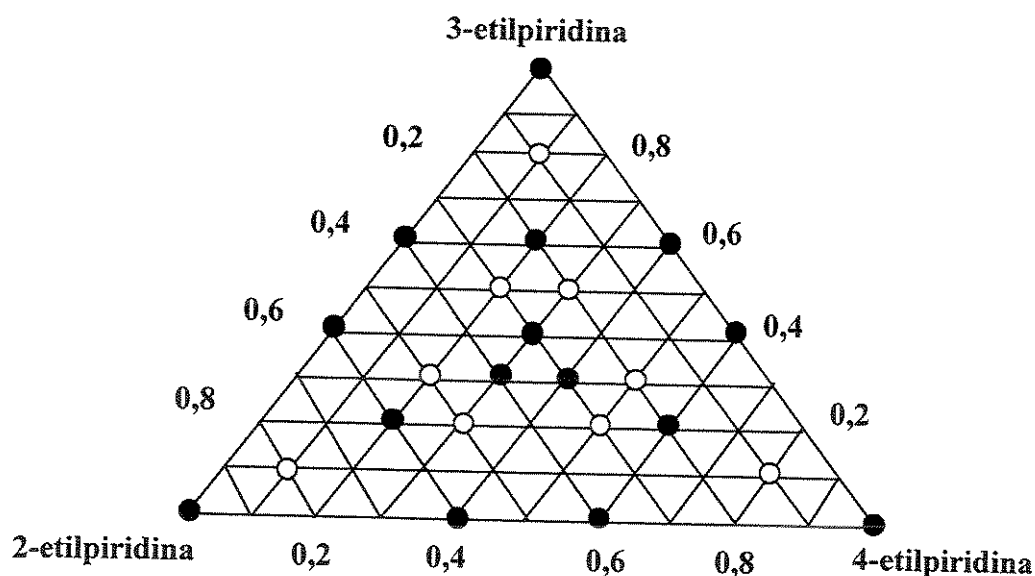


Figura 35. Planejamento experimental com as respectivas frações molares dos isômeros de etilpiridina nas misturas: (●) amostras no modelo de calibração e (○) amostras para testar o modelo.

Tabela 10. Frações molares das 24 amostras analisadas por MS/MS

amostras modelo	2-etilpiridina	3-etilpiridina	4-etilpiridina
1	1.000	0.000	0.000
2	0.599	0.401	0.000
3	0.399	0.601	0.000
4	0.000	1.000	0.000
5	0.599	0.200	0.201
6	0.199	0.600	0.201
7	0.399	0.300	0.301
8	0.599	0.000	0.401
9	0.299	0.400	0.301
10	0.299	0.300	0.401
11	0.000	0.600	0.400
12	0.362	0.000	0.638
13	0.199	0.200	0.601
14	0.000	0.400	0.600
15	0.000	0.000	1.000
amostras teste			
1	0.799	0.100	0.100
2	0.100	0.800	0.100
3	0.499	0.300	0.201
4	0.299	0.501	0.200
5	0.499	0.200	0.301
6	0.199	0.500	0.301
7	0.299	0.200	0.501
8	0.199	0.300	0.501
9	0.100	0.100	0.800

Cada espectro de massas MS/MS consiste de 11 íons precursores (m/z 39, 50, 51, 52, 65, 77, 78, 79, 92, 106 e 107) que foram selecionados em Q_3 depois da ionização por EI a 70 eV na fonte de íons. Estes 11 íons precursores foram subsequentemente dissociados por CID, resultando em 11 espectros de massas com os respectivos produtos da dissociação. O quadrupolo Q_5 foi utilizado para varrer estes espectros, em uma faixa entre m/z 10 a 120 com incremento de massas de m/z 0,1, resultando em um tensor de dados X ($24 \times 1101 \times 11$). A representação esquemática do processo é mostrada na Figura 36.

O espectro de massas final para cada amostra é a média de cinco espectros registrados sucessivamente e cada um dos cinco espectros é por sua vez o resultado de múltiplas varreduras (tempo de integração 33,3ms), minimizando assim, o ruído presente em cada medida. Cada espectro de massas tridimensional é fornecido na forma binária, para diminuir o espaço em disco, e foi necessária inicialmente a transformação no formato ASCII para serem tratados em ambiente MATLAB 5.3 [63], através de rotinas do N-way Toolbox [86].

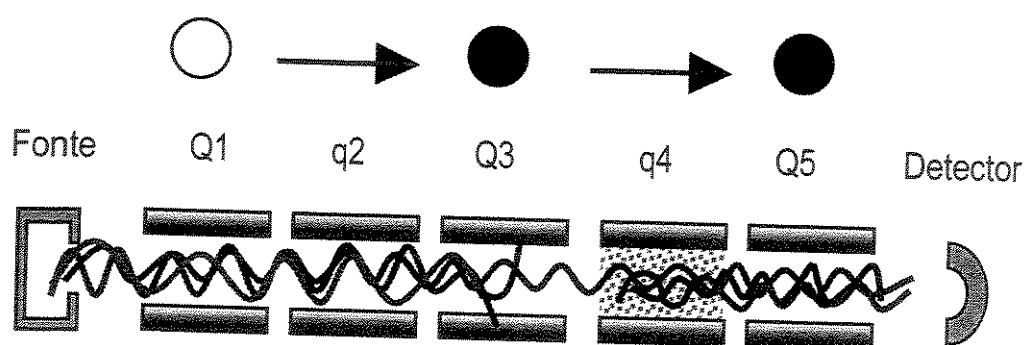


Figura 36. Esquema de um espectrômetro de massas pentaquadrupolar e a análise sequencial aplicada.

2.2. Quantificação de misturas de BTX dissolvidos em água, através do PLS e N-PLS.

Foram analisadas pelas técnicas MIMS e MIMS/MS, 39 amostras contendo misturas de benzeno, tolueno e p-xileno (BTX) solubilizados em água. Cada amostra foi diluída em balões volumétricos de 100 mL, no qual as concentrações variaram entre 300 e 1000 $\mu\text{g L}^{-1}$ aproximadamente, a partir de soluções estoques de 100 mg L^{-1} . As soluções estoques foram diluídas na presença de metanol em água e mantidas em geladeira. As concentrações das 39 amostras estão na Tabela 11, das quais foram selecionadas 18 amostras de calibração e 21 amostras de previsão. Estes valores de concentrações foram convertidos em valores de fração molar, para a construção do modelo de calibração.

Nos dois experimentos MIMS e MIMS/MS foi utilizada a amostragem direta através da sonda de membrana com uma vazão de 2 mL/min. Os dados bidimensionais foram obtidos por simples ionização por EI a 70 eV. Todos os quadrupolos, exceto o Q_5 , foram mantidos sob uma radio frequência permitindo que todos os íons produzidos na fonte de ionização atingissem Q_5 para serem selecionados e detectados. Os íons entre m/z 45 e 110 foram registrados, com incremento de massas de 0,1, resultando em uma matriz X (39×651).

Os dados tridimensionais foram obtidos selecionando 11 íons precursores em Q_3 (m/z 50, 51, 52, 65, 77, 78, 79, 91, 92, 105 e 106) e dissociados em q_4 através da dissociação por CID, cujos produtos da dissociação foram selecionados em Q_5 entre valores de m/z 40 e 110, resultando um tensor X ($39 \times 701 \times 11$).

Tabela 11. Valores de concentração em $\mu\text{g L}^{-1}$ para as amostras de BTX

amostras modelo	benzeno	tolueno	p-xileno
1	300	300	1000
2	600	300	700
3	700	300	600
4	1000	300	300
5	400	400	800
6	800	400	400
7	500	500	600
8	600	500	500
9	300	600	700
10	500	600	500
11	700	600	300
12	300	700	600
13	600	700	300
14	400	800	400
15	300	1000	300
16	1000	0	0
17	0	1000	0
18	0	0	1000
amostras teste			
1	400	300	900
2	500	300	800
3	800	300	500
4	900	300	400
5	300	400	900
6	500	400	700
7	600	400	600
8	700	400	500
9	900	400	300
10	300	500	800
11	400	500	700
12	700	500	400
13	800	500	300
14	400	600	600
15	600	600	400
16	400	700	500
17	500	700	400
18	300	800	500
19	500	800	300
20	300	900	400
21	400	900	300

O espectro de massas final para cada amostra, assim como para as etilpiridinas, é a média de cinco espectros registrados sucessivamente e cada um dos cinco espectros é por sua vez o resultado de múltiplas varreduras (tempo de integração 33,4 ms), minimizando assim, o ruído presente em cada medida. Foram utilizadas rotinas do PLS_toolbox [62] para tratamento de dados

bidimensionais e rotinas do N-way Toolbox [86] para tratamento de dados tridimensionais, todas escritas em ambiente MATLAB 5.3 [63].

3. Resultados e discussão

3.1. Análise seqüencial de misturas de etilpiridina e quantificação por U-PLS, N-PLS e MCR.

Os espectros de massas MS/MS para os três isômeros puros de etilpiridina (2-etilpiridina, 3-etilpiridina e 4-etilpiridina) são apresentados na Figura 37. Como pode ser verificado, eles são bastante similares, apresentando os mesmos fragmentos, que diferenciam apenas pelas intensidades relativas. A única exceção é o 2-etilpiridina que não apresenta o íon de m/z 92.

A análise seqüencial por MS/MS para cada amostra produz uma matriz $\mathbf{X}$ ($J \times K$), onde cada um dos íons precursores K é dissociado gerando um novo espectro de massas com J valores de m/z que descrevem os produtos da segunda fragmentação. Isto significa que uma série de amostras produzirá um arranjo tridimensional de dados, $\underline{\mathbf{X}}$ ($I \times J \times K$), onde I é o número de amostras. Se a concentração de R analitos diferentes na mistura é conhecida, esta informação pode ser coletada em uma matriz bidimensional $\mathbf{Y}$ ($I \times R$). Esta situação é mostrada na Figura 7 do capítulo 2. Desta forma as 24 misturas foram coletadas e ordenadas em um arranjo tridimensional $\underline{\mathbf{X}}$ ($24 \times 1101 \times 11$).

Antes, porém de analisar os diferentes métodos de calibração multivariada e suas vantagens e desvantagens, diferentes pré-processamentos foram utilizados para investigar qual é o mais adequado no tratamento dos dados de espectrometria de massas. No total foram utilizados cinco pré-processamentos diferentes e os resultados foram comparados a partir da calibração multivariada por N-PLS e U-PLS [87].

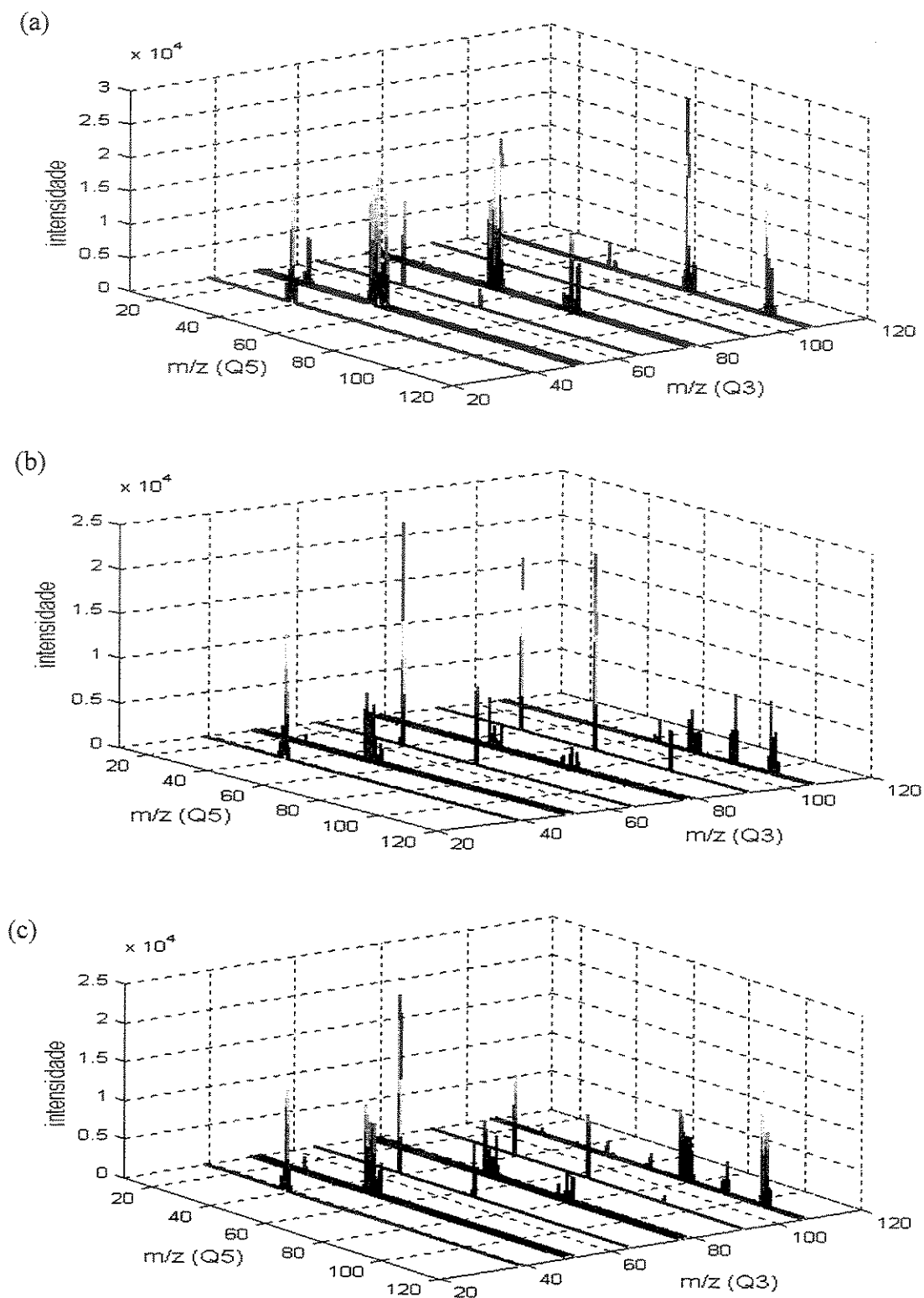


Figura 37. Espectros de massas MS/MS dos isômeros: (a) 2-etilpiridina, (b) 3-etilpiridina, (c) 4-etilpiridina.

Os pré-processamentos utilizados foram:

1. Sem pré-processamento
2. N - Normalização pelo pico máximo:

$$x_{jk}^* = \frac{x_{jk}}{\max(X)} \quad (19)$$

3. L - Transformação logarítmica:

$$x_{jk}^* = \ln(x_{jk} + 1) \quad (20)$$

4. N,L - Normalização pelo pico máximo seguido da transformação logarítmica:

$$x_{jk}^* = \ln\left(\frac{x_{jk}}{\max(X)} + 1\right) \quad (21)$$

5. L,N - Transformação logarítmica seguida pela normalização pelo pico máximo:

$$x_{jk}^* = \frac{\ln(x_{jk} + 1)}{\max(\ln(X + 1))} \quad (22)$$

3.1.1. U-PLS e N-PLS

Para cada um dos pré-processamentos utilizados, antes da calibração multivariada ser aplicada, os espectros de massas e os valores de concentração foram centrados na média. Os dois métodos de calibração multivariada U-PLS e N-PLS foram então aplicados e comparados e nos dois casos, diferentes modelos de calibração foram construídos para cada analito presente nas misturas. O número de componentes utilizados no modelo foram selecionados pelo método da validação cruzada, preferindo o número de componentes que produzisse o menor erro de previsão e ao mesmo tempo, que fosse o mais baixo para construção do modelo.

Na Figura 38 encontram-se os gráficos dos valores reais contra os valores previstos, para cada um dos isômeros, obtidos pelo modelo de calibração U-PLS. Na Figura 39 encontram-se os gráficos dos valores reais contra os valores previstos, para cada um dos isômeros, obtidos pelo modelo de calibração N-PLS.

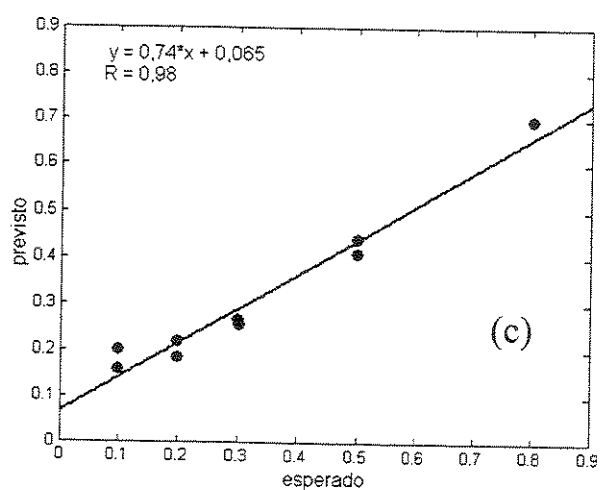
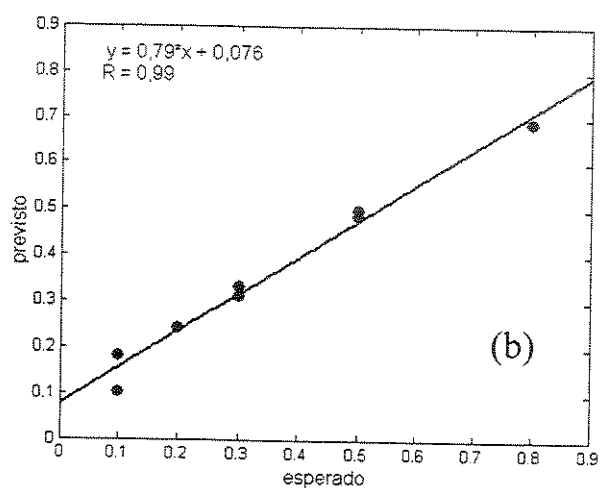
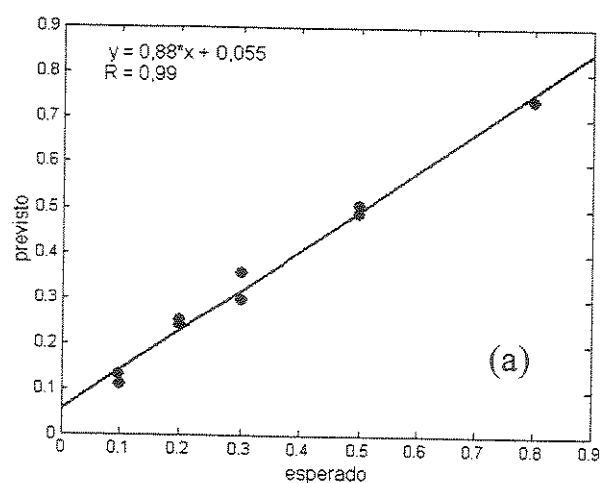


Figura 38. Valor esperado contra valor previsto obtido por U-PLS para: (a) 2-etilpiridina, (b) 3-etilpiridina e (c) 4-etilpiridina

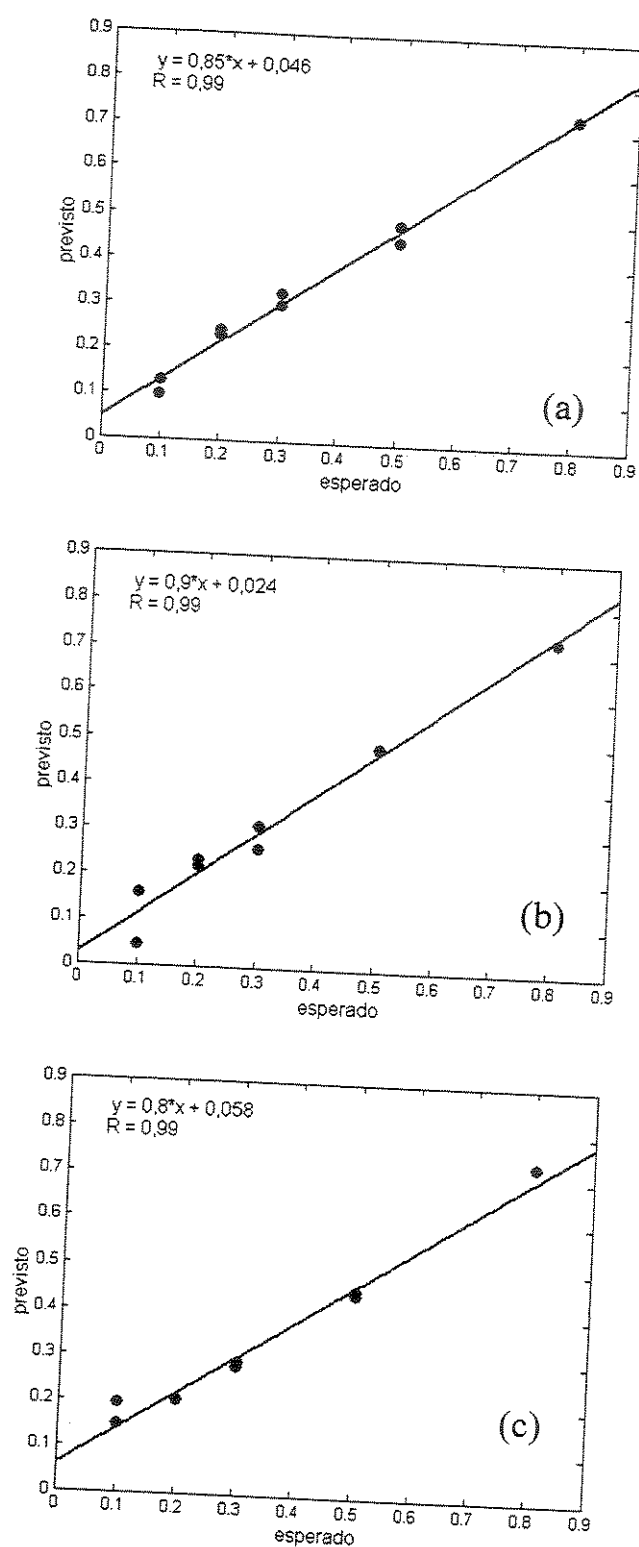


Figura 39. Valor esperado contra o valor previsto obtido por N-PLS para: (a) 2-etilpiridina, (b) 3-etilpiridina e (c) 4-etilpiridina

Os erros de previsão %SEP obtidos por U-PLS e N-PLS para as 9 amostras teste foram calculados conforme a equação 18 e estão na Tabela 12 e 13.

Tabela 12. Resultados da calibração multivariada através do U-PLS a partir de diferentes métodos de pré-processamento.

Pré-processamento:	%SEP		
	2-etilpiridina	3-etilpiridina	4-etilpiridina
1. Sem pré-processamento	11,7	47,7	31,7
2. N	14,7	19,9	32,7
3. L	11,4	15,3	20,8
4. N,L	13,5	15,6	26,4
5. L,N	11,2	15,3	20,3

Tabela 13. Resultados da calibração multivariada através do N-PLS a partir de diferentes métodos de pré-processamento.

Pré-processamento:	%SEP		
	2-etilpiridina	3-etilpiridina	4-etilpiridina
1. Sem pré-processamento	12,6	68,0	55,8
2. N	19,9	34,8	52,8
3. L	10,9	13,4	14,3
4. N,L	18,1	26,5	42,6
5. L,N	11,1	12,3	15,0

Os dois métodos de calibração multivariada mostraram ser eficientes e provaram como é importante utilizar o pré-processamento de dados correto. Pode-se dizer que é fundamental aplicar o pré-processamento através da transformação logarítmica antes da utilização dos espectros de massas para a quantificação das amostras. Deve-se destacar que a melhor opção de pré-processamento, para este conjunto de dados, foi o L,N, ou seja, depois da aplicação da transformação logarítmica foi importante a normalização pelo pico máximo para garantir a correção da linha base.

O N-PLS, entretanto, mostrou ser um método de calibração multivariada mais adequado que o U-PLS, apresentando menores erros de previsão principalmente para o 4-etilpiridina, quando aplicado L,N como etapas de pré-processamento. O N-PLS é mais sensível quando utilizado um método de pré-processamento incorreto, provavelmente porque é um modelo com estrutura mais rígida e menos flexível à não linearidade nos dados.

Verificou-se que a quantificação do 2-etilpiridina é mais fácil que dos outros dois isômeros, pois foram necessários apenas três componentes para a construção do modelo PLS para este analito, já para os demais foram necessários cinco componentes. Embora o espectro do 2-etilpiridina não tenha um pico seletivo ou específico para este composto, como os outros dois isômeros, seu espectro de massas é mais diferenciado, porque o íon precursor de m/z 92 está presente com uma intensidade muito baixa e não aparece no espectro tridimensional, como pode ser verificado na Figura 37(a). É possível verificar apenas um pequeno sinal do seu fragmento m/z 65. O íon de m/z 92 decorrente do isômero 4-etilpiridina também é pouco intenso se comparado com o isômero 3-etilpiridina. Isto porque, a densidade eletrônica na piridina é baixa na posição *orto* e *para* e a perda de um radical metila para a formação do m/z 92 é mais difícil. Já na posição *meta* a densidade eletrônica na piridina é relativamente alta e a perda de um radical metila é bastante favorável formando o m/z 92. A fragmentação do íon de m/z 92, pela quebra na ligação da posição β e a conseqüente eliminação de CHN, resulta no íon de m/z 65 [88]. Mesmo assim, uma baixa quantidade do m/z 92 é formada para as posições *orto* e *para*, pois é possível verificar uma pequena quantidade do íon de m/z 65 sendo formada para estes isômeros, Figuras 37(a) e (c). Além do íon precursor de m/z 92 não aparecer no espectro do 2-etilpiridina outra diferença deste isômero é que a pressão de vapor é mais alta (4,88 mmHg) em relação ao 3-etilpiridina e 4-etilpiridina (2,55 mmHg e 2,22 mmHg respectivamente). Isto significa que este isômero deve sofrer menor influência, devido à variação de pressão ou temperatura, durante o transporte da mistura a partir da ampola para dentro do instrumento.

É também possível verificar que o modelo de calibração para o 2-etilpiridina não é muito dependente do tipo de pré-processamento usado. Para o 3-etilpiridina e 4-etilpiridina, a aplicação da transformação logarítmica para diminuir o ruído heteroscedástico é claramente vantajosa, reduzindo drasticamente o valor de erro quando se aplica a transformação logarítmica antes da normalização (L,N). Para estes dados, a perda da linearidade devido à aplicação da transformação logarítmica pode ser compensada utilizando componentes extras na construção do modelo por U-PLS e N-PLS.

Além da calibração multivariada por N-PLS ter apresentado um melhor resultado, através do pré-processamento L,N em relação ao U-PLS, outro ponto muito importante deste modelo de calibração é a possibilidade de visualizar os valores de pesos, nas duas dimensões dos dados. Na Figura 40, por exemplo, estão os valores de pesos, para cada uma das componentes selecionadas, dos íons precursores que descrevem a primeira dimensão dos dados.

O isômero 2-etilpiridina foi descrito por 3 componentes e os isômeros 3-etilpiridina e 4-etilpiridina foram descritos por 5 componentes. Como discutido anteriormente, íons precursores com pesos molares mais altos, como os íons de m/z 107, 106 e 92, são mais importantes para o desenvolvimento do modelo de calibração por N-PLS, porque seus espectros de massas resultantes da segunda dissociação apresentam um grande número de produtos ou fragmentos e conseqüentemente fornecem mais informação.

É possível também encontrar os fragmentos mais importantes em relação à segunda dimensão dos dados. Na Figura 41 são mostrados, para cada um dos isômeros, os pesos para cada componente selecionada que estão graficadas em diferentes cores. É possível verificar que todos os íons produzidos em toda a extensão do espectro de massas são significativos. Isto podia ser esperado para a calibração multivariada de misturas de isômeros, que apresentam similaridades espectrais muito grandes. Como não existem picos seletivos para cada composto, a calibração multivariada é baseada nas diferenças de intensidades relativas dos picos de cada fragmento produzido. Embora todos os compostos produzam íons de m/z 106 e 107, estes picos na segunda dimensão mostram ser relativamente menos importantes, pois apresentam valores de peso menores e conseqüentemente contribuem pouco na diferenciação dos isômeros. O íon de m/z 92, produzido a partir dos íons precursores de m/z 106 e 107, por sua vez, continua sendo importante depois da segunda dissociação, mantendo um valor de peso alto para os três isômeros. Isto ocorre porque esse íon apresenta maior variação entre os espectros de massas destes isômeros.

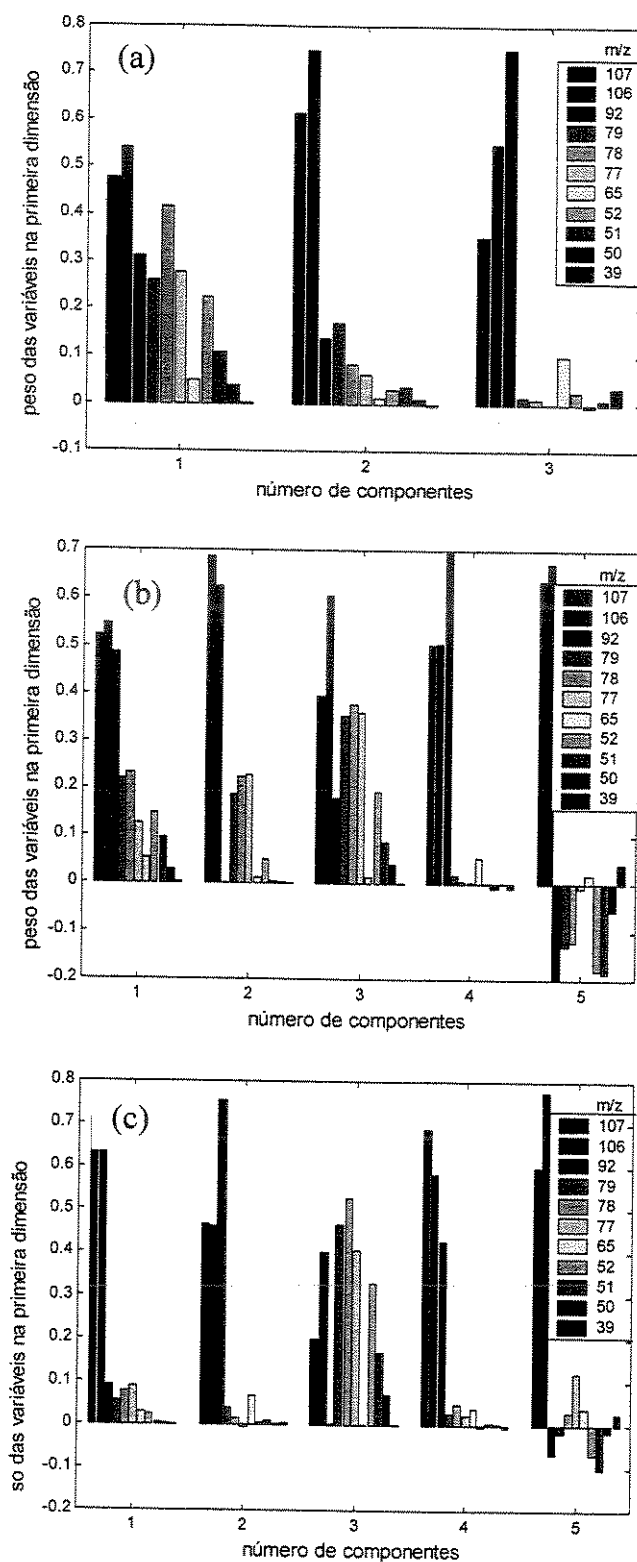


Figura 40. Valores de pesos correspondentes a cada componente para cada um dos íons precursores: (a) 2-etilpiridina (3 componentes), (b) 3-etilpiridina (5 componentes), (c) 4-etilpiridina (5-componentes).

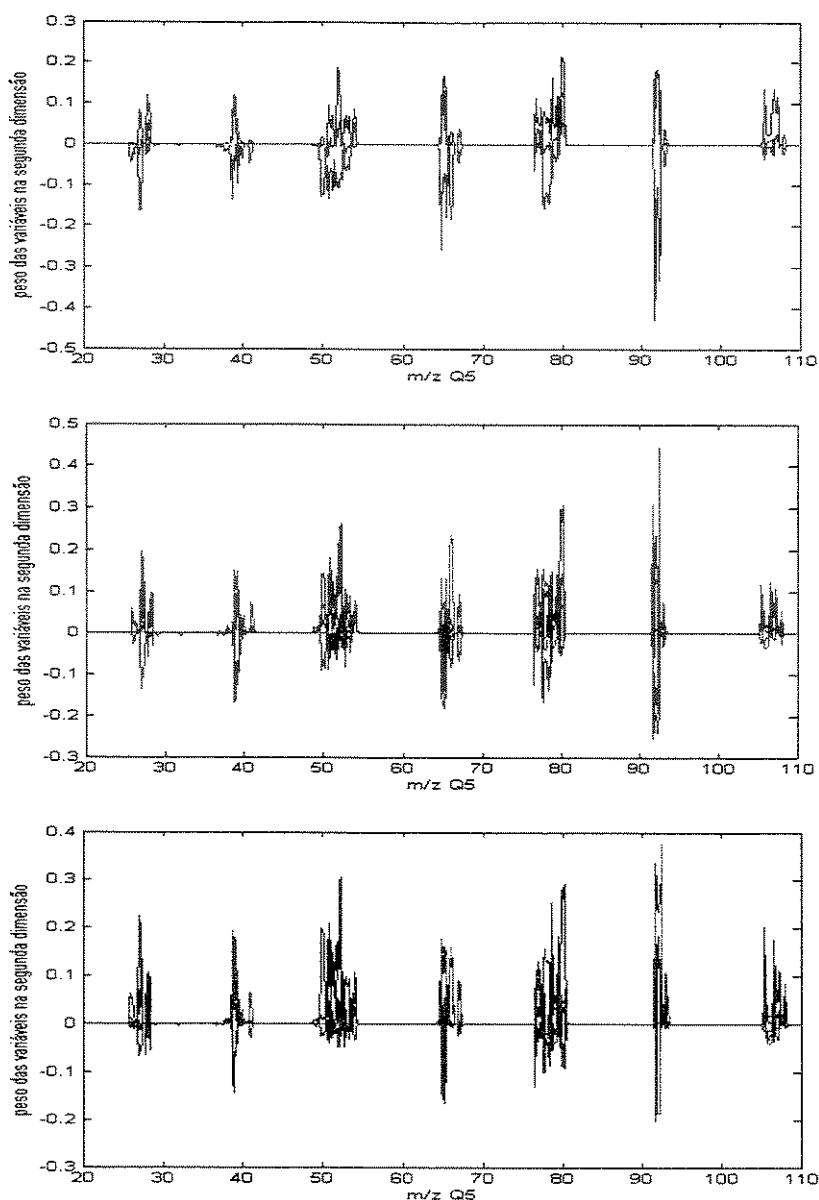


Figura 41. Pesos das variáveis (m/z) para cada componente na segunda dimensão dos dados: (a) 2-etilpiridina (3 componentes), (b) 3-etilpiridina (5 componentes), (c) 4-etilpiridina (5 componentes).

Quando utilizada a transformação logarítmica e o N-PLS, o %SEP encontrado é em torno de 10-15%. Esta faixa de erro é bastante próxima de resultados de outros trabalhos na literatura, que utilizam dados de espectros de massas bidimensionais por amostragem direta [36,37].

O N-PLS impõe uma estrutura trilinear para o conjunto de dados de MS/MS usado no modelo. Entretanto, não é muito claro qual a extensão da estrutura trilinear que está realmente presente nos dados. E, como discutido anteriormente, a matriz produzida a partir de uma amostra

através da análise por MS/MS, certamente não é bilinear como, por exemplo, outros instrumentos que fornecem dados de segunda ordem como GC/MS, no qual cada espectro de massas obtido em diferentes tempos apresenta uma relação linear entre si. Para os dados de MS/MS não há uma relação linear direta entre cada um dos espectros de massas dos produtos (dimensão J), obtidos a partir da dissociação dos precursores (dimensão K). Entretanto, para o N-PLS, esta falta de clareza na estrutura trilinear, não é considerada como um problema, pois o principal objetivo é encontrar um subespaço a partir de $\underline{X}$, que pode ser usado para a regressão.

Uma vantagem do N-PLS quando aplicado a dados de MS/MS, como já mostrado, é a possibilidade de visualizar separadamente as dimensões de pesos para os íons precursores (K) e os íons produzidos a partir da dissociação destes precursores (J), como observado nas Figuras 40 e 41. No U-PLS, as duas dimensões são misturadas quando se realiza o desdobramento dos dados, tornando mais complicada a interpretação do modelo. Além do mais, a aplicação da estrutura multilinear, como no N-PLS, utiliza poucos parâmetros na construção do modelo ($1101+11 = 1112$) em comparação ao U-PLS ($1101 \times 11 = 12111$) para descrever o conjunto de dados. Isto significa que o modelo N-PLS é mais robusto em relação à influência do ruído nos dados.

3.1.2. MCR

O mesmo conjunto de dados utilizado na calibração multivariada por U-PLS e N-PLS, com misturas de 2-etilpiridina, 3-etilpiridina e 4-etilpiridina, foi utilizado na calibração multivariada por MCR. O pré-processamento através da transformação logarítmica e a normalização pelo pico máximo também foram utilizados. O modelo de MCR pode encontrar, a partir do conjunto de dados, as concentrações das amostras e os perfis dos espectros dos analitos puros.

Para aplicar a rotina do MCR, o tensor de dados contendo as 24 misturas das etilpiridinas, foi desdobrado como no modelo U-PLS, mostrado na Figura 8. A partir dos dados desdobrados, o valor de posto químico (ou 'pseudoposto') da matriz de dados foi estimado utilizando o cálculo dos valores singulares (SVD) [34] antes de aplicar o MCR. O SVD também realiza a decomposição da matriz de dados como a PCA, a diferença é que SVD decompõe a matriz entre todos os componentes possíveis, a PCA como é um processo iterativo pode calcular apenas os primeiros componentes que possuem o maior significado químico.

Teoricamente esperava-se um posto químico igual a três, pois cada amostra contém três analitos misturados. Entretanto, como pode ser verificado no gráfico dos valores singulares versus o número de componentes, mostrado na Figura 42, foi encontrado um posto químico igual a dois, uma vez que a partir do terceiro componente não existem valores singulares significativos. Isto confirma a grande correlação ou similaridade entre os espectros de massas de cada um dos compostos puros nas misturas.

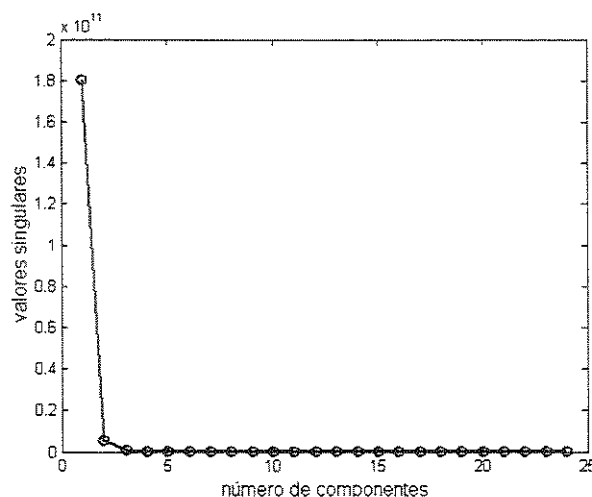


Figura 42. Estimativa do número de componentes por decomposição dos valores singulares (SVD)

Decompondo a matriz de dados pela PCA para os 7 primeiros componentes, mostrado na Tabela 14, verifica-se da mesma forma, que os dois primeiros componentes são suficientes para explicar o conjunto de dados.

Tabela 14. Porcentagem de variância explicada pelas 8 primeiras componentes principais.

componente principal	% variância explicada para cada CP	% variância total explicada
1	92,64	92,64
2	6,29	98,93
3	0,37	99,31
4	0,24	99,55
5	0,20	99,75
6	0,16	99,91
7	0,04	99,95

Baseado na estimativa do posto e sabendo que existe uma proporção linear direta entre as intensidades de pico e a concentração na matriz de dados desdobrada, é possível utilizar a equação 7 do capítulo 2. O MCR foi aplicado para recuperar o perfil de concentração e os espectros dos três isômeros puros. A concentração de novas amostras que não foram incluídas no modelo de calibração pode ser prevista a partir dos espectros dos três isômeros puros recuperados. Este cálculo é baseado nas equações 11, 12 e 13 do capítulo 2.

Os resultados obtidos utilizando todos os espectros de massas decorrente de todos os íons precursores, não se mostraram satisfatórios. Assim, para o MCR, além de aplicar a função logarítmica e a normalização pelo pico máximo, foi necessário também selecionar os espectros dos íons precursores que apresentavam picos mais intensos como m/z 107, 106, 51 e 50 [89]. Os resultados mostraram ser razoáveis na previsão do primeiro isômero (2-etilpiridina), entretanto para os dois últimos isômeros (3-etilpiridina e 4-etilpiridina) não foi possível ajustar um bom modelo de calibração. Entretanto, ao retirar as amostras testes de número 1, 2 e 9 da Tabela 10, que apresentam valores de concentrações mais baixos e mais altos, bons coeficientes de correlação foram obtidos, como pode ser verificado na Figura 43. Deve ser destacado, no entanto, que apesar dos bons resultados, isto não significa que o MCR é melhor que o N-PLS, pois estes valores de erros são referentes a apenas 6 amostras que excluem os erros obtidos para amostras com valores de concentrações extremas, que para os isômeros 3-etilpiridina e 4-etilpirina foram maiores que 30%.

Os valores de %SEP, mostrados na Tabela 15, foram obtidos para os três isômeros e comprovam bons resultados.

Tabela 15. Valores de %SEP resultantes do MCR para 6 amostras teste.

	2-etilpiridina	3-etilpiridina	4-etilpiridina
%SEP	8,4	12,9	13,7

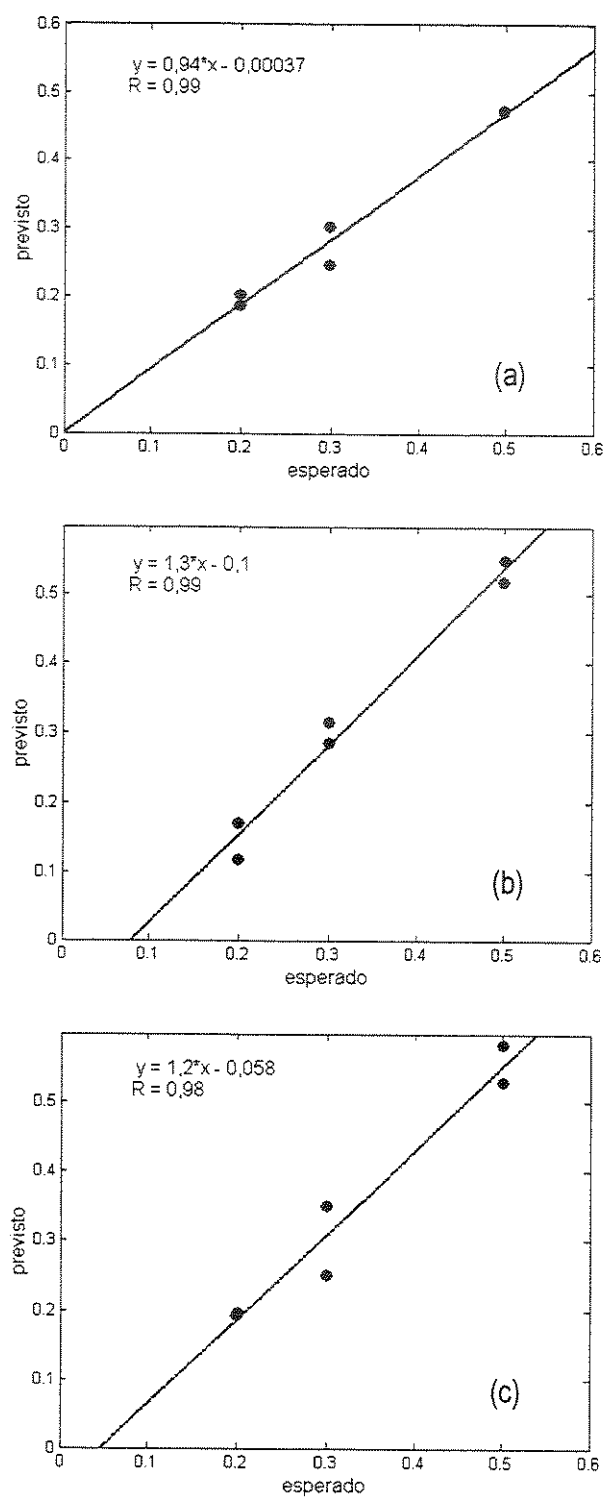


Figura 43. Valor esperado contra o valor previsto obtido por MCR para: (a) 2-etilpiridina, (b) 3-etilpiridina e (c) 4-etilpiridina

Na Figura 44 são apresentados os espectros reais e recuperados para cada isômero puro obtido a partir do modelo construído por MCR e utilizados para a previsão das amostras teste.

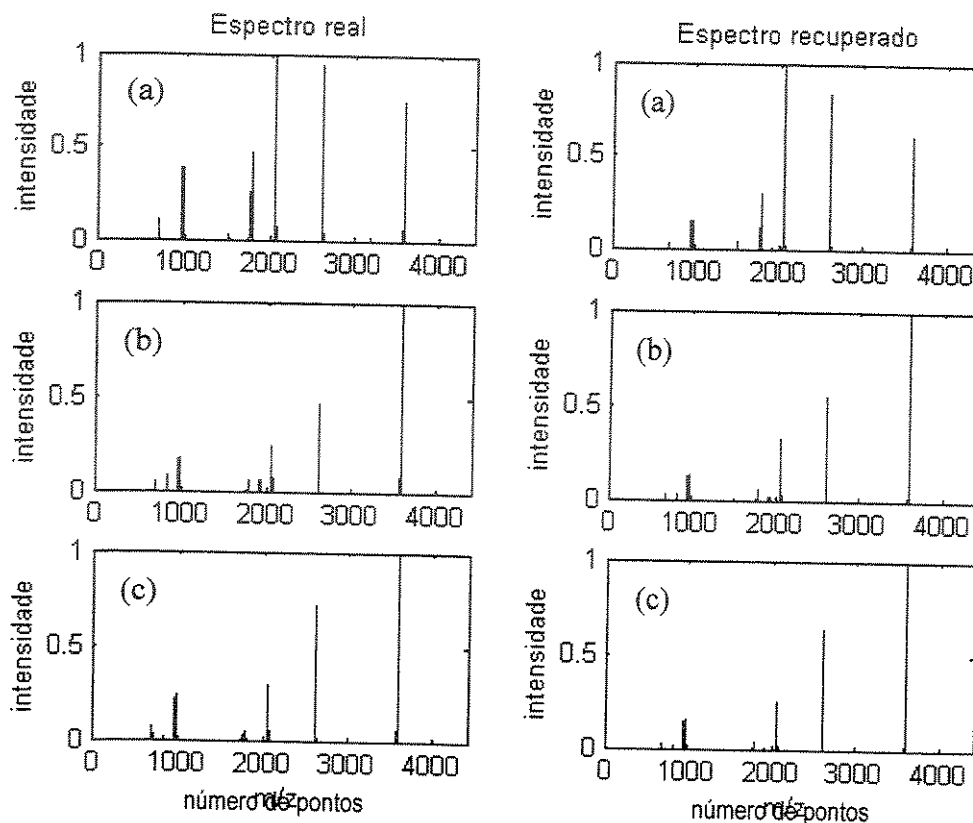


Figura 44. Espectros de massas reais e espectros de massas recuperados através do MCR para: (a) 2-etilpiridina, (b) 3-etilpiridina e (c) 4-etilpiridina

Uma boa concordância é observada entre os perfis dos espectros obtidos experimentalmente e os recuperados pelo modelo, mas que não são suficientes para uma boa previsão de amostras desconhecidas com valores de concentrações muito baixo ou muito alto. Isto se deve, provavelmente, à perda de linearidade do sinal analítico com a concentração, depois que a função logarítmica é aplicada. O MCR é um modelo muito mais sensível a esta perda de linearidade em relação ao N-PLS, como é comprovado com este experimento das misturas de etilpiridina. Desta forma pode-se dizer que os resultados a partir do N-PLS, ainda mostram-se superiores em relação ao MCR.

3.1.3. Métodos de calibração de segunda ordem

Além dos métodos de calibração apresentados U-PLS, N-PLS e MCR, existem os métodos de calibração de segunda ordem, dos quais são mais estritamente confiáveis na presença de uma estrutura trilinear nos dados. Métodos como aniquilação de posto generalizado (generalized rank annihilation, GRAM) [90], bilinearização residual (residual bilinearization, RBL) [46], aniquilação de posto para dados não bilineares (non-bilinear rank annihilation, NBRA) [45] e Tucker com restrições (restricted Tucker) [91], têm atraído a atenção, pois estes métodos oferecem a possibilidade de lidar com interferentes não conhecidos, presentes em novas amostras de previsão. Esta característica destes métodos de calibração de segunda ordem, de prever amostras contendo interferentes que não estavam presentes no conjunto de calibração é conhecida como 'vantagem da calibração de segunda ordem' [45], algo que o método PLS não pode resolver. Quando aplicado o PLS e qualquer outro modelo que seja uma extensão dele (U-PLS e N-PLS), os interferentes presentes nas amostras de previsão devem estar presentes também nas amostras de calibração.

Embora estes métodos tenham sido desenvolvidos para dados bilineares, alguns trabalhos têm mostrado que podem ser estendidos para dados não bilineares como MS/MS. Em particular, Wang *et al.* [45] escreveram um trabalho, no qual mostrou-se um exemplo limitado do uso do NBRA para a quantificação de dados de MS/MS. Entretanto, as tentativas de aplicar ambos os métodos NBRA e RBL nos dados de MS/MS descrito neste trabalho não foram bem sucedidas, produzindo erros de previsão bastante altos. A provável razão é que a aproximação bilinear realizada a partir destes métodos não foi possível, devido à complexidade dos espectros de massas MS/MS, considerados como métodos não bilineares.

3.2. Quantificação de misturas de BTX dissolvidos em água através do PLS e N-PLS.

As 39 amostras contendo BTX solubilizados em água na presença de metanol foram analisadas no espectrômetro de massas pentaquadrupolar, através das técnicas MIMS e MIMS/MS. Os espectros de massas obtidos por MIMS foram arranjados em uma matriz de dados X (39×651), que corresponde a um intervalo de m/z entre 45 e 110. Na Figura 45 tem-se um espectro de massas bidimensional resultante da amostra de número 5 (400 : 400 : 800), utilizada no modelo de calibração. Note que neste caso as concentrações das amostras são baixas e por isso não foi

possível registrar íons com m/z abaixo de 45, pois nesta região a interferência da linha base é bastante alta, devido a presença de CO_2 da água, íons resultantes de resíduos de outros compostos voláteis presentes na fonte de ionização, etc.

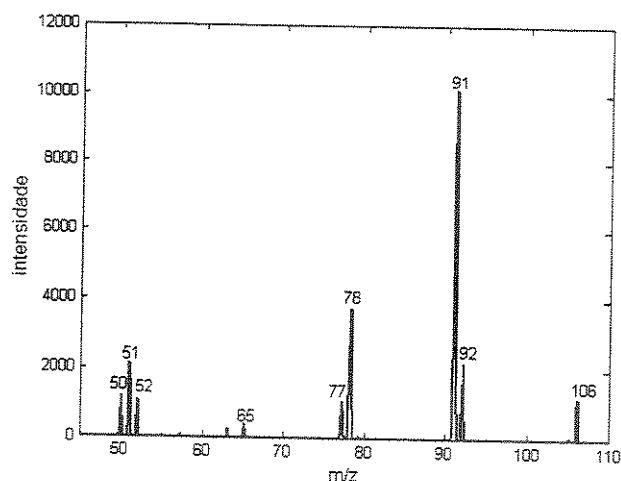


Figura 45. Espectro de massas bidimensional da mistura (400:400:800) de BTX

Os espectros de massas resultantes da análise sequencial MIMS/MS foram arranjados em um tensor de dados $\underline{X}$ ($39 \times 701 \times 11$), onde 39 corresponde ao número de amostras analisadas. Neste caso cada espectro tridimensional é formado por uma dimensão igual a 11, correspondente aos íons precursores selecionados em Q3 (m/z 50, 51, 52, 65, 77, 78, 79, 91, 92, 105 e 106), que depois da dissociação por CID produzem espectros de massas definidos por um total de 701 pontos, equivalente a um intervalo entre m/z 40 a 110. Na Figura 46 tem-se um espectro de massas tridimensional obtido também a partir da amostra 5 (400 : 400 : 800).

Foi aplicada a função logarítmica nos dois casos e em seguida os dados foram normalizados pelo pico máximo e centrados na média, tanto os espectros de massas, como os valores de concentrações.

O PLS foi utilizado na construção do modelo de calibração a partir dos dados de MIMS, sendo que um modelo foi construído separadamente para cada um dos analitos a partir de 18 amostras de calibração e 21 amostras teste. Os modelos de calibração utilizaram 2 componentes para o benzeno e tolueno e 4 componentes para o p-xileno.

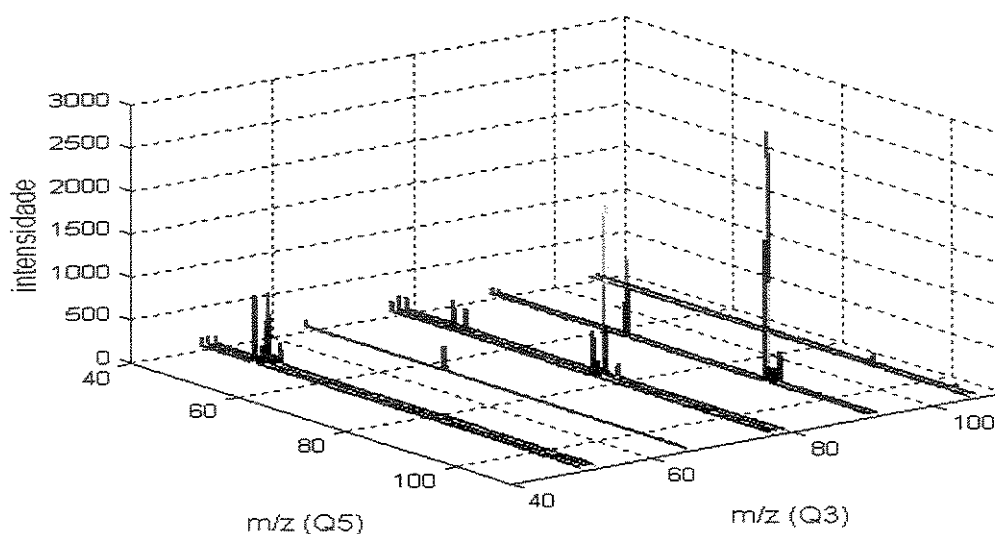


Figura 46. Espectro de massas tridimensional da mistura (400:400:800) de BTX

O N-PLS foi utilizado para a construção do modelo de calibração a partir dos dados de MIMS/MS e da mesma forma que o PLS foi construído separadamente para cada um dos analitos um modelo de calibração diferente, no qual foram utilizados 2 componentes para o benzeno e tolueno e 3 componentes para o p-xileno.

Os valores de concentração esperado contra a concentração prevista encontram-se nas Figuras 47 e 48. Os resultados de porcentagem de erro %SEP encontram-se na Tabela 16. Notem que para a construção do modelo de calibração foram utilizados os valores de concentração em termos de fração molar, minimizando assim o efeito das pressões de vapores diferentes, entre os analitos.

Tabela 16. Valores de %SEP resultantes do PLS e do N-PLS para 21 amostras teste

	%SEP		
	benzeno	tolueno	p-xileno
MIMS	10,6	15,3	20,9
MIMS/MS	12,4	15,3	17,5

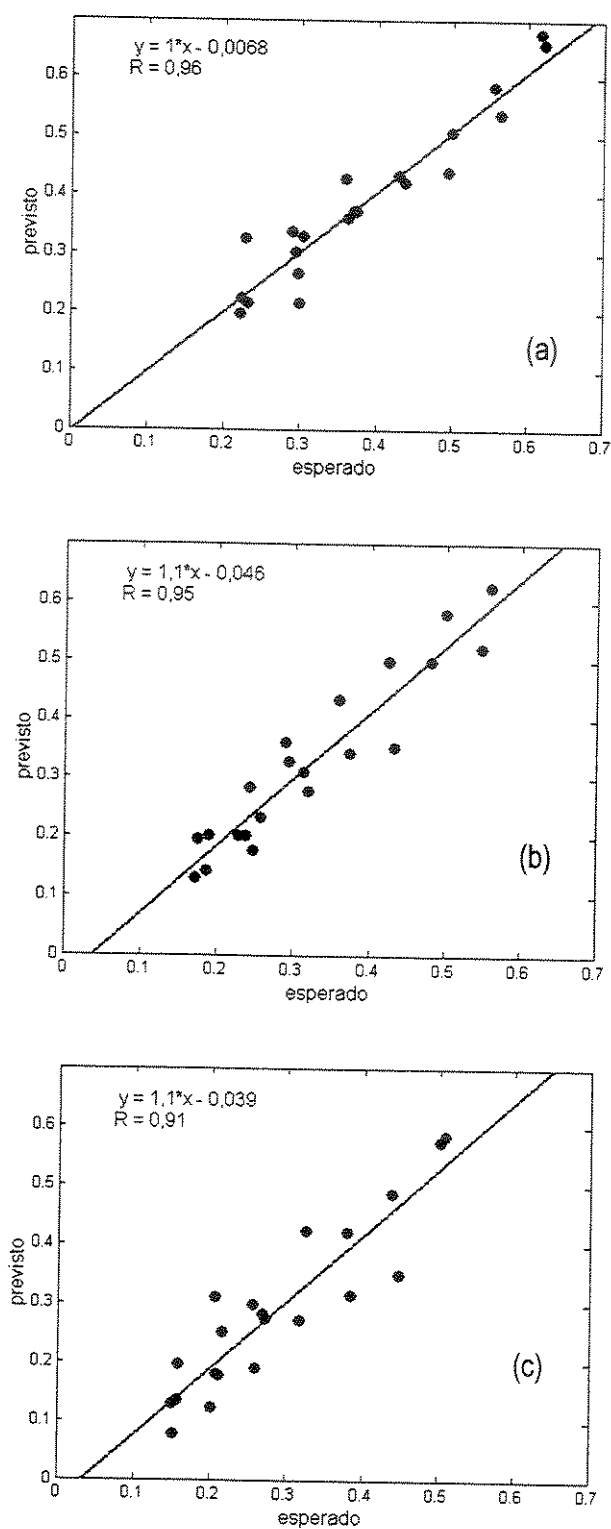


Figura 47. Valor esperado contra o valor previsto obtido por PLS a partir da análise MIMS para: (a) Benzeno, (b) Tolueno e (c) p-Xileno

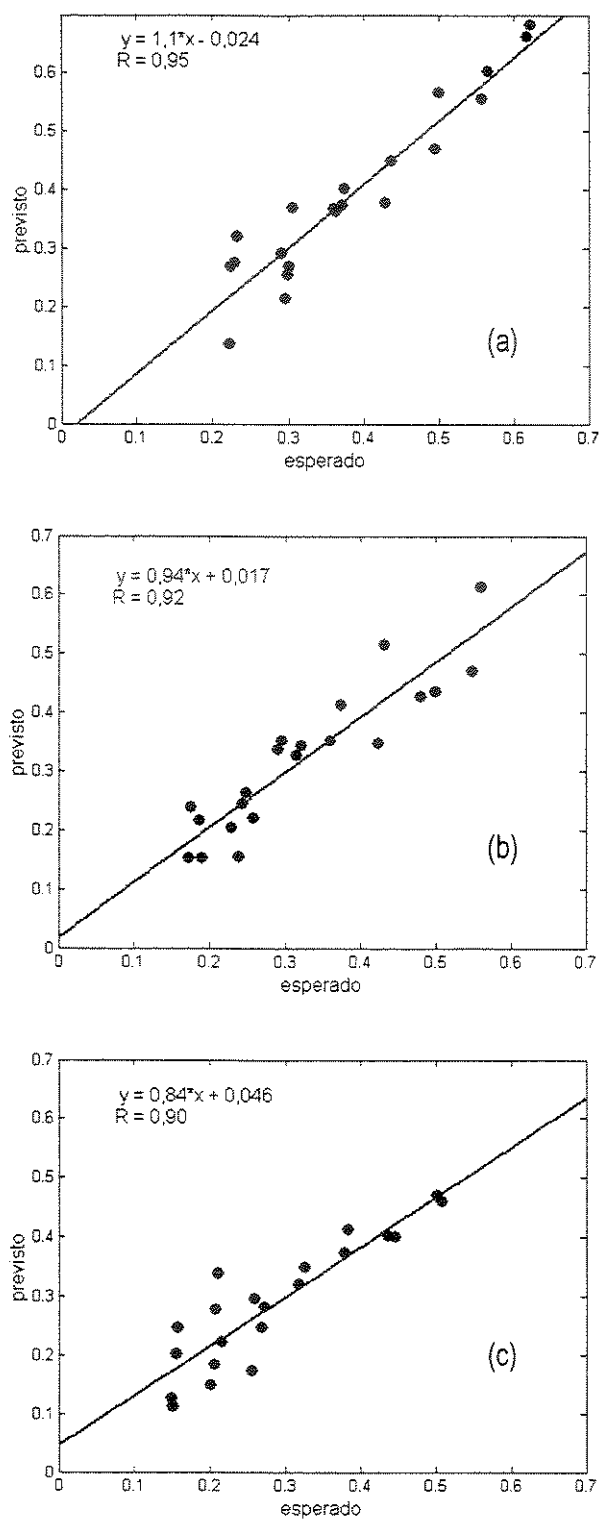


Figura 48. Valor esperado contra o valor previsto obtido por N-PLS a partir da análise MIMS/MS para: (a) Benzeno, (b) Tolueno e (c) p-Xileno

A técnica MIMS/MS (equivalente a MS/MS), como já discutido anteriormente apresenta a vantagem de aumentar a seletividade, facilitando assim a diferenciação de espectros de massas similares. Entretanto, para este caso do BTX os espectros MIMS não são tão similares, sendo possível encontrar íons intensos, quase específicos para cada analito (m/z 78 benzeno, m/z 92 tolueno e m/z 106 p-xilenos). Entretanto, o p-xileno pode interferir na quantificação do tolueno assim como na do benzeno, se a calibração univariada for utilizada, pois no seu espectro existe fragmento de intensidade baixa do m/z 78 e 92. Desta forma, como pode ser verificado na Tabela 16, os resultados para as técnicas MIMS e MIMS/MS não apresentam uma diferença significativa entre os valores de %SEP. Pode-se dizer que os valores de erro de previsão são razoáveis para os dois experimentos (11–21% e 12-18%, MIMS e MIMS/MS respectivamente). Outro ponto que deve ser levado em consideração é que os valores de erro de previsão são bastante próximos dos valores de erro encontrados na literatura para este tipo de experimento [36,37].

Note que o p-xileno apresenta erro de previsão mais alto, isto porque sofre maior interferência dos outros analitos devido à sobreposição de picos como o de m/z 91, 65 e 51. Entretanto, para que estes valores fossem atingidos o modelo PLS utilizou 4 componentes para o p-xileno e o modelo N-PLS utilizou 3 componentes. Isto significa que mesmo os dois métodos apresentando resultados muito parecidos, os dados de MIMS/MS ainda são mais seletivos e o N-PLS pode ser calibrado com menos componentes.

De acordo com experimentos previamente realizados, verificou-se que amostras com concentrações abaixo de $200 \mu\text{g L}^{-1}$ não são possíveis de serem analisadas pela análise seqüencial MS/MS, devido ao comprometimento do sinal analítico em relação ao ruído. Por isso sugere-se que na análise de amostras com concentrações baixas utilize-se outras técnicas de pré-concentração junto a membrana para melhorar a sensibilidade.

4. Conclusão

Um dos principais pontos deste trabalho foi a avaliação dos diferentes métodos de pré-processamentos dos dados de espectrometria de massas, em que foi possível perceber como é importante este processo e que o uso da transformação logarítmica antes da normalização pelo pico máximo fornece melhores resultados.

Todos os métodos de calibração multivariada aplicados em dados de espectrometria de massas seqüencial MS/MS por amostragem direta, que tem por propriedade aumentar a seletividade dos espectros, mostraram ser adequados para a quantificação de isômeros como as etilpiridinas. Entretanto, o N-PLS foi o método de calibração multivariada que forneceu resultados melhores, com erros de previsão da ordem de 11-15%, um resultado que não seria possível sem a utilização da técnica de calibração multivariada para a análise de espectros de massas similares. Além disso, o N-PLS é o único método que fornece informações em relação aos pesos das variáveis nas duas dimensões dos dados, íons precursores (Q3) e produtos (Q5) que podem ser visualizados através dos gráficos de pesos das variáveis. Estas informações podem ser importantes para avaliar o sistema e tirar conclusões sobre o conjunto de dados.

Este trabalho demonstrou que a resolução de curvas multivariadas é um método apropriado para obter-se informações qualitativas e quantitativas a partir de dados complexos de segunda ordem, como os de MS/MS. O MCR recuperou os espectros de massas dos três isômeros puros com uma boa precisão, mas não foi capaz de prever amostras com concentrações muito baixas e concentrações muito altas. Isto porque o MCR, diferentemente do N-PLS, é mais sensível à perda de linearidade devido à aplicação da função logarítmica.

Conclusões gerais e perspectivas futuras

A aplicação da técnica FIA-MIMS para a quantificação de compostos orgânicos voláteis em amostras de água contaminadas com gasolina mostrou-se muito eficiente e rápida. A análise multivariada por PCA, dos espectros de massas obtidos a partir destas amostras, identificou rapidamente os compostos orgânicos que mais solubilizaram em água devido ao efeito de co-solvência exercida pelo etanol. Dentre estes compostos estão o benzeno, tolueno, etilbenzeno, xilenos e dois outros compostos que dão origem a fragmentos de m/z 41 e 55.

Para as amostras de cerveja analisadas pela mesma técnica MIMS e a classificação por PCA, foi possível identificar a partir dos espectros de massas íons de m/z 70, 88, 91 e 104, como sendo produtos da degradação das cervejas. Estes íons foram detectados em concentrações mais elevadas para uma das amostras que apresentou indícios de degradação. Tudo indica que estes compostos são decorrentes de alguns álcoois e ésteres que mesmo em concentrações baixas prejudicam a qualidade do produto. Para que este trabalho seja mais bem finalizado um estudo mais aprofundado em relação à identificação destes compostos deve ser realizado, utilizando técnicas de derivatização ou ainda a análise por CG/MS para que uma melhor seletividade seja alcançada. Estudos têm sido realizados para um melhor controle dos subprodutos obtidos a partir do processo de fabricação de cervejas e conseqüentemente melhorar a qualidade do produto final. Entretanto, sabe-se que mais de 800 compostos orgânicos são produzidos durante o processo de fermentação e muitas enzimas continuam a atuar na cerveja mesmo depois de engarrafada, podendo provocar alterações no produto final.

Verificou-se desta forma que a união da técnica MIMS com a análise multivariada, PCA, tem um grande potencial para o desenvolvimento de métodos de análise exploratória ou de classificação de amostras, em que os espectros de massas são agrupados de acordo com a variação da concentração dos compostos orgânicos voláteis detectados.

Um dos principais pontos deste trabalho foi à avaliação do tipo de ruído presente nos espectros de massas, que é um dos principais passos antes de aplicar a calibração multivariada. Foi identificada a presença de um tipo de ruído heteroscedástico que apresenta um nível de ruído proporcional à intensidade do sinal analítico. Este ruído heteroscedástico deve ser transformado para um nível de ruído constante para que possa ser identificado pelos modelos de calibração multivariada e conseqüentemente eliminado na forma de resíduo a partir da decomposição da matriz

de dados. Para esta transformação foi utilizada a função logarítmica aplicada em toda a matriz de dados. Entretanto, este artifício matemático implica a perda de linearidade do sinal analítico com a concentração.

A quantificação de amostras contendo misturas de 2-etilpiridina, 3-etilpiridina e 4-etilpiridina através da calibração multivariada por PLS, a partir de espectros de massas obtidos pela análise seqüencial por reação íon-molécula/CID, apresentaram resultados superiores aos obtidos pela simples ionização por elétrons a 70 eV. Isto porque ao utilizar a análise seqüencial é possível uma melhor diferenciação entre os espectros de massas, melhorando assim a seletividade em relação à ionização convencional. Estes espectros foram necessariamente pré-processados através da função logarítmica e normalizados pelo pico máximo antes de utilizados para a construção do modelo de calibração multivariado.

Um novo conjunto de misturas de etilpiridina foi analisado através da espectrometria de massas seqüencial por CID e dados tridimensionais foram obtidos. Diferentes métodos de pré-processamento foram aplicados e foi possível observar que, para a obtenção de um bom modelo de calibração, esta etapa é demasiadamente crítica. A utilização da transformação logarítmica antes da normalização pelo pico máximo forneceu os melhores resultados a partir da calibração multivariada para dados multidimensionais (U-PLS, N-PLS e MCR). Todos os métodos de calibração aplicados nos dados de espectrometria de massas tridimensionais mostraram ser adequados para a quantificação de isômeros como as etilpiridinas. Entretanto, foi o N-PLS que forneceu o melhor resultado, com erros de previsão da ordem de 11 a 15%. Além do N-PLS fornecer o melhor resultado é o único método que fornece informações em relação aos pesos das variáveis nas duas dimensões dos dados, íons precursores (Q3) e produtos (Q5) que podem ser visualizados através dos gráficos de pesos das variáveis. Estas informações podem ser importantes para avaliar o sistema e tirar conclusões sobre o conjunto de dados.

Este trabalho demonstrou que a resolução de curvas multivariadas (MCR) é um método apropriado para obter-se informações qualitativas e quantitativas a partir de dados de segunda ordem complexos, como os de MS/MS. O MCR recuperou os espectros de massas dos três isômeros puros com uma boa precisão, mas não foi capaz de prever amostras com concentrações muito baixas ou muito altas. Isto porque o MCR, diferentemente do N-PLS, é mais sensível à perda de linearidade devido à aplicação da função logarítmica.

No último experimento realizado através da técnica MIMS e a calibração multivariada para verificar a viabilidade de utilização destes métodos na determinação de BTX (benzeno, tolueno e p-xileno), bons resultados foram obtidos. Foi comprovada desta forma que a espectrometria de massas pode ser uma ferramenta muito valiosa em análise de rotina e monitoramento em linha, nos quais espectros de massas podem ser adquiridos rapidamente e utilizados para a quantificação através da calibração multivariada. Erros da ordem de 12 a 17% foram obtidos para amostras contendo misturas de BTX com concentrações variando de 300 a 1000 $\mu\text{g L}^{-1}$.

Em todos os experimentos em que métodos de calibração multivariada foram aplicados para a quantificação de compostos orgânicos voláteis, pode ser observado que a previsão de amostras com concentração muito baixa ou muito alta apresentaram problemas. Isto se deve, como já descrito anteriormente, a perda de linearidade devido à aplicação da função logarítmica, necessária para a transformação do ruído heteroscedástico. Uma alternativa para este problema seria utilizar métodos de regressão pesados (weighted regression methods) [93-95]. Entretanto, estes métodos requerem um conhecimento detalhado da distribuição do ruído através do espectro, o qual nem sempre pode ser avaliado. Uma das propostas de trabalhos futuros seria um estudo mais aprofundado em relação a estes métodos de regressão como forma de contornar o problema do ruído heteroscedástico.

Um outro ponto que poderia trazer um grande benefício analítico seria a automatização, ou melhor, a mecanização do sistema de análise e não somente o de aquisição de dados, incluindo todo o processo de amostragem por injeção em fluxo. Isto resultaria em um melhor controle no tempo de amostragem e conseqüentemente a diminuição no erro de análise decorrente da injeção manual. Depois desta etapa acredita-se que um sistema como este seria de grande utilidade em análise de rotina, beneficiando áreas como análises ambientais e de controle de qualidade em indústrias de alimentos e bebidas.

Referências bibliográficas

1. N. Srinivasan, R.C. Johnson, N. Kasthurikrishnan, P. Wong, R. G. Cooks, *Membrane introduction mass spectrometry*, Anal. Chim. Acta, 350, **1997**, 257.
2. T. Kotiaho, F. R. Lauritsen, T. K. Choudhury, R. G. Cooks, G. T. Tsao, *Membrane introduction mass spectrometry*, Anal. Chem. 63, **1991**, 875A.
3. F. R. Lauritsen, T. Kotiaho, *Advances in membrane inlet mass spectrometry (MIMS)*, Rev. Anal. Chem. 15, **1996**, 237.
4. R. C. Johnson, R. G. Cooks, T. M. Allen, M. E. Cisper, P. H. Hemberger, *Membrane introduction mass spectrometry: trends and applications*, Mass Spectrom. Rev. 19, **2000**, 1.
5. L. S. Riter, Z. Takáts, L. Charles, R. G. Cooks, *High surface area membrane introduction mass spectrometry for analysis of volatile and semi-volatile organic compounds in air*, Rapid Commun. In Mass Spectrom. 15, **2001**, 1520.
6. T. Aggerholm, F. R. Lauritsen, *Direct detection of polyaromatic hydrocarbons, estrogenic compounds and pesticides in water using desorption chemical ionization membrane inlet mass spectrometry*, Rapid Commun. In Mass Spectrom. 15, **2001**, 1826.
7. C. Shang, E. R. Blatchley, *Differentiation and quantification of free chlorine and inorganic chloramines in aqueous solution by MIMS*, Environ. Sci. Technol. 33, **1999**, 2218.
8. T. M. Allen, T. M. Falconer, M. E. Cisper, A. J. Borgerding, C. W. Wilkerson Jr., *Real -time analysis of methanol in air and water by membrane introduction mass spectrometry*, Anal. Chem. 73, **2001**, 4830.
9. T. Kotiaho, *On-site environmental and in situ process analysis by mass spectrometry*, J. Mass Spectrom. 31, **1996**, 1.
10. F. R. Lauritsen, S. Gylling, *On-line monitoring of biological reactions at low parts-per-trillion levels by membrane inlet mass spectrometry*, Anal. Chem. 67, **1995**, 1418.
11. R. C. Johnson, N. Srinivasan, R. G. Cooks, D. Schell, *Membrane introduction mass spectrometry in a pilot plant: on-line monitoring of fermentation broths*, Rapid Commun. In Mass Spectrom. 11, **1997**, 363.
12. M. A. LaPack, J. C. Tou, C. G. Enke, *Membrane mass spectrometry for the direct trace analysis of volatile organic compounds in air and water*, Anal. Chem. 62, **1990**, 1265.

13. M. Ojala, R. A. Ketola, T. Mansikka, T. Kotiaho, R. Kostianen, *Determination of mono- and sesquiterpenes in water samples by membrane inlet mass spectrometry and static headspace gas chromatography*, *Talanta*, 49, **1999**, 179.
14. M. E. Cisper, C. G. Gill, L. E. Townsend, P. H. Hemberger, *On-line detection of volatile organic compounds in air at parts-per-trillion levels by membrane introduction mass spectrometry*, *Anal. Chem.* 67, **1995**, 1413.
15. R. Haddad, M. A. Mendes, N. F. Höehr, M. N. Eberlin, *Amino acid quantitation in aqueous matrices via trap and release membrane introduction mass spectrometry: homocysteine in human plasma*, *Analyst*, 126, **2001**, 1212.
16. R. A. Ketola, V. T. Virkki, M. Ojala, V. Komppa, T. Kotiaho, *Comparison of different methods for the determination of volatile organic compounds in water samples*, *Talanta*, 44, **1997**, 373.
17. R. Kostianen, T. Kotiaho, I. Mattila, T. Mansikka, M. Ojala, R. A. Ketola, *Analysis of volatile organic compounds in water and soil samples by purge-and-membrane mass spectrometry*, *Anal. Chem.* 70, **1998**, 3028.
18. M. A. Mendes, R. Sparrapan, M. N. Eberlin, *Headspace membrane introduction mass spectrometry for trace level analysis of VOCs in soil and other solid matrixes*, *Anal. Chem.* 72, **2000**, 2166.
19. M. A. Mendes, R. S. Pimpim, T. Kotiaho, M. N. Eberlin, *A cryotrap membrane introduction mass spectrometry system for analysis of volatile organic compounds in water at the low parts-per-trillion level*, *Anal. Chem.* 68, **1996**, 3502.
20. P. Geladi, and B.R. Kowalski, *Partial least-squares regression: a tutorial*, *Anal. Chim. Acta*, 185, **1986**, 1.
21. H. Martens, T. Næs, *Multivariate calibration*, Wiley, Chichester, **1989**.
22. P. Geladi, *Some recent trends in the calibration literature*, *Chemom. Intell. Lab. Systems*, 60, **2002**, 211.
23. K.L. Busch, G.L. Glish, and S.A. McLuckey, *Mass spectrometry/mass spectrometry techniques and applications of tandem mass spectrometry*, VCH Publishers, Inc. **1989**.
24. <http://www.spectroscopynow.com/Spy/basehtm...>
25. J. Barker, *Mass spectrometry, analytical chemistry by open learning*, 2a. ed., University of Greenwich, Ed. David J. Ando, John Wiley & Sons, **2000**.

26. J. R. Chapman, *Practical organic mass spectrometry, a guide for chemical and biochemical analysis*, 2a. ed., John Wiley & Sons, **1993**.
27. V. F. Juliano, *Controle, aquisição e tratamento de dados para espectrometria de massas seqüencial em aparelho pentaquadrupolar: espectros multidimensionais MS^n ($n=1,2,3$) e monitoramento de íons em função do tempo e potencial de ionização*, Tese de Doutorado, Unicamp, II.5, **1996**.
28. F. W. McLafferty (Ed.), *Tandem mass spectrometry*, Wiley: New York, **1983**.
29. V. F. Juliano, F. C. Gozzo, M. N. Eberlin, C. Kascheres, C. L. Lago, *Fast multidimensional (3D and 4D) MS^2 and MS^3 scans in a high-transmission pentaquadrupole mass spectrometer*, Anal. Chem. 68, **1996**, 1328.
30. J. Ruzicka, E. H. Hansen, *Flow injection analysis*, Wiley, New York, **1988**.
31. J. S. Morgan, W. A. Bryden, R. F. Vertes, S. Bauer, *Detection of chemical agents in water by membrane introduction mass spectrometry*, Johns Hopkins Apl. Tech. Digest. 20, **1999**, 381.
32. M. Soni, S. Bauer, J. W. Amy, P. Wong, R. G. Cooks, *Direct determination of organic compounds in water at parts-per-quadrillion levels by membrane introduction mass spectrometry*, Anal. Chem. 67, **1995**, 1409.
33. K. S. Booksh, B. R. Kowalski, *Theory of analytical chemistry*, Anal. Chem. 66, **1994**, 782A.
34. D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. Jong, P. J. Lewi, J. Smeyers-Verbeke, *Handbook of chemometrics and qualimetrics: part A*, Elsevier, Amsterdam, The Netherlands, **1997**.
35. B. G. M. Vandeginste, D. L. Massart, L. M. C. Buydens, S. Jong, P. J. Lewi, J. Smeyers-Verbeke, *Handbook of chemometrics and qualimetrics: part B*, Elsevier, Amsterdam, The Netherlands, **1998**.
36. S. K. Ohorodnik, R. E. Shaffer, J. H. Callahan, S. L. Rose-Pehrsson, *Mixture analysis using membrane introduction mass spectrometry and multivariate calibration*, Anal. Chem. 69, **1997**, 4721.
37. W. P. Gardner, R. E. Shaffer, J. E. Girard, J. H. Callahan, *Application of quantitative chemometric analysis techniques to direct sampling mass spectrometry*, Anal. Chem. 73, **2001**, 596.
38. S. Wold, P. Geladi, K. Esbensen, J. Öhman, *Multi-way principal component and PLS-analysis*, J. Chemom. 1, **1987**, 41.

39. R. Bro, *Multiway calibration multilinear PLS*, J. Chemom. 10, **1996**, 47.
40. A. K. Smilde, D. A. Doornbos, *Simple validity tools for judging the predictive performance of parafac and three-way PLS*, J. Chemom. 6, **1992**, 11.
41. A. K. Smilde, *Comments on multilinear PLS*, J. Chemom. 11, **1997**, 367.
42. S. De Jong, *Short communication regression coefficients in multilinear PLS*, J. Chemom. 12, **1998**, 77.
43. R. Tauler, I. Marqués, E. Casassas, *Multivariate curve resolution applies to three-way trilinear data: study of a spectrofluorimetric acid-base titration of salicylic acid at three excitation wavelengths*, J. Chemom. 12, **1998**, 55.
44. R. Tauler, *Multivariate curve resolution applied to second order data*, Chemom. and Intel. Lab. Systems, 30, **1995**, 133.
45. Y. Wang, O. S. Borgen, B. R. Kowalski, M. Gu, F. Turecek, *Advances in second-order calibration*, J. Chemom. 7, **1997**, 117.
46. J. Öhman, P. Geladi, S. Wold, *Residual bilinearization. Part 1: theory and algorithms*, J. Chemom. 4, **1990**, 79.
47. M. Esteban, C. Ariño, J. M. Díaz-Cruz, M. S. Díaz-Cruz, R. Tauler, *Multivariate curve resolution with alternating least squares optimisation: a soft-modelling approach to metal complexation studies by voltammetric techniques*, Trends in Anal. Chem. 19, **2000**, 49.
48. R. Tauler, B. Kowalski, *Multivariate curve resolution applied to spectral data from multiple runs of an industrial process*, Anal. Chem. 65, **1993**, 2040.
49. C. M. Kao, C. C. Wang, *Control of BTEX migration by intrinsic bioremediation at a gasoline spill site*, Water Res. 34, **2000**, 3413.
50. W. R. Hartley, A. J. Englande Jr., *Health risk assessment of the migration of unleaded gasoline a model for petroleum products*, Water Sci Technol. 25, **1992**, 65.
51. P. J. Squillace, J. S. Zogorski, W. G. Wilber, C. V. Price, *Preliminary assessment of the occurrence and possible sources of MTBE in groundwater in the united states, 1993-1994*, Environ. Sci. Technol. 30, **1996**, 1721.
52. <http://www.conpet.gov.br>
53. S. R. Cater, N. I. Stefan, J. R. Bolton, A. Safarzadeh-Amiri, *UV/H₂O₂ treatment of methyl tert-butyl ether in contaminated waters*, Environ. Sci. Technol. 34, **2000**, 659.

54. H. X. Corseuil, P. J. J. Alvarez, *Natural bioremediation perspective for BTX-contaminated groundwater in Brazil: effect of ethanol*, Water Sci. Technol. 34, **1996**, 311.
55. H. X. Corseuil, C. S. Hunt, R. C. F. Dos Santos, P. J. J. Alvarez, *The influence of the gasoline oxygenates ethanol on aerobic and anaerobic BTX biodegradation*, Water Res. 32, **1998**, 2065.
56. F. R. Groves, Jr., *Effect of cosolvents on the solubility of hydrocarbons in water*, Environ. Sci. Technol. 22, **1988**, 282.
57. R. M. Alberici, C. G. Zampronio, R. J. Poppi, M. N. Eberlin, *Water solubilization of ethanol and BTEX from gasoline: on-line monitoring by membrane introduction mass spectrometry*, Analyst, 127, **2002**, 230.
58. P. V. Cline, J. J. Delfino, P. S. C. Rao, *Partitioning of aromatic constituents into water from gasoline and other complex solvent mixtures*, Environ. Sci. Technol. 25, **1991**, 914.
59. R. M. Alberici, R. Sparrapan, W. F. Jardim, M. N. Eberlin, *Selective trace level analysis of phenolic compounds in water by flow injection analysis-membrane introduction mass spectrometry*, Environ. Sci. Technol. 35, **2001**, 2084.
60. M. A. Mendes, R. S. Pimpim, T. Kotiaho, J. S. Barone, M. N. Eberlin, *Construção de uma sonda de membrana e sua aplicação na análise de compostos orgânicos voláteis em água através da técnica MIMS e MIMS/MS*, Quim. Nova, 19, **1996**, 480.
61. R. Alberici, R. Sparrapan, M. N. Eberlin, D. Windmler, R. Augusti, *Polyetherimide-silicone: a 10 μ m ultrathin composite membrane for faster and more sensitive membrane introduction mass spectrometry analysis*, Anal. Commun. 36, **1999**, 221.
62. B. M. Wise, N. B. Gallagher, *PLS_Toolbox 2.0 for use with MATLAB*, **1998**.
63. MATLAB v. 5.3, *The Math Works Inc.*, **1999**.
64. <http://www.enter.com.br/noticiascrea/16crea28.htm>
65. R. M. Alberici, C. G. Zampronio, R. J. Poppi, M. N. Eberlin, *Water solubilization of ethanol and BTEX from gasoline: on-line monitoring by membrane introduction mass spectrometry*, Analyst, 127, **2002**, 230.
66. W. A. Hardwick (PhD), *Handbook of brewing*, Marcel Dekker, Inc. New York, **1995**.
67. S. Araki, T. Kimura, C. Shimizu, S. Furusho, M. Takashio, K. J. Shinotsuka, *Estimation of antioxidative activity and its relationship to beer flavor stability*, Am. Soc. Brew. Chem. 57, **1999**, 34.

68. S. Noël, C. Liégeois, G. Lermusieau, E. Bodart, C. Badot, S. Collin, *Release of deuterated nonenal during beer aging from labelled precursors synthesized in the boiling kettle*, J. Agric. Food Chem. 47, **1999**, 4323.
69. F. Chevance, C. Guyot-Declerck, J. Dupont, S. Collin, *Investigation of the β -damascenone level in fresh and aged commercial beers*, J. Agric. Food Chem. 50, **2002**, 3818.
70. S. Masuda, K. Kikuchi, K. Harayama, *Determination of lightstruck character in beer by gas chromatography-mass spectroscopy*, J. Am. Soc. Brew. Chem. 58, **2000**, 152.
71. D. Madigan, A. Perez, M. Clements, *Furanic aldehyde analysis by HPLC as a method to determine heat-induced flavour damage to beer*, J. Am. Soc. Brew. Chem. 56, **1998**, 146.
72. M. Nyborg, H. Outtrup, T. Dreyer, *Investigations of the protective mechanism of sulphite against beer staling and formation of adducts with trans-2-nonenal*, J. Am. Soc. Brew. Chem. 57, **1999**, 24.
73. M. Ojala, T. Kotiaho, J. Surila, M. L. Sihvonen, *Analysis of aldehydes and ketones from beer as o-(2,3,4,5,6-pentafluorobenzyl) hydroxylamine derivatives*, Talanta, 41, **1994**, 1297.
74. S. Sakuma, H. Amano, M. Ohkochi, *Identification of off-flavour compounds in beer*, J. Am. Soc. Brew. Chem. 58, **2000**, 26.
75. C. Zeros, R. H. Albert, *Chemometrics: The use of multivariate methods for the determination and characterization of off-flavors*, in: G. Charalambous (Ed.), *Off-flavors in food and beverages*, Elsevier Science Publishers B. V., Amsterdam, **1992**, 669.
76. P.R.Griffiths (Ed.), *Transform techniques in chemistry*, Heyden, London, **1978**.
77. S.J.Orfanidis, *Introduction to signal processing*, prentice-hall, Engelwood Cliffs, **1996**.
78. O. M. Kvalheim, F. Brakstad, Y. Liang, *Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise*, Anal. Chem. 66, **1994**, 43.
79. F. Brakstad, *The feasibility of latent variables applied to GC-MS data*, Chemom. Intel. Lab. Systems, 29, **1995**, 157.
80. A. M. Woodward, B. K. Alsberg, D. B. Kell, *The effect of heteroscedastic noise on the chemometric modelling of frequency domain data*, Chemom. Intel. Lab. Systems, 40, **1998**, 101.
81. L. A. B. Moraes, F. C. Gozzo, M. N. Eberlin, P. Vainiotalo, *Transacetalization with acylium ions. A structurally diagnostic ion/molecule reaction for cyclic acetals and ketals in the gas phase*, J. Org. Chem. 62, **1997**, 5096.

82. M. N. Eberlin, *Triple-estage pentaquadrupole (QqQqQ) mass spectrometry and ion-molecule reactions*, Mass Spec. Rev. 16, **1997**, 113.
83. M. N. Eberlin, T. K. Majumdar, R. G. Cooks, *Structures and mechanisms of reactions of isomeric C₂H₃S⁺ ions revealed through ion-molecule reactions in conjunction with 2D and 3D mass spectrometry*, J. Am. Chem. Soc. 114, **1992**, 2884.
84. F. C. Gozzo, M. N. Eberlin, *The ionized methylene transfer from the distonic radical cation ⁺CH₂ – O – CH₂[•] to heterocyclic compounds. A pentaquadrupole mass spectrometric study*, J. Am. Soc. Mass Spectrom. 6, **1995**, 554.
85. D. L. Heavner, W. T. Morgan, M. W. Ogden, *Determination of volatile organic compounds and environmental tobacco smoke (ETS) apportionment in 49 homes*, Environ. Intern. 21, **1995**, 3.
86. C. A. Anderson, R. Bro, *The n-way toolbox for MATLAB*, Chemom. Intell. Lab. System, 52, **2000**, 1.
87. C. G. Zampronio, S. P. Gurden, L. A. B. Moraes, M. N. Eberlin, A. K. Smilde, R. J. Poppi, *Direct sampling tandem mass spectrometry (MS/MS) and multiway calibration for isomer quantitation*, Analyst, 127, **2002**, 1054.
88. H. Budzikiewicz, C. Djerassi, D. H. W. Holden-Day, *Mass spectrometry of organic compounds*, Holden-Day, San Francisco, California, **1967**, 566.
89. C. G. Zampronio, L. A. B. Moraes, M. N. Eberlin, R. J. Poppi, *Multivariate curve resolution applied to MS/MS data obtained from isomeric mixtures*, Anal. Chim. Acta, 446, **2001**, 495.
90. E. Sanchez, L. S. Ramos, B. R. Kowalski, *Generalized rank annihilation method 1. application to liquid-chromatography diode-array ultraviolet detection data*, J. Chromatogr. 358, **1987**, 151.
91. A. K. Smilde, R. Tauler, J. M. Henshaw, L. W. Burgess, B. R. Kowalski, *Multicomponent determination of chlorinated hydrocarbons using a reaction-based chemical sensor. 3. medium-rank second-order calibration with restricted tucker models*, Anal. Chem. 66, **1994**, 3345.
92. N. M. Faber, *The price paid for the second-order advantage when using the generalized rank annihilation method (GRAM)*, J. Chemom. 15, **2001**, 743.
93. M. Davidian, P. D. Haaland, *Regression and calibration with no constant error variance*, Chemom. Intell. Lab. System, 9, **1990**, 231.

94. M. E. Zom, R. D. Gibbons, W. C. Sonzogni, *Weighted least squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration*, Anal. Chem. 69, **1997**, 3069.
95. F. Laborda, J. Medrano, J. R. Castillo, *Quality of quantitative and semi quantitative results in inductively coupled plasma mass spectrometry*, J. Anal. At. Spectrom. 16, **2001**, 732.