



UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE QUÍMICA  
PÓS-GRADUAÇÃO EM FÍSICO-QUÍMICA

ANDRÉ MESSIAS KRELL PEDRO

**DESENVOLVIMENTO DO MÉTODO MULTIVARIADO ACELERADO PARA  
DETERMINAÇÃO DO PRAZO DE VALIDADE DE PRODUTOS  
UNINDO QUIMIOMETRIA E CINÉTICA QUÍMICA**

Tese apresentada ao Instituto de Química  
como parte dos requisitos para obtenção  
do título de Doutor em Química

**Orientadora:** Professora Doutora Márcia Miguel Castro Ferreira

**CAMPINAS – SÃO PAULO**

**Março de 2009**

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

P343d Pedro, André Messias Krell.  
Desenvolvimento do método multivariado acelerado para determinação do prazo de validade de produtos unindo quimiometria e cinética química / André Messias Krell Pedro. -- Campinas, SP: [s.n], 2009.

Orientadora: Márcia Miguel Castro Ferreira.

Tese - Universidade Estadual de Campinas, Instituto de Química.

1. Teste de estabilidade. 2. PCA. 3. PARAFAC. 4. Multi-modos. I. Ferreira, Márcia Miguel Castro. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

**Título em inglês:** Development of the multivariate accelerated shelf-life test for determining the shelf-life of products using chemometrics and chemical kinetics

**Palavras-chaves em inglês:** Stability test, PCA, PARAFAC, Multi-way

**Área de concentração:** Físico-Química

**Titulação:** Doutor em Ciências

**Banca examinadora:** Profa. Dra. Márcia Miguel Castro Ferreira (orientadora), Prof. Dr. Roy Edward Bruns (IQ-UNICAMP), Prof. Dr. Ronei Jesus Poppi (IQ-UNICAMP), Prof. Dr. Marco Flôres Ferrão (DFQ-UNISC), Profa. Dra. Paula Fernandes de Aguiar (IQ-UFRJ)

**Data de defesa:** 30/03/2009

“Models are to be used, not believed”  
Henry Theil in *Principles of Econometrics*

À minha família

**André Messias Krell Pedro** finalizou seu doutorado em Físico-Química, sub-área quimiometria, pela Universidade Estadual de Campinas (UNICAMP) em 2009, desenvolvendo trabalhos na área de quimometria e cinética química para determinação do prazo de validade de produtos. Conquistou o título de Mestre em Química pela mesma instituição, na mesma área de concentração, em 2004, por seus trabalhos na área de calibração multivariada para determinação simultânea e não-destrutiva de micro-nutrientes em alimentos. Graduou-se bacharel em química em 2002 também pela UNICAMP, quando foi agraciado com o Prêmio Lavoisier, concedido pelo Conselho Regional de Química IV Região, por desempenho acadêmico excepcional.

Atualmente é cientista chefe na divisão de cosméticos da Unilever Brasil, liderando os times de *data modelling* e analítico, cargo que ocupa desde 2004. Seus trabalhos na área cosmética resultaram no depósito de 4 patentes, sendo também agraciado com várias premiações internas pela concepção, desenvolvimento e lançamento no mercado de diversos produtos. Trabalhou na divisão de alimentos da mesma Companhia, como químico sênior, de 1998 a 2004, onde desenvolveu estratégias para monitoramento de produtos transgênicos, novos métodos analíticos e parcerias com instituições externas à Unilever. De 1996 a 1998 atuou como Coordenar de Qualidade, onde estabeleceu estratégias de monitoramento de materiais e processos, bem como diversas parcerias com fornecedores.

Foi agraciado, em 2008, com o *NIR Young Scientist Award* por seu trabalho sobre a utilização de espectroscopia NIR para determinação do prazo de validade de produtos.

Publicou trabalhos em importantes revistas científicas:

Pedro AMK, Ferreira MMC, The Use of NIR Spectroscopy and Chemometrics for Determining the Shelf-Life of Products, *App. Spectrosc.*, no prelo.

Pedro AMK, Ferreira MMC, Simultaneously calibrating solids, sugars and acidity of tomato products using PLS2 and NIR spectroscopy, *Analytica Chimica Acta*, **2007**, 595, 221-227.

Pedro AMK, Ferreira MMC, Nondestructive Determination of Solids and Carotenoids in Tomato Products by Near-Infrared Spectroscopy and Multivariate Calibration, *Anal. Chem.*, **2005**, 77, 2505-2511.

Pedro AMK, Ferreira MMC, Multivariate accelerated shelf-life testing: a novel approach for determining the shelf-life of foods, *J. Chemometrics*, **2006**, 20, 76-83.

## RESUMO

DESENVOLVIMENTO DO MÉTODO MULTIVARIADO ACELERADO PARA DETERMINAÇÃO DO PRAZO DE VALIDADE DE PRODUTOS UNINDO QUIMIOMETRIA E CINÉTICA QUÍMICA descreve um novo conceito de análise de dados que permite avaliar os mecanismos que governam a degradação de bens de consumo e determinar o período de tempo no qual os bens de consumo mantêm suas características dentro de níveis aceitáveis. O algoritmo une técnicas quimiométricas com a teoria cinética química formal, constituindo um avanço sobre os protocolos para determinação da validade de produto existentes.

Além de ser de fácil interpretação, suas principais vantagens incluem a capacidade de unir informações provenientes de diferentes disciplinas – analíticos, físico-químicos, sensoriais - e a possibilidade de utilização direta de dados oriundos de instrumentação analítica como espectroscópios, cromatógrafos, calorímetros, etc.

O trabalho é dividido em seis capítulos. O Capítulo I traz uma revisão da cinética química e dos métodos convencionais para determinação do prazo de validade de produtos. O algoritmo do Método Multivariado Acelerado, bem como suas premissas, vantagens e desvantagens, são descritos no capítulo II. Os Capítulos III e IV trazem aplicações do Método Multivariado Acelerado. No Capítulo III dados físico-químicos e sensoriais foram utilizados para determinar o prazo de validade de um produto alimentício. No Capítulo IV utilizaram-se dados instrumentais, oriundos de espectroscopia no infravermelho próximo (NIR), como fonte de informação quantitativa que, aliados a avaliações sensoriais, permitiram a determinação da validade de um produto cosmético e forneceram informações relevantes sobre seus modos de degradação.

Uma comparação com o método tri-linear PARAFAC é apresentada no Capítulo V. Apesar de constituir ferramenta importante para avaliação dos modos de degradação de produtos, o método PARAFAC carece de estratégias para determinação efetiva do prazo de bens de consumo.

Finalmente, o Capítulo VI mostra o desenvolvimento de um protocolo para determinação de várias propriedades de produtos concentrados de tomate utilizando regressão PLS2. Demonstrou-se que, apesar da preferência pelo método de calibração PLS1 por parte dos pesquisadores, a regressão PLS2 é vantajosa quando há correlação entre propriedades cujos métodos de referência são precisos e outras, cuja quantificação é imprecisa.

## ABSTRACT

DEVELOPMENT OF THE MULTIVARIATE ACCELERATED SHELF-LIFE TEST (MASLT) FOR DETERMINING THE SHELF-LIFE OF PRODUCTS BY MERGING CHEMOMETRICS AND CHEMICAL KINETICS describes a new concept of data analysis for evaluating the mechanisms that rule out the degradation of consumer goods. The algorithm herein developed took advantage of the soft modeling concepts from chemometrics and of chemical kinetics for determining the period of time within which products would be able to keep their quality characteristics within acceptable levels. Besides being of easy interpretation, the main advantages of the MASLT are the capacity to merge data from different sources – analytical, physical chemical, sensorial – and the ability of handling instrumental data (spectroscopic, chromatographic, calorimetric, etc.) directly for determining the shelf-life of products.

This work and some applications are presented in six chapters. In Chapter I, a review of the chemical kinetics theory, as well as the concepts of conventional shelf-life methods, is presented. The MASLT algorithm, together with its assumptions, characteristics and advantages, is presented in Chapter II.

Chapters III e IV describe applications where the shelf-life was determined successfully by using the MASLT. In Chapter III, physical chemical and sensorial data were merged not only for determining the shelf-life of an industrialized tomato product, but also to study the correlation between the various properties. In Chapter IV, instrumental data from NIR spectroscopy were used as source of quantitative information and, together with sensory analyses, allowed the determination of the shelf-life of a cosmetic product as well as provided valuable information with regards to its degradation mechanisms.

A comparison between the tri-linear method PARAFAC is presented in Chapter V. Despite being an important tool for evaluating the degradation mechanisms of consumer goods, the PARAFAC method lacks strategies for effectively determining the shelf-life of products.

Finally, Chapter VI describes the development of a new protocol for determining several properties of tomato concentrate products using PLS2 regression. It was demonstrated that, despite most researchers having a preference for using PLS1 regression, PLS2 is advantageous when there is strong correlations between properties that can be determined quite accurately by their reference methods with those presenting lower accuracy.

# SUMÁRIO

ÍNDICE DE TABELAS.....	xxi
ÍNDICE DE FIGURAS .....	xxiii
CONSIDERAÇÕES GERAIS .....	1
CAPÍTULO I: CINÉTICA QUÍMICA E DETERMINAÇÃO DO PRAZO DE VALIDADE DE PRODUTOS .....	5
1 – INTRODUÇÃO .....	7
2 - OBJETIVOS .....	9
3 – CINÉTICA QUÍMICA .....	10
3.1 - INTRODUÇÃO.....	10
3.2 – LEIS DE VELOCIDADE.....	12
3.3 – REAÇÕES REVERSÍVEIS E EM EQUILÍBRIO.....	22
3.3 – DEPENDÊNCIA DA LEI DE VELOCIDADE COM A TEMPERATURA .....	24
4 – ESTUDOS DE VALIDADE E ESTABILIDADE DE PRODUTOS.....	29
4.1 - ESTUDOS ACELERADOS DE VALIDADE (ASLT - ACCELERATED VALIDADE TESTS) .....	32
4.1.1 – ESTOCAGEM EM CONDIÇÕES DE TEMPERATURA ELEVADAS .....	33
4.1.2 – ESTUDOS DE VALIDADE COM VARIAÇÃO DA UMIDADE RELATIVA DE ESTOCAGEM .....	36
4.2 – RELAÇÕES ENTRE DIFERENTES CONDIÇÕES DE ESTOCAGEM E CÁLCULO DA VALIDADE.....	40
4.3 - TESTES DE ESTABILIDADE DE PRODUTOS .....	42
4.4 – CONDUÇÃO PRÁTICA DE ESTUDOS DE VALIDADE .....	42
5 – CONCLUSÕES .....	47
6 – BIBLIOGRAFIA.....	48



<b>CAPÍTULO II: MÉTODO MULTIVARIADO ACELERADO PARA DETERMINAÇÃO DO PRAZO DE VALIDADE DE PRODUTOS (MASLT) .....</b>	<b>52</b>
<b>1 – INTRODUÇÃO .....</b>	<b>54</b>
<b>2 - OBJETIVOS .....</b>	<b>56</b>
<b>3 – ANÁLISE DE COMPONENTES PRINCIPAIS .....</b>	<b>56</b>
<b>3.1 - O MODELO MATEMÁTICO DA PCA .....</b>	<b>59</b>
<b>3.2 – ALGORITMOS PARA REALIZAÇÃO DA PCA .....</b>	<b>63</b>
<b>4 – MULTIVARIATE ACCELERATED SHELF-LIFE TEST (MASLT) .....</b>	<b>65</b>
<b>4.1 – PROPRIEDADES E PREMISSAS .....</b>	<b>65</b>
<b>4.2 – ALGORITMO DO MASLT .....</b>	<b>67</b>
<b>5 – CONCLUSÕES .....</b>	<b>70</b>
<b>6 – BIBLIOGRAFIA.....</b>	<b>71</b>
<b>CAPÍTULO III: O MASLT APLICADO A DADOS FÍSICO-QUÍMICOS E SENSORIAIS PRODUTO CONCENTRADO DE TOMATE .....</b>	<b>74</b>
<b>1 – INTRODUÇÃO .....</b>	<b>76</b>
<b>2 - OBJETIVO .....</b>	<b>78</b>
<b>3 – PARTE EXPERIMENTAL .....</b>	<b>78</b>
<b>3.1 – PREPARO E ESTOCAGEM DAS AMOSTRAS.....</b>	<b>78</b>
<b>3.2 – ANÁLISES FÍSICO-QUÍMICAS E SENSORIAIS .....</b>	<b>79</b>
<b>3.3 – ANÁLISE DO PRAZO DE VALIDADE .....</b>	<b>80</b>
<b>4 – RESULTADOS E DISCUSSÃO .....</b>	<b>81</b>
<b>4.1 – RESULTADOS FÍSICO-QUÍMICOS E SENSORIAIS .....</b>	<b>81</b>
<b>4.2 – APLICAÇÃO DO MASLT A PRODUTOS CONCENTRADOS DE TOMATE .....</b>	<b>88</b>
<b>5 – CONCLUSÕES .....</b>	<b>94</b>
<b>6 – REFERÊNCIAS .....</b>	<b>95</b>

<b>CAPÍTULO IV: APLICAÇÃO DO MASLT A DADOS INSTRUMENTAIS: DETERMINAÇÃO DO PRAZO DE VALIDADE DE COSMÉTICOS POR ESPECTROSCOPIA NIR .....</b>	<b>98</b>
1 – INTRODUÇÃO .....	100
2 - OBJETIVO .....	102
3 – PARTE EXPERIMENTAL .....	102
4 – RESULTADOS E DISCUSSÃO .....	103
5 – CONCLUSÕES .....	112
6 – REFERÊNCIAS .....	113
<b>CAPÍTULO V: COMPARAÇÃO ENTRE O MASLT E O MÉTODO TRI-LINEAR PARAFAC .....</b>	<b>116</b>
1 – INTRODUÇÃO .....	118
2 - OBJETIVOS .....	119
3 – PARALLEL FACTOR ANALYSIS (PARAFAC) .....	120
3.1 – REPRESENTAÇÃO MATEMÁTICA DO MÉTODO PARAFAC .....	120
4 – EXPERIMENTAL.....	124
5 – RESULTADOS E DISCUSSÃO .....	126
5.1 – DADOS DE ESPECTROSCOPIA NIR .....	126
5.2 – DADOS FÍSICO-QUÍMICOS E SENSORIAIS .....	133
5.3 – CONSIDERAÇÕES GERAIS SOBRE A UTILIZAÇÃO DO MÉTODO PARAFAC PARA DETERMINAÇÃO DO PRAZO DE VALIDADE .....	139
6 – CONCLUSÕES .....	140
7 – BIBLIOGRAFIA.....	141

**CAPÍTULO VI: DESENVOLVIMENTO DE MÉTODOS RÁPIDOS E NÃO-DESTRUTIVOS DE ANÁLISE DE PRODUTOS DE TOMATE PARA DETERMINAÇÃO DO PRAZO DE VALIDADE**144

<b>1 – INTRODUÇÃO .....</b>	<b>146</b>
<b>2 - OBJETIVO .....</b>	<b>148</b>
<b>3 – PARTE EXPERIMENTAL .....</b>	<b>149</b>
<b>3.1 – MÉTODOS FÍSICO-QUÍMICOS E ESPECTROSCÓPICOS .....</b>	<b>149</b>
<b>3.2 – MÉTODOS QUIMIOMÉTRICOS .....</b>	<b>150</b>
<b>4– RESULTADOS E DISCUSSÃO .....</b>	<b>154</b>
<b>5 – CONCLUSÕES .....</b>	<b>162</b>
<b>6 – REFERÊNCIAS .....</b>	<b>163</b>
<b>CONSIDERAÇÕES FINAIS EPERSPECTIVAS FUTURAS .....</b>	<b>166</b>

## ÍNDICE DE TABELAS

Tabela 1.1 – Principais etapas dos estudos de validade.....	66
Tabela 3.1 – Perfil mínimo desejável para produtos concentrados de tomate .....	104
Tabela 3.2 – Valores iniciais e finais para os parâmetros de concentrado de tomate durante o estudo de validade. ....	105
Tabela 3.3 – Parâmetros cinéticos univariados das propriedades selecionadas .....	107
Tabela 3.4 – Prazos de validade calculados para as principais propriedades do estudo univariado.....	110
Tabela 3.5 – Coeficientes de correlação entre variáveis físico-químicas. ....	113
Tabela 3.6 – Parâmetros cinéticos multivariados para os produtos de tomate .....	115
Tabela 4.1 – Parâmetros cinéticos multivariados para PC1 e PC2 para o conjunto de dados espectrais sem aplicação de pré-processamento.....	132
Tabela 4.2 – Pontuações das avaliações sensoriais para amostras estocadas a 45°C..	133
Tabela 5.1 – <i>Loadings</i> e parâmetros cinéticos do fator 2 para o modo das condições de estocagem (C).....	154
Tabela 5.2 – <i>Loadings</i> do modo das condições de estocagem para o conjunto de dados espectroscópicos após aplicação de pré-processamento por centralização na média ...	155
Tabela 5.3 – <i>Loadings</i> e parâmetros cinéticos do modo das condições de estocagem..	161
Tabela 5.4 – Somatória quadrática dos resíduos ( $SS_E$ ) para os modelos MASLT e PARAFAC .....	162
Tabela 6.1 – Número de fatores e erros de previsão para os modelos obtidos com regressão PLS2.....	181
Tabela 6.2 – Comparação entre os melhores modelos de calibração PLS2 e PLS1 .....	183
Tabela 6.3 – Coeficientes de correlação entre as propriedades em estudo.....	183
Tabela 6.4 - Valores medidos, valores previstos e desvios de previsão para as amostras do conjunto de validação externa.....	184

## ÍNDICE DE FIGURAS

Figura 1.1 – Diagrama de energia pictórico de uma reação química elementar .....	35
Figura 1.2 – Diagrama de energia para uma reação complexa, composta por duas reações elementares. ....	36
Figura 1.3 – Gráfico da variação da concentração da espécie A com o tempo para uma reação de cinética de ordem zero. ....	42
Figura 1.4 - Gráfico da variação da concentração da espécie A com o tempo para uma reação de cinética de primeira ordem. ....	43
Figura 1.5 - Gráfico da variação da concentração da espécie A com o tempo para uma reação de cinética de segunda ordem, de acordo com a Equação 1.24.....	45
Figura 1.6 – Comportamentos típicos das velocidades de reações químicas em função da temperatura.....	49
Figura 1.7 – Diagrama pictórico da distribuição de energia em um sistema reacional. ..	52
Figura 1.8 – Distribuição de energia de Boltzmann para o sistema reacional em duas temperaturas distintas .....	52
Figura 1.9 – Oscilações de temperatura por funções de onda senoidal e quadrada.....	57
Figura 1.10 – Comparação entre as velocidades de reação de produtos estocados em condições constantes e com oscilação de temperatura. ....	58
Figura 1.11 – Exemplo de histórico de temperatura de um produto estocado em condição ambiente não-controlada.....	59
Figura 1.12 – Isotermas típicas de absorção de água. A linha sólida representa um produto menos higroscópico que aquele representado pela linha tracejada.....	61
Figura 1.13 – Diferenças entre os prazos de validade estimados para cinéticas de pseudo-zero e pseudo-primeira ordem .....	69
Figura 2.1 – Representação geométrica da aplicação da PCA para um sistema com duas variáveis .....	80
Figura 2.2 – Representação geométrica dos <i>scores</i> e <i>loadings</i> para PC1.....	81
Figura 2.3 – Representação da PCA como a somatória dos produtos.....	84

Figura 2.4 – Gráfico <i>scree</i> ou de variância, comumente utilizado para selecionar o número de PCs a manter na PCA. ....	85
Figura 2.5 – Representação pictórica do MASLT.....	88
Figura 2.6 - Procedimento de desdobramento do arranjo de dados X.....	92
Figura 2.7 – Gráfico de <i>socres</i> vs. tempo para seleção das PCs relevantes para estudo da degradação do produto. ....	92
Figura 3.1 – Gráficos cinéticos para a) cor e b) vitamina C, que exemplificam cinéticas de pseudo zero e pseudo-primeira ordem, respectivamente .....	108
Figura 3.2 – Estruturas dos carotenóides a) $\beta$ -caroteno e b) licopeno.....	109
Figura 3.3 – Gráficos de PC1 (68%) vs. PC2 (12%) para a) <i>scores</i> e b) <i>loadings</i> . ....	112
Figura 3.4 – Comportamento das primeiras duas PCs em função do tempo para amostras estocadas a 8 °C, 25 °C e 35 °C.....	114
Figura 3.5 – Gráficos cinéticos multivariados para amostras mantidas a 8, 25 e 35 °C..	115
Figura 4.1 – Espectros de amostras estocadas a 45°C: a) sem ajuste das linhas de base; b) ajustando as linhas de base.....	127
Figura 4.2 – Espectros de amostras estocadas a 45°C após aplicação de pré-processamento por derivada-primeira.....	128
Figura 4.3 – Micrografia da emulsão da loção corporal após 70 dias de estocagem a 45°C. As gotículas maiores são oriundas de coalescência .....	129
Figura 4.4 – Gráficos de <i>loadings</i> para PC1 e PC2.....	130
Figura 4.5 – Gráfico cinético dos <i>scores</i> para PC1 e PC2 de amostras estocadas a 8, 25, 35 e 45°C. As linhas tracejadas representam os critérios de corte .....	131
Figura 4.6 – Espectros de amostras estocadas a 45°C após aplicação de pré-processamento por centralização na média .....	135
Figura 5.1 – Representação geométrica do PARAFAC .....	144
Figura 5.2 – Pré-processamento de um arranjo tri-linear, proposto por Smilde <i>et al.</i> A centralização é realizada no primeiro modo, enquanto que o escalamento é realizado no segundo modo.....	145
Figura 5.3 – Representação do PARAFAC como um produto tensorial de matrizes .....	147

Figura 5.4 – Construção de nova amostra e determinação dos <i>loadings</i> temporais correspondentes para determinação do prazo de validade utilizando os modelos PARAFAC .....	148
Figura 5.5 – Representação tridimensional dos espectros NIR em função do tempo para amostras estocadas em a) 25°C e b) 45°C .....	150
Figura 5.6 – Gráficos Corcondia para modelos PARAFAC com a) 2 fatores e b) 3 fatores, para o conjunto de dados de espectros NIR .....	151
Figura 5.7 – Gráficos de <i>loadings</i> do modo espectral (B) para o conjunto de dados de espectros NIR: a) fator 1 e b) fator 2.....	152
Figura 5.8 – Gráfico de <i>loadings</i> do modo temporal (A) para os dados espectroscópicos: a) fator 1 e b) fator 2.....	153
Figura 5.9 – Gráfico cinético dos <i>loadings</i> temporais para o segundo fator do modelo PARAFAC aplicado aos dados de espectroscopia NIR .....	155
Figura 5.10 – Gráficos de diagnóstico de consistência do núcleo para modelos PARAFAC com a) 2 fatores e b) 3 fatores, para o conjunto de dados físico-químicos e sensoriais sem pré-processamento. ....	157
Figura 5.11 – Gráfico de <i>loadings</i> do modo das variáveis (B) para o conjunto de dados físico-químicos e sensoriais para um modelo PARAFAC construído sem pré-processamento de dados .....	158
Figura 5.12 - Gráfico de <i>loadings</i> do modo das variáveis (B) para o conjunto de dados físico-químicos e sensoriais para um modelo PARAFAC construído com pré-processamento por auto-escalamento .....	159
Figura 5.13 – Gráfico de diagnóstico de consistência do núcleo para 2 fatores para o conjunto de dados físico-químicos e sensoriais com pré-processamento por auto-escalamento .....	159
Figura 5.14 – Gráfico de <i>loadings</i> do modo temporal (A) para o conjunto de dados físico-químicos e sensoriais: a) fator 1 e b) fator 2 .....	160
Figura 5.15 – Gráfico cinético dos <i>loadings</i> temporais (A) para o primeiro fator do modelo PARAFAC aplicado aos dados físico-químicos e sensoriais.....	162
Figura 6.1 – Alisamento por Fourier de um sinal genérico com unidades arbitrárias: (a) pico apresentando ruído; (b) uma função Gaussiana de convolução (linha descontínua) é aplicada ao interferograma (linha sólida) para remover os componentes de alta frequência que caracterizam o ruído; (c) sinal após alisamento .....	174

Figura 6.2 – Gráfico de *leverage* vs. resíduos de Student. As linhas horizontais representam os resíduos de Student com 95% de significância enquanto que a linha vertical tracejada representa o valor crítico de *leverage* ..... 176

Figura 6.3 – (a) Espectros originais e (b) espectros alisados. As linhas contínuas e tracejadas representam, respectivamente, amostras contendo teores baixos e elevados de sólidos de tomate ..... 179

Figura 6.4 – Pré-processamento de espectros por a) MSC; b) primeira- e c) segunda-derivada após alisamento por filtro de Fourier com janela de 17 pontos ..... 180

Figura 6.5 - Diferença na eficiência de alisamento pela média (linha sólida) e por filtro de Fourier (linha tracejada) ..... 182



# **CONSIDERAÇÕES**

## **GERAIS**

Em um cenário competitivo agressivo, a inovação é a ferramenta mais eficiente para que as indústrias sofisticem seu portfólio, criando diferenciação, ampliando margens de lucro e aumentando participações no mercado que, em última instância, garantem a sobrevivência das organizações. De fato, indústria desenvolve novos produtos a taxas impressionantes: uma miríade de novos alimentos, cosméticos, produtos de higiene e limpeza e produtos para limpeza doméstica, bem como seus intermediários industriais (matérias-primas) são lançados no mercado anualmente.

Mas o desenvolvimento de novos produtos demanda uma quantidade considerável de recursos e tempo. Além disso, deve-se garantir que novos produtos mantenham suas características de qualidade até o momento e seu uso e, para este fim, testes para determinação do prazo de validade são conduzidos. Existem atualmente diversos protocolos e algoritmos para determinação do prazo de validade de produtos descritos na literatura, como ver-se-á no Capítulo I deste trabalho. Entretanto, estes métodos são univariados por natureza e, deste modo, não são capazes de lidar com a grande quantidade de informação fornecida pela química analítica e pela instrumentação científica, bem como pela análise sensorial.

Este trabalho descreve o desenvolvimento de um algoritmo multivariado inovador para determinação do prazo de validade de produtos que complementa os protocolos atuais. O algoritmo desenvolvido, denominado Método Multivariado para Determinação do Prazo de Validade de Produtos (*Multivariate Accelerated Shelf-Life Test – MASLT*) une técnicas quimiométricas com a teoria cinética química formal e traz um novo conceito de análise de dados que permite avaliar os mecanismos que governam a degradação de bens de consumo utilizando conjuntos de dados multi-dimensionais.

Além de ser de fácil interpretação, suas principais vantagens do MASLT incluem a capacidade de unir informações provenientes de diferentes disciplinas – analíticas, físico-químicas, sensoriais - e a possibilidade de utilização direta de dados oriundos de instrumentação analítica como espectrômetros, cromatógrafos, calorímetros, etc.

# **CAPÍTULO I**

## **CINÉTICA QUÍMICA E DETERMINAÇÃO DO PRAZO DE VALIDADE DE PRODUTOS**

# 1 – Introdução

A legislação brasileira estabelece que todos os produtos industrializados, como alimentos, cosméticos, medicamentos, materiais de limpeza doméstica, produtos de higiene pessoal, etc., devem trazer claramente em suas embalagens seus prazos de validade [1, 2]. Além disso, os consumidores atualmente estão mais exigentes por produtos que apresentam desempenho superior: alimentos devem não apenas possuir valor nutricional adequado, mas também características como sabor, odor e textura devem estar de acordo com as preferências do mercado; produtos de beleza devem cumprir seu papel cosmético e/ou funcional, bem como apresentar características relevantes como suavidade, fragrância agradável, etc; e fármacos devem garantir não apenas plena eficácia do princípio ativo, mas também que nenhuma interação entre compostos secundários ocorra durante todo o período de validade do produto [3-6].

O prazo de validade é definido como o período de tempo, a partir da data de fabricação, no qual um produto manterá suas características de qualidade dentro de padrões aceitáveis. A análise da validade pode então ser definida como um protocolo de pesquisa que tem por objetivo reunir evidências de que o produto de fato atende às especificações legais e de mercado até o momento de seu uso pelos consumidores [4-7]. Assim, ao iniciar um estudo de prazo de validade, os pesquisadores devem determinar quais são os parâmetros que definem a qualidade do produto, bem como seus níveis mínimos aceitáveis.

A atribuição correta e livre de ambigüidades dos parâmetros que definem a qualidade de um bem de consumo não é tarefa trivial, pois estes dependem não apenas do produto propriamente dito, mas também de especificidades do mercado uma vez que diferentes grupos de consumidores não atribuem a mesma relevância a parâmetros de qualidade distintos. Todavia, em termos gerais, a qualidade de um produto pode ser dividida em três grupos principais [8, 9]:

**Segurança:** os produtos não podem conter nenhum agente nocivo ao consumidor, quer seja antes, durante ou após o uso. Por exemplo, as embalagens devem conter pigmentos e corantes inertes e não devem apresentar riscos de lesões durante a abertura ou manuseio; os produtos em si também não podem conter nem permitir o desenvolvimento de agentes nocivos, quer sejam microbiológicos ou químicos.

A escassez de recursos e a crescente conscientização dos consumidores levaram as indústrias a considerar também os impactos ambientais da produção, distribuição, uso e descarte de seus produtos. Exemplos de medidas ambientais incluem o uso de materiais de embalagem biodegradáveis e/ou oriundos de fontes sustentáveis e o desenvolvimento e distribuição de produtos mais concentrados – aumentando a eficácia do transporte, reduzindo a emissão de CO<sub>2</sub> e otimizando a utilização de água.

O pleno atendimento dos requisitos de segurança, entretanto, não adiciona valor comercial ao produto do ponto de vista financeiro uma vez que os consumidores entendem que a observância destes quesitos é condição mínima para comercialização de qualquer produto. Além disso, existem órgãos governamentais de regulação, como a Agência Nacional de Vigilância Sanitária (ANVISA), responsáveis por garantir que apenas produtos que atendam a padrões rígidos de qualidade, do ponto de vista de segurança ao consumidor e meio-ambiente, sejam comercializados. Todavia, qualquer mínima falha no atendimento dos requisitos de segurança causa grande dano à imagem de uma marca ou indústria.

**Características primárias ou funcionais** são aquelas que os produtos devem inexoravelmente possuir para cumprir sua função principal. Por exemplo, um fármaco deve possuir eficácia comprovada sobre a doença ou moléstia à qual se destina; uma loção corporal deve hidratar a pele; um composto energético deve possuir uma quantidade mínima de calorias, e assim por diante. Geralmente as características primárias são as que atendem a necessidades específicas dos consumidores e, por esta razão, devem ser eficazes no momento do uso. De fato, a falha em entregar o benefício funcional prometido é prontamente percebida pelo consumidor. Em alguns casos, como o de fármacos, esta falha constitui crime. Em outras instâncias, que podem ocorrer com cosméticos e alimentos, esta falha pode provocar a retirada do produto do mercado, com altíssimos custos envolvidos, além de prejuízos à imagem da marca e da empresa.

Todavia, para alguns ramos da indústria (especialmente no caso dos bens de consumo) as propriedades funcionais não são necessariamente as que mais agregam valor de mercado ao produto, nem tampouco as que direcionam a preferência dos consumidores. Essas propriedades não-funcionais são conhecidas como secundárias.

**Características secundárias** não adicionam nenhuma funcionalidade específica aos produtos, mas trazem benefícios adicionais para os consumidores. Geralmente agregam tanto quanto ou, não raras vezes, maior valor de mercado aos produtos que aquele trazido

por suas próprias características principais. Além disso, dentro de uma mesma classe de produtos, são as características secundárias que trazem diferenciação entre marcas e variantes distintas e, por este motivo, são extensivamente exploradas nas campanhas publicitárias dos bens de consumo.

Exemplos de características secundárias são cores e texturas em alimentos; fármacos que dissolvem por efervescência; produtos em pó de dissolução instantânea; e detergentes e sabonetes líquidos de maior viscosidade e rendimento. Fora do âmbito da composição dos produtos, embalagens de fácil manuseio e abertura, ou que podem ser fechadas depois de abertas, ou ainda que possam ser reutilizadas para outras finalidades após o uso do produto também constituem bons exemplos de características secundárias.

Além da determinação dos parâmetros que definem a qualidade, outro grande desafio na condução de estudos de validade de produtos é a determinação de seus níveis aceitáveis. Apesar de existir legislação específica para uma série de parâmetros, estas normas usualmente se restringem às características relacionadas à segurança do consumidor e, em alguns casos, a algumas características funcionais dos produtos. Mas qual seria, por exemplo, o tempo de dissolução aceitável para um medicamento efervescente? Até que ponto a coloração de um detergente em pó pode alterar-se com o tempo? Ou qual seria a textura ideal de uma barra de cereais? Esses parâmetros não têm impacto na segurança ou nas funcionalidades (ação medicamentosa, limpeza de roupas ou valor nutricional), mas podem causar a imediata rejeição dos produtos por parte dos consumidores. Por este motivo, as indústrias gastam quantidades significativas de recursos em pesquisas de mercado a fim de determinar quais são as características relevantes para os consumidores, bem como seus níveis mínimos aceitáveis, com o objetivo de desenvolver produtos de qualidade superior que lhes tragam vantagem competitiva no mercado [10, 11].

## **2 - Objetivos**

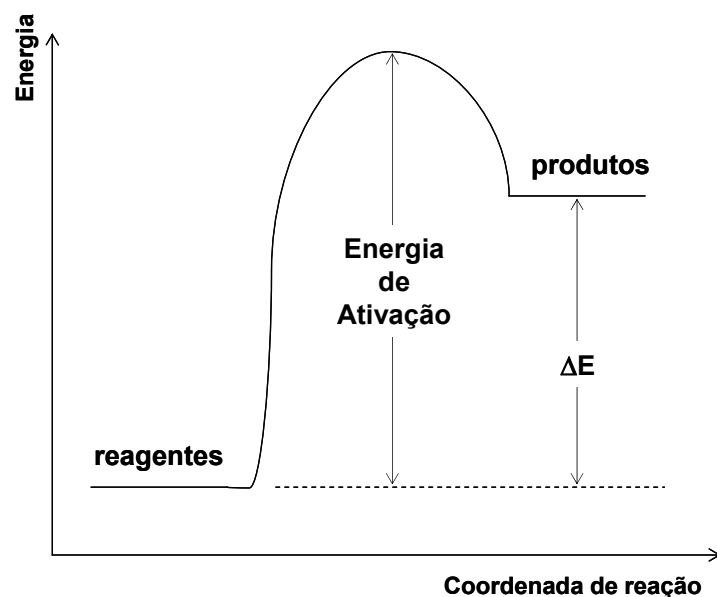
Este capítulo tem por objetivos apresentar a teoria atualmente utilizada na determinação do prazo de validade de bens de consumo, as relações que guarda com a teoria cinética das reações químicas, suas aproximações e considerações, bem como os aspectos práticos para a condução de estudos robustos.

## 3 – Cinética Química

### 3.1 - Introdução

As leis da termodinâmica determinam quando uma reação química será espontânea comparando as energias e entropias dos estados inicial e final do sistema. Entretanto, a termodinâmica não considera em quanto tempo esse processo ocorrerá, focando apenas nas diferenças entre os estados inicial e final do sistema [12, 13]. O ramo da ciência que se dedica a estudar a velocidade em que os processos químicos ocorrem é denominado cinética (do grego *kinesis*, movimento). A cinética química inclui a investigação de como diferentes condições do sistema podem influenciar a velocidade dos processos químicos, provendo informações sobre mecanismos de reação e estados de transição, além de modelos matemáticos que descrevem as características destes processos [12-16].

A experiência mostra que os reagentes não são instantaneamente convertidos em produtos em uma reação química. Várias observações demonstram que algumas reações simplesmente não ocorrem a menos que uma certa quantidade de energia seja fornecida ao sistema. Pode-se então postular que um sistema em transformação deve passar por um máximo de energia durante o curso da reação, localizado entre as energias de seu estado inicial e final. Esta posição de energia máxima é conhecida como estado de transição da reação, e a diferença com a energia inicial do sistema é conhecida como barreira de energia da reação. A Figura 1.1 mostra um diagrama de energia genérico para uma reação elementar, ou seja, que ocorre em uma única etapa [13-18].

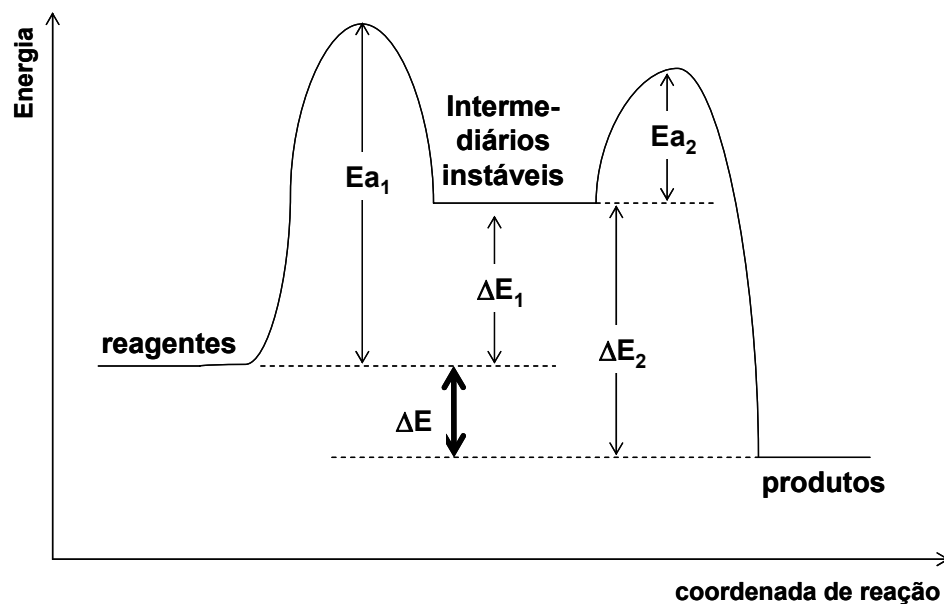


**Figura 1.1** – Diagrama de energia pictórico de uma reação química elementar.

Vários sistemas químicos em transformação são constituídos por duas ou mais reações elementares e, neste caso, são denominados reações complexas ou compostas (Figura 1.2). Neste caso, há a formação de um intermediário instável no ponto de energia mínima entre os dois estados de transição. Este intermediário não pode ser caracterizado como um estado de transição porque não ocorre em um ponto de máxima energia no caminho de reação. Uma interpretação alternativa consiste em considerar o intermediário como o produto da primeira reação elementar e que constitui, por sua vez, o reagente da segunda reação elementar [12, 13, 14-16].

Através da descrição seqüencial de todas as reações elementares que ocorrem entre o estado inicial e final do sistema pode-se postular o mecanismo da reação, assim como também é possível estimar as estruturas químicas e a estequiometria dos estados de transição de cada uma das reações elementares. Também é possível determinar equações matemáticas capazes de quantificar o comportamento do sistema, isto é, de definir as concentrações dos reagentes e/ou produtos em qualquer intervalo de tempo transcorrido após um tempo inicial arbitrário. A estas equações dá-se o nome de leis de velocidade [12-18].





**Figura 1.2** – Diagrama de energia para uma reação complexa, composta por duas reações elementares.

### 3.2 – Leis de Velocidade

Uma reação química genérica ocorrendo em um sistema fechado (onde não há perda nem ganho de matéria) pode ser representada pela expressão genérica [14-18]:

$$\sum_i^r \nu_i R_i = \sum_j^p \nu_j P_j \quad (1.1)$$

Em que  $\nu$  representa coeficientes estequiométricos,  $r$  e  $p$  correspondem ao número total de reagentes e produtos, respectivamente,  $R$  representa um reagente genérico e  $P$  um produto genérico da reação. A Equação 1.1 pode ser ainda re-escrita como:

$$\sum \nu_i A_i = 0 \quad (1.2)$$

onde  $A$  representa tanto produtos quanto reagentes, mas  $\nu$  é negativo para reagentes e positivo para produtos.

Se  $n_i$  representar a quantidade de matéria do  $i$ -ésimo composto, em mols, o grau de avanço da reação ( $\xi$ ) pode ser definido como [12-15]:

$$\xi = \frac{n_i - n_i^0}{\nu_i} \quad (1.3)$$

em que  $n_i^0$  é a quantidade de matéria do composto  $i$  em um tempo<sup>1</sup>  $\theta = 0$  arbitrário. A velocidade da reação pode então ser assim definida, em termos do  $i$ -ésimo composto, como [12-15]:

$$\frac{d\xi}{d\theta} = \frac{1}{\nu_i} \frac{dn_i}{d\theta} \quad (1.4)$$

em que, como na Equação 1.2,  $\nu_i$  é positivo para produtos e negativo para reagentes. Se o volume total do sistema for mantido constante – uma condição comum para reações em solução ou no estado sólido – é mais conveniente expressar a velocidade da reação em termos de concentração molar. Dividindo ambos os termos da Equação 1.4 pelo volume ( $V$ ) tem-se que [12-18]:

$$v = \frac{d(\xi/V)}{d\theta} = \frac{1}{\nu_i} \frac{dn_i}{Vd\theta} = \frac{1}{\nu_i} \frac{d[C]_i}{d\theta} \quad (1.5)$$

em que  $v$  é a velocidade da reação e  $[C]_i$  é a concentração molar de um composto genérico da reação. Note que a convenção de sinais de  $\nu_i$  garante que a velocidade da reação será sempre uma grandeza positiva.

Apesar de  $d[C]_i/d\theta$  ser uma taxa de variação, a negligência do coeficiente estequiométrico  $\nu_i$  nas Equações 1.4 e 1.5 pode gerar ambigüidade na determinação da velocidade da reação. Por exemplo, a velocidade da reação genérica [12-18]:



será, de acordo com a Equação 1.5, dada por:

---

1 - neste trabalho tempo será representado pela letra grega  $\theta$  para que não haja ambigüidade com a notação da Análise de Componentes Principais (PCA – *Principal Component Analysis*), onde  $t$  representa *scores*.

$$v = -\frac{d[A]}{d\theta} = -\frac{1}{2} \frac{d[B]}{d\theta} = \frac{1}{3} \frac{d[C]}{d\theta} \quad (1.6)$$

Certamente a taxa de variação da concentração do composto *A* é duas vezes maior que aquela do composto *B* porque duas moléculas deste último reagem com uma única molécula de *A* [12-18].

À relação matemática que descreve a velocidade de uma reação química em função da concentração das espécies do sistema (usualmente a temperatura e pressão constantes) dá-se o nome de *equação* ou *lei de velocidade*, que é genericamente expressa como [12-18]:

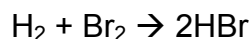
$$v = \frac{1}{\nu_P} \frac{d[P]}{d\theta} = \frac{1}{\nu_Q} \frac{d[Q]}{d\theta} = \dots = k[P]^m[Q]^s \dots \quad (1.7)$$

em que *k* é a constante de velocidade (também conhecida como coeficiente de velocidade) e os expoentes *m* e *s* são definidos como as ordens da reação com relação, respectivamente, às espécies *P* e *Q*. A constante de velocidade *k* tem sinal positivo se a velocidade da reação é expressa como função de um produto da reação, e negativo, se a velocidade for expressa como função de um reagente [12-16].

A constante de velocidade *k* independe das concentrações de quaisquer das espécies envolvidas na reação. É importante também salientar que a constante de velocidade e a velocidade da reação propriamente dita são grandezas diferentes. Entretanto, em uma equação de velocidade simples, *k* pode ser interpretada como o número de mols por litro reagindo por unidade de tempo quando as concentrações de todas as espécies do sistema são iguais a 1 mol.L<sup>-1</sup>. Devido a esta interpretação, a constante de velocidade é também conhecida como velocidade específica [12-18].

A soma dos expoentes de uma lei de velocidade é a ordem geral da reação. Entretanto, algumas leis de velocidade têm formatos diferentes daquele apresentado na Equação 1.7 – por exemplo, as concentrações podem aparecer no denominador ou a ordem de reação com relação a algum composto pode ser fracionária – e, nestes casos, o conceito de ordem geral da reação não é aplicável. Leis de velocidade que seguem o formato da Equação 1.7 são denominadas equações de velocidade simples, enquanto que aquelas

que apresentam formatos diferentes são conhecidas como equações de velocidade complexas [12, 13, 15-17]. Por exemplo, determinou-se experimentalmente que a reação em estado gasoso:



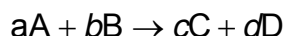
é dada pela seguinte equação de velocidade complexa [13]:

$$v = \frac{k_1[\text{H}_2][\text{Br}_2]^{1/2}}{1 + k_2[\text{HBr}][\text{Br}_2]} \quad (1.8)$$

Neste caso, nota-se que a reação não possui ordem definida e, além disso, apresenta duas constantes de velocidade distintas. A lei de velocidade expressa pela Equação 1.8 sugere ainda que a formação de HBr a partir de  $\text{H}_2$  e  $\text{Br}_2$  dá-se de acordo com um mecanismo de reação complexo [12, 13, 16-19].

Apesar de ser possível que a ordem da reação em relação a um determinado composto seja igual ao seu coeficiente estequiométrico na reação química, este não é o caso geral e a ordem da reação com relação a cada espécie deve necessariamente ser determinada empiricamente através de técnicas experimentais bem definidas. Dentre as técnicas mais empregadas encontram-se os métodos do isolamento, das concentrações iniciais e do ajuste [12-15].

Assumindo-se que a reação genérica:



possua a lei de velocidade simples:

$$v = k[\text{A}]^m[\text{B}]^n$$

(1.9)

então a constante de velocidade  $k$  poderia ser determinada se as ordens de reação  $m$  e  $n$  fossem conhecidas. No método do isolamento um dos reagentes é adicionado em grande

excesso ao sistema reacional de modo que sua concentração durante a reação possa ser considerada constante. Assim, a Equação 1.9 pode ser re-escrita, de maneira simplificada, como [12-15]:

$$v = k'[B]^n \quad (1.10)$$

em que  $k' = k[A]^m$  é constante e é denominada constante de velocidade observável. A Equação 1.10 representa uma reação de **pseudo**-ordem  $n$  e pode ser determinada medindo-se como a concentração de  $B$  varia no tempo.

Analogamente, se a concentração de  $B$  for mantida em excesso durante a reação, a Equação 1.9 pode ser re-escrita como:

$$v = k''[A]^m \quad (1.11)$$

onde  $k'' = k[B]^n$  também é uma constante e a pseudo-ordem da reação com relação a  $A$  pode ser determinada de maneira análoga àquela descrita acima. Apesar deste exemplo genérico apresentar apenas dois reagentes, é importante notar que a técnica do isolamento pode ser empregada a sistemas contendo um número maior de espécies [12-15].

Em algumas reações a adição de um reagente em largo excesso é impraticável, quer seja porque este reagente afeta significativamente o equilíbrio – e conseqüentemente a velocidade – da reação, quer seja porque altera o mecanismo que a reação seguiria caso estivesse em concentrações mais baixas, tornando impossível aplicar o método do isolamento. Uma alternativa para estes casos é o método das concentrações iniciais. Se for possível medir a velocidade da reação sob várias concentrações de  $A$  e  $B$ , então as ordens da reação com relação às espécies envolvidas na lei de velocidade, bem como a constante de velocidade  $k$ , podem ser determinadas de maneira direta ajustando-se os dados experimentais à Equação 1.9. Apesar de a mediação direta da diferencial  $d[A]/d\theta$  ser impraticável para a maioria das reações, a alteração na concentração de  $[A]$  pode ser medida para um intervalo finito de tempo  $\Delta\theta$ . Deste modo [12-15]:

$$v = -\frac{1}{\nu_A} \frac{d[A]}{d\theta} \approx -\frac{1}{\nu_A} \frac{\Delta A}{\Delta\theta} = k[A]^m[B]^n \quad (1.12)$$

Quanto menor for o intervalo de tempo  $\Delta\theta$  menor será a diferença entre  $d[A]/d\theta$  e  $\Delta A/\Delta\theta$ . Considerando-se que duas medidas da velocidade da reação são realizadas em um intervalo finito de tempo e que a concentração inicial de  $A$  seja igual a  $[A]_0$  em ambas as medidas mas a concentração de  $B$  seja  $[B]_1$  e  $[B]_2$ , as velocidades para cada reação serão dadas por [12, 13]:

$$v_1 = -\frac{1}{\nu_A} \left( \frac{\Delta A}{\Delta\theta} \right)_1 = k[A]_0^m [B]_1^n \quad (1.13)$$

$$v_2 = -\frac{1}{\nu_A} \left( \frac{\Delta A}{\Delta\theta} \right)_2 = k[A]_0^m [B]_2^n \quad (1.14)$$

em que os índices 1 e 2 são utilizados para diferenciar os dois experimentos realizados com concentrações iniciais diferentes de  $B$ . Dividindo-se a Equação 1.13 pela 1.14 e tomando-se os logaritmos, é possível resolver o sistema para  $n$ , sendo:

$$n = \frac{\ln \frac{v_1}{v_2}}{\ln \frac{[B]_1}{[B]_2}} \quad (1.15)$$

Analogamente, se o experimento for repetido para concentrações iniciais constantes de  $B$  (ou seja,  $[B] = [B]_0$ ) enquanto as concentrações iniciais de  $A$  são variadas, a ordem da reação com relação a  $A$  pode ser determinada.

Outra maneira prática de determinar as ordens e a constante de velocidade de reações simples é através do ajuste dos dados experimentais às soluções integradas das leis de velocidade. Por exemplo, a integração da equação de velocidade de ordem zero<sup>1</sup> [12-15]:

---

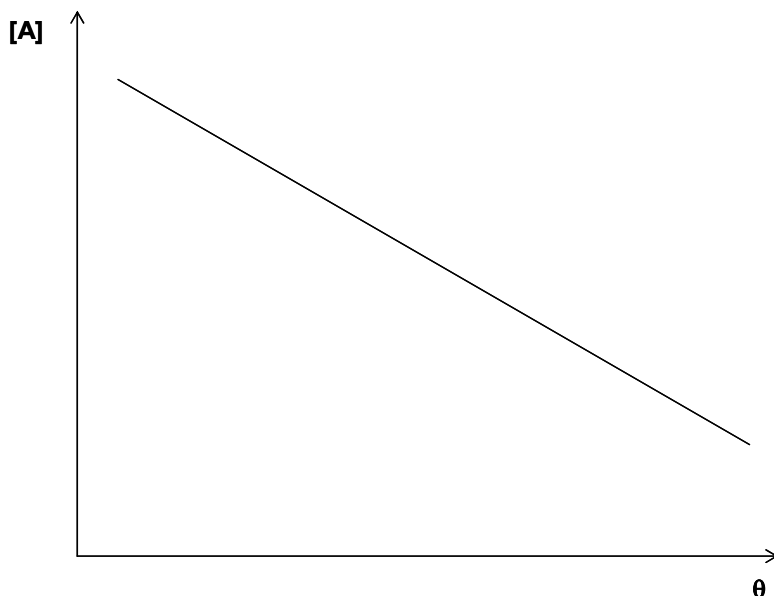
1 - é interessante notar que a velocidade de uma reação que segue cinética de ordem zero é independente da concentração dos reagentes ou dos produtos no tempo. fornece a solução:

$$v = -\frac{1}{\nu_A} \frac{d[A]}{d\theta} = k[A]^0 = k \quad (1.16)$$

$$[A]_{\theta} - [A]_{\theta_0} = k(\theta - \theta_0) \quad (1.17)$$

em que  $k$  deve ter dimensões de velocidade ( $\text{mol.L}^{-1}.\text{s}^{-1}$ ) e tem sinal negativo se  $A$  é consumido na reação e positivo caso contrário.

Assim, se a cinética da reação for realmente de ordem zero, um gráfico de  $\Delta[A]$  por  $\Delta\theta$  deve fornecer uma reta com coeficiente angular cujo módulo vale  $k$  (Figura 1.3).



**Figura 1.3** – Gráfico da variação da concentração da espécie  $A$  com o tempo para uma reação de cinética de ordem zero.

Reações que possuem cinética de primeira ordem (Figura 1.4) possuem a seguinte lei de velocidade [12-18]:

$$v = -\frac{1}{\nu_A} \frac{d[A]}{d\theta} = k[A]^1 \quad (1.18)$$

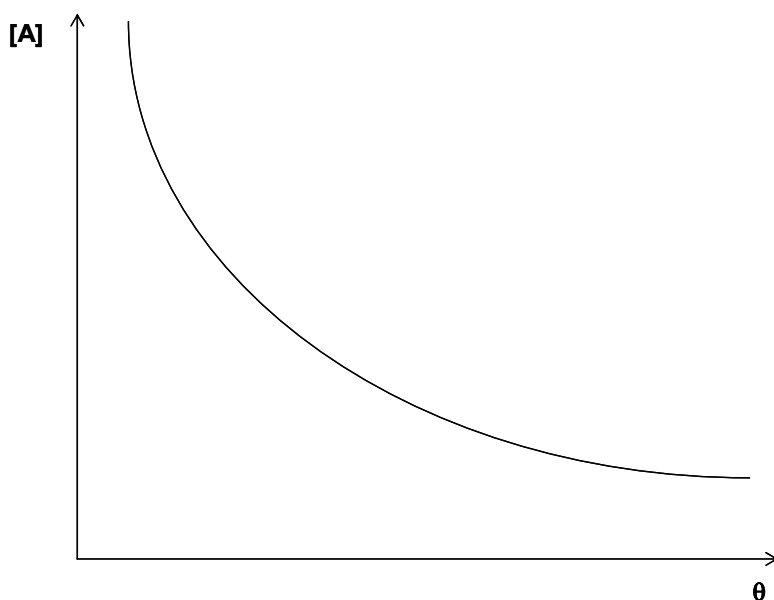
cuja solução integrada é dada por:

$$[A]_{\theta} = [A]_{\theta_0} e^{k(\theta - \theta_0)} \quad (1.19)$$

onde  $k$  tem dimensões de frequência ( $s^{-1}$ ) e é negativo se  $A$  é consumido na reação e positivo caso contrário. A constante de velocidade  $k$  pode ser determinada re-escrevendo a Equação 1.19 como [12-18]:

$$\ln\left(\frac{[A]_{\theta}}{[A]_{\theta_0}}\right) = k(\theta - \theta_0) \quad (1.20)$$

A Equação 1.20 mostra que um gráfico de  $\ln([A]_{\theta}/[A]_{\theta_0})$  por  $\Delta\theta$  fornecerá uma reta cujo coeficiente angular é igual a  $k$ .



**Figura 1.4** - Gráfico da variação da concentração da espécie  $A$  com o tempo para uma reação de cinética de primeira ordem.

Um conceito importante em cinética química é o de meia-vida ( $\theta_{1/2}$ ), que é definido como o intervalo de tempo necessário para que a concentração de um reagente seja igual à metade do valor que possuía no início da reação. Deste modo, para uma reação de ordem zero  $\theta_{1/2} = \theta_{[A]=1/2[A]_0}$ . Para uma reação de cinética de primeira-ordem, o tempo necessário para que a concentração de  $A$  seja  $1/2[A]_0$  será dado, de acordo com a Equação 1.20, por [12-15]:



$$k\theta_{1/2} = -\ln\left(\frac{1/2[A]_{\theta_0}}{[A]_{\theta_0}}\right) = -\ln\frac{1}{2} = \ln 2 \quad (1.21)$$

$$\theta_{1/2} = \frac{\ln 2}{k} \quad (1.22)$$

onde, arbitrariamente,  $\theta_0 = 0$ . A Equação 1.22 mostra que a meia-vida de um reagente é independente de sua concentração inicial para reações de primeira ordem [12-15].

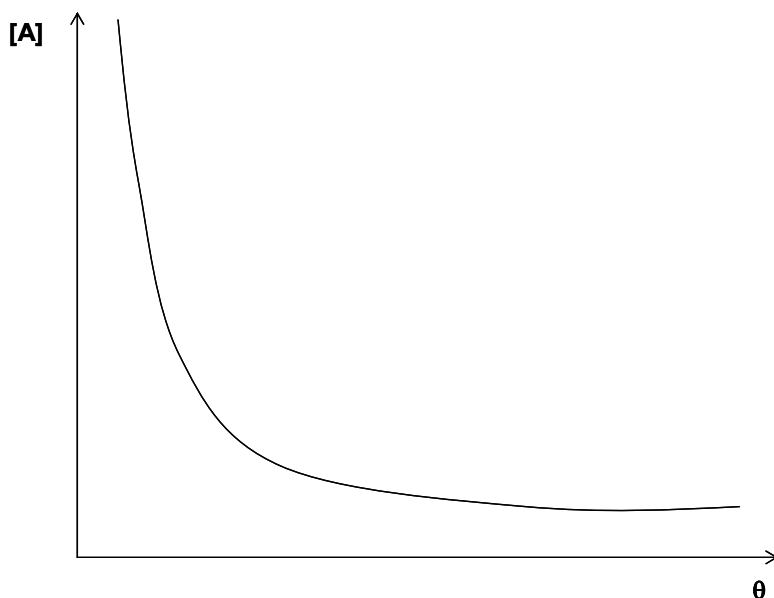
Existem duas maneiras possíveis de expressar leis de velocidade de reações que seguem cinética de segunda ordem (Figura 1.5) Na primeira, a velocidade da reação depende apenas da concentração de um dos reagentes elevado ao quadrado [12-15]:

$$v = -\frac{1}{\nu_A} \frac{d[A]}{d\theta} = k[A]^2 \quad (1.23)$$

A integração da Equação 1.23 fornece a seguinte solução:

$$\frac{1}{[A]_{\theta}} - \frac{1}{[A]_{\theta_0}} = k(\theta - \theta_0) \quad (1.24)$$

e mostra que  $k$  tem dimensões de  $\text{L}\cdot\text{mol}^{-1}\cdot\text{s}^{-1}$  e será negativo se  $A$  for consumido no tempo.



**Figura 1.5** - Gráfico da variação da concentração da espécie A com o tempo para uma reação de cinética de segunda ordem, de acordo com a Equação 1.24.

Se a reação for de segunda ordem com relação a A, um gráfico de  $1/[A]_{\theta}$  por  $\theta$  fornecerá uma linha reta com coeficiente angular igual a  $k$  e intercepto igual a  $1/[A]_{\theta_0}$ .

Para reações que seguem cinética de segunda ordem descritas conforme a equação 1.24, se  $\theta = \theta_{1/2}$  e  $[A]_{\theta} = 1/2[A]_{\theta_0}$ , então a meia-vida do reagente A será dada por:

$$\theta_{1/2} = \frac{1}{k[A]_{\theta_0}} \quad (1.25)$$

Note que, diferentemente de reações de primeira ordem, a meia-vida de reagentes que seguem cinética de segunda ordem depende de sua concentração inicial. Uma consequência prática deste fato é que compostos que decaem de acordo com uma cinética de segunda ordem tendem a persistir por longos períodos quando em baixas concentrações porque suas meias-vidas são longas nestas condições [12-18].

Uma reação também pode seguir cinética de segunda ordem quando dois reagentes quaisquer que fazem parte da lei de velocidade possuírem cinética de primeira ordem, como mostrado na Equação 1.26 [12-16]:

$$v = -\frac{d[A]}{d\theta} = k[A][B] \quad (1.26)$$

cuja integração fornece a solução:

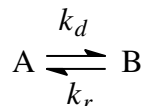
$$\frac{1}{[A]_{\theta_0} - [B]_{\theta_0}} \ln \left( \frac{[A]_{\theta} [B]_{\theta_0}}{[B]_{\theta} [A]_{\theta_0}} \right) = k(\theta - \theta_0) \quad (1.27)$$

Note que a Equação 1.27 é indeterminada na condição  $[A]_{\theta_0} = [B]_{\theta_0}$ . Neste caso, esta condição deve ser incluída na integração da Equação 1.26 [12-16].

### 3.3 – Reações Reversíveis e em Equilíbrio

Uma reação reversível pode ocorrer em ambas direções - isto é, produtos formam reagentes, e vice-versa - e é representada por duas setas, uma referente à reação direta e outra, à reversa. Exemplos clássicos de reações reversíveis são a ionização de eletrólitos em solução e a isomerização *cis-trans* [12-16].

Considere a seguinte reação reversível genérica:



em que  $k_d$  e  $k_r$  representam, respectivamente, as constantes de velocidade da reação direta e reversa (alguns autores também adotam a simbologia  $k_1$  e  $k_{-1}$ , onde a constante de velocidade da reação direta recebe um subscrito positivo, enquanto à reversa é atribuído um índice negativo) [13]. Um sistema reversível sempre tende ao equilíbrio para quaisquer concentrações iniciais  $[A]_0$  e  $[B]_0$ . Neste ponto as velocidades das reações direta e reversa se igualam (Equação 1.28) e as concentrações de  $A$  e  $B$  (dadas pela expressão da constante de equilíbrio, Equação 1.29) mantêm-se constantes [12-15]:

$$-\frac{d[A]}{d\theta} = \frac{d[B]}{d\theta} \quad (1.28)$$

$$K_c = \frac{[B]_{\text{eq}}}{[A]_{\text{eq}}} \quad (1.29)$$

É importante notar que, ao contrário do senso comum, os processos reativos não cessam no equilíbrio. De fato, as Equações 1.28 e 1.29 mostram que  $A$  é convertido em  $B$ , e vice-versa, num equilíbrio dinâmico.

Considere-se agora o caso específico em que a reação genérica no equilíbrio tenha uma lei de velocidade de primeira ordem com relação a cada um dos compostos:

$$v = -\frac{d[A]}{d\theta} = \frac{d[B]}{d\theta} = k_d[A] - k_r[B] \quad (1.30)$$

Nota-se que a lei de velocidade consiste da soma de dois termos: o primeiro representa a velocidade da reação direta e o segundo, da reação reversa. A diferença nos sinais de  $k_d$  e  $k_r$  refletem o fato de que o reagente  $A$  é consumido na reação direta e formado na reação reversa [12-15].

Se  $[B]_0 = 0$  então, pela estequiometria da reação,  $[B]_\theta = [A]_0 - [A]$  e a equação 1.30 pode ser re-escrita como:

$$-\frac{d[A]}{d\theta} = (k_d + k_r)[A]_\theta - k_r[A]_{\theta_0} \quad (1.31)$$

A integração da Equação 1.31, sujeita às condições de contorno do equilíbrio, fornece a solução [13]:

$$[A]_\theta = ([A]_{\theta_0} - [A]_{\text{eq}})e^{-(k_d+k_r)\theta} + [A]_{\text{eq}} \quad (1.32)$$

Que, por rearranjo, pode ser re-escrita como:

$$\ln([A]_\theta - [A]_{\text{eq}}) = \ln([A]_{\theta_0} - [A]_{\text{eq}}) - (k_d + k_r)\theta \quad (1.33)$$

A Equação 1.33 mostra que um gráfico de  $\ln([A]_{\theta} - [A]_{\theta q})$  em função do tempo fornece uma reta cujo coeficiente angular será igual a  $k_d + k_r$  e o intercepto corresponderá a  $\ln([A]_{\theta_0} - [A]_{eq})$ .

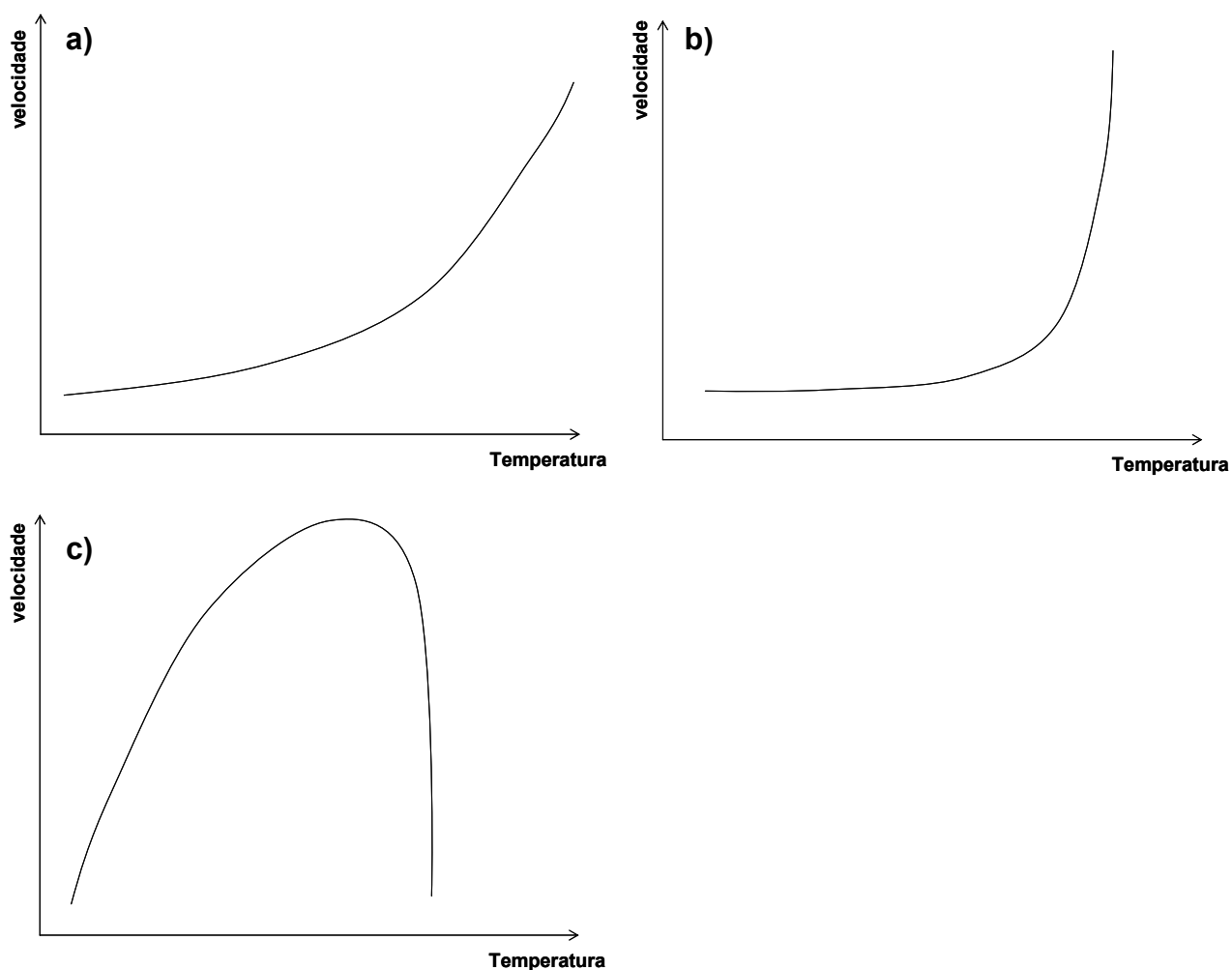
A relação entre a lei de velocidade e a constante de equilíbrio pode então ser utilizada para determinar os valores individuais de  $k_d$  e  $k_r$ . Como no equilíbrio  $d[A]/d\theta = 0$ , a Equação 1.29 pode ser re-escrita como [12-16]:

$$\frac{k_d}{k_r} = \frac{[B]_{eq}}{[A]_{eq}} = K_c \quad (1.34)$$

### 3.3 – Dependência da Lei de Velocidade com a Temperatura

A experiência mostra que a velocidade de uma reação química é dependente da temperatura do sistema. De fato, a síntese prática de muitos compostos de interesse industrial só é possível quando o sistema é mantido sob temperaturas adequadas [12-15, 20-24].

A velocidade de uma grande gama de reações apresenta comportamento exponencial com o aumento da temperatura do sistema (Figura 1.6a). Entretanto, em alguns casos específicos, outros comportamentos são observados. A Figura 1.6b mostra uma reação que se torna explosiva assim que uma temperatura mínima é atingida, enquanto que a Figura 1.6c representa uma típica reação catalisada por enzimas: acima de uma certa temperatura - onde a velocidade de conversão é máxima -, a desnaturação da enzima faz com que a velocidade da reação caia rapidamente [12-19, 22-24].



**Figura 1.6** – Comportamentos típicos das velocidades de reações químicas em função da temperatura: a) aumento exponencial; b) comportamento explosivo e c) catálise enzimática.

Todavia, as relações observadas entre a velocidade de uma reação química e a temperatura do meio reacional não estão explicitamente representadas nas equações cinéticas até aqui apresentadas.

Em 1887, van't Hoff observou, baseado na variação das constantes de equilíbrio ( $K_c$ ), que era a constante de velocidade de uma reação química que dependia da temperatura do sistema. Mais especificamente, ele observou que o logaritmo da constante de velocidade de uma reação deveria seguir uma relação linear com o recíproco da temperatura absoluta, conforme a equação [12-19, 25, 26]:

$$\log k = B - \frac{A}{T} \quad (1.35)$$

onde  $A$  e  $B$  são constantes determinadas empiricamente.

Mais tarde, Arrhenius estendeu a idéia inicial de van't Hoff e a aplicou com sucesso em uma miríade de reações químicas. Arrhenius argumentou que, se a variação na constante de equilíbrio obedece à lei [12-19, 25, 26]:

$$\frac{d \ln K_c}{dT} = \frac{\Delta E}{RT^2} \quad (1.36)$$

onde  $\Delta E$  é a mudança de energia do sistema, na hipótese de um equilíbrio dinâmico (em que a Equação 1.34 é válida), a Equação 1.36 pode ser expressa como:

$$\frac{d(\ln k_d)}{dT} - \frac{d(\ln k_r)}{dT} = \frac{E}{RT^2} \quad (1.37)$$

que pode ser resolvida por separação de variáveis [14-18]:

$$\frac{d(\ln k_d)}{dT} = \frac{E_d}{RT^2} + C \quad (1.38)$$

e:

$$\frac{d(\ln k_r)}{dT} = \frac{E_r}{RT^2} + C \quad (1.39)$$

onde  $E_d - E_r = E$  e  $C$  é uma constante cujo valor Arrhenius demonstrou experimentalmente ser igual a zero. A velocidade da reação é então relacionada à temperatura de acordo com a equação [14-18]:

$$\frac{d(\ln k_d)}{dT} = \frac{E_d}{RT^2} \quad (1.40)$$

que possui a seguinte solução analítica:

$$\ln k_d = -\frac{E_d}{RT} + \text{const.} \quad (1.41)$$

e que, após rearranjo, leva à conhecida Equação de Arrhenius:

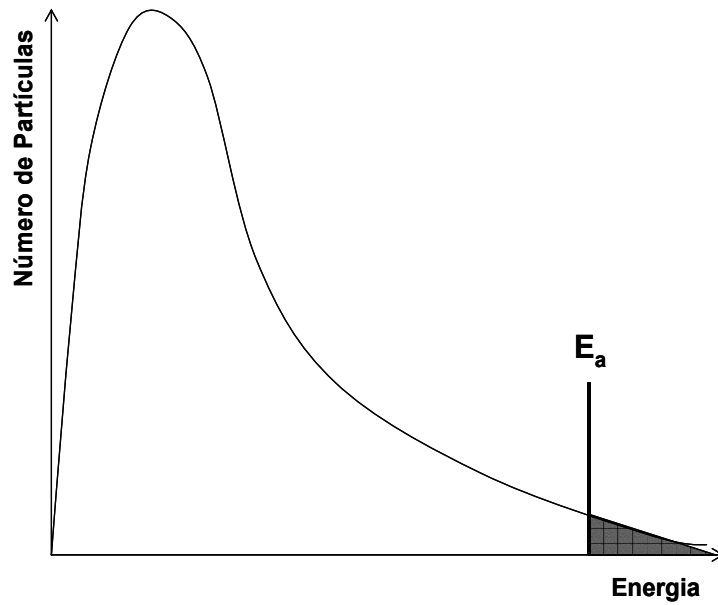
$$k_T = Ae^{-E_a/RT} \quad (1.42)$$

onde  $A$  é conhecido como termo pré-exponencial ou fator de frequência e tem as mesmas unidades de  $k$ , e  $E_a$  é denominada energia de ativação do sistema. Se a constante de velocidade for medida a várias temperaturas, a Equação 1.42 revela que  $E_a$  pode ser determinada experimentalmente através da inclinação da reta do gráfico de  $\ln(k_T)$  por  $1/T$  [12-18].

As considerações e resultados obtidos por Arrhenius continuam válidos até os dias atuais e, por este motivo, a Equação 1.42 é conhecida como Lei de Arrhenius. De fato, a observação de que uma reação química não obedece à lei de Arrhenius constitui evidência de que a reação se dá através de mecanismos complexos [12-18].

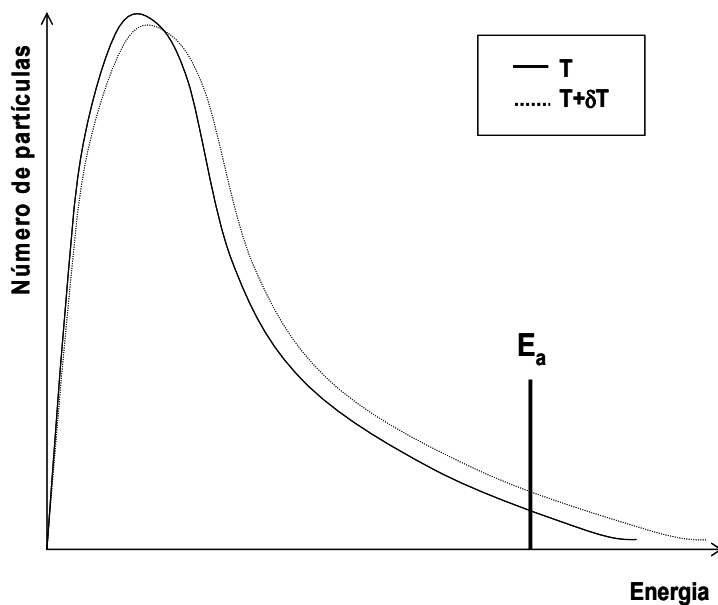
Foge ao escopo deste trabalho fornecer interpretações mecânicas precisas dos parâmetros da Lei de Arrhenius. Estas podem ser encontradas nos trabalhos de Tolman, Truhlar, Leider e Menzinger [26-29]. Uma interpretação simplista, mas conveniente, pode ser formulada da seguinte maneira: para que uma reação química ocorra é necessário que os reagentes colidam com uma energia (cinética mais potencial) mínima igual à energia de ativação ( $E_a$ ). Por exemplo, em um sistema gasoso podem ocorrer inúmeras colisões entre os reagentes, mas estes transformar-se-ão em produtos apenas se as colisões tiverem energia maior ou igual à  $E_a$  (Figura 1.7). Deste modo, o termo pré-exponencial pode ser entendido como uma medida da taxa total de colisões no sistema, independente de sua energia, enquanto que o termo exponencial ( $e^{-E_a/RT}$ ) na Equação 1.42 pode ser interpretado como a fração de colisões que possuem energia suficiente para proporcionar a reação. Assim, o produto  $A \times e^{-E_a/RT}$  fornece uma estimativa do número de colisões reativas no sistema [12-18].





**Figura 1.7** – Diagrama pictórico da distribuição de energia em um sistema reacional.

Quando a temperatura do sistema é aumentada um número maior de partículas popula regiões de mais alta energia (cinética, vibracional, interna) na distribuição de Boltzmann (Figura 1.8), aumentando a probabilidade de ocorrência de colisões reativas e, deste modo, a reação procede mais rapidamente [12-18].



**Figura 1.8** – Distribuição de energia de Boltzmann para o sistema reacional em duas temperaturas distintas. Quando a temperatura do sistema é aumentada de  $T$  para  $T+\delta T$  um número maior de moléculas possui energia maior que  $E_a$ .

## 4 – Estudos de Validade e Estabilidade de Produtos

Apesar de seu reconhecido poder de previsão e sua incrível solidez, a teoria cinética química formal é raramente utilizada para determinação da validade de bens de consumo. Salvo raras exceções, a natureza complexa de produtos de higiene e limpeza, cosméticos e fármacos, que podem possuir dezenas de ingredientes em sua formulação, e de produtos alimentícios, que usualmente apresentam interações entre centenas de compostos diferentes, torna impraticável a definição do prazo de validade através da determinação direta de concentrações de espécies químicas específicas. Mesmo produtos farmacêuticos, nos quais a concentração do composto químico ativo deve ser garantida durante toda sua vida útil, devem ter suas características cuidadosamente estudadas em função do tempo de modo a garantir que interações adversas entre os materiais excipientes não comprometam a qualidade do medicamento, o que poderia resultar em efeitos colaterais indesejados ou rejeição por parte dos consumidores.

Entretanto, a experiência mostra que uma aproximação da teoria cinética química formal poderia ser utilizada para determinação do prazo de validade de produtos. Por outro lado, se a variação da concentração molar de uma espécie química fosse substituída pelos valores medidos de uma propriedade de interesse qualquer, a Equação 1.7 poderia ser reescrita como:

$$v = \frac{dP}{d\theta} = k_T P^n \quad (1.43)$$

onde  $v$  é a velocidade de alteração da propriedade  $P$  no tempo,  $k_T$  é a constante de velocidade observável (ou experimental) e  $n$  é a pseudo-ordem do processo de degradação do produto. Como no caso da teoria cinética química clássica,  $k_T$  será positivo se o valor numérico de  $P$  aumentar com o tempo, e negativo se este valor diminuir. Todavia, é comum em estudos de validade converter o valor medido para porcentagem de alteração (ou porcentagem de perda de qualidade) no tempo, o que eliminaria a necessidade da convenção de sinais.

A equação 1.43 traz implícitas duas premissas importantes. Primeiramente, assume-se que o valor medido da propriedade  $P$  em um tempo  $\theta$  qualquer é devido ao efeito líquido de uma série de reações químicas ou processos físicos. Deste modo, tais processos possuem, por sua vez, uma cinética química total mensurável através do estudo da variação de  $P$ . Uma das vantagens da teoria é que o conhecimento das cinéticas das reações que resultam na alteração da medida de  $P$  não precisam necessariamente ser determinadas para que o prazo de validade do produto seja definido. O conhecimento de alguns processos reacionais e/ou processos físicos envolvidos na degradação do produto é, todavia, fundamental quando se desejam adotar medidas para ampliar o prazo de validade do bem de consumo. Entretanto, para este fim, estudos específicos, geralmente independentes dos testes de validade propriamente ditos, são empregados.

É importante notar que não é necessário que  $P$  seja uma medida instrumental. De fato, há vários relatos de prazos de validade na literatura calculados utilizando-se dados de painéis sensoriais [3-5, 30, 31]. A única condição para que a Equação 1.43 seja aplicada adequadamente é que os dados devem ser numéricos e contínuos.

Outra premissa da teoria, implícita na Equação 1.43, é que a variação da propriedade  $P$  independe da variação de outras propriedades. Esta aproximação é equivalente ao método do isolamento na teoria cinética química clássica e, de fato, é por esse motivo que a reação possui uma pseudo-ordem e a constante de velocidade é denominada “observável” ou “experimental”. Apesar de apresentarem certo grau de arbitrariedade, estas premissas são usualmente robustas no sentido em que desvios da idealidade não alteram significativamente as conclusões dos estudos de validade.

Por conseqüência das premissas da teoria dos estudos de validade, a experiência mostra que a grande maioria das propriedades dos bens de consumo segue cinéticas de pseudo-zero, pseudo-primeira ou pseudo-segunda ordem (ou seja,  $n = 0, 1$  ou  $2$ ) – e, assim, a integração da Equação 1.43 leva às seguintes soluções analíticas:

$$P_{\theta} - P_{\theta_0} = k_T \theta, \quad \text{para processos de ordem pseudo-zero} \quad (1.44)$$

$$\ln\left(\frac{P_{\theta}}{P_{\theta_0}}\right) = k_T \theta, \quad \text{para processos de pseudo-1ª ordem} \quad (1.45)$$

$$\frac{1}{P_{\theta}} - \frac{1}{P_{\theta_0}} = k_T \theta, \quad \text{para processos de pseudo-2ª ordem} \quad (1.46)$$

onde  $P_{\theta}$  e  $P_{\theta_0}$  são, respectivamente, os valores da propriedade  $P$  nos tempos  $\theta = \theta_0 + \delta\theta$  e  $\theta_0$ . Deste modo, tanto as pseudo-ordens como as constantes de velocidade experimentais dos processos de degradação do produto podem ser determinadas construindo-se gráficos lineares do tempo em função da transformação apropriada dos valores medidos da propriedade  $P$  no tempo, conforme já explanado na seção 3.2 deste capítulo.

Leis de velocidade diferentes da expressa na Equação 1.43 acima foram observadas para propriedades específicas em alguns produtos. Por exemplo, Farmer *et al.* determinaram que a oxidação de lipídeos segue cinética complexa (Equação 1.47), o que é esperado para reações radicalares [3-5, 30, 31]. Além disso, apesar do crescimento microbiológico seguir cinética de primeira ordem (ou crescimento exponencial) após a fase lag para uma grande gama de produtos, estudos específicos levaram à definição de vários outros modelos, entre eles o de Zwietering (Equação 1.48)<sup>1</sup> [7, 33, 34].

$$-\frac{d[O_2]}{d\theta} = \frac{ka[RH][ROOH]}{1 + \lambda[RH]/p} \quad (1.47)$$

$$\log(\text{CFU}) = K + A \times \exp\left\{-\exp\left\{\left[2.7182 \times \mu_{\max} \frac{\lambda - t}{A}\right] + 1\right\}\right\} \quad (1.48)$$

---

1 – As equações 1.47 e 1.48 estão apresentadas apenas com função demonstrativa. Foge ao escopo deste trabalho detalhar suas premissas, bem como explicar seus parâmetros.

#### **4.1 - Estudos Acelerados de Validade (ASLT – *Accelerated Shelf-Life Tests*)**

Pela teoria até aqui apresentada, a determinação do prazo de validade de um produto requereria que o estudo de validade fosse conduzido até que um parâmetro atingisse um nível inaceitável de qualidade quando estocado em temperatura ambiente. Apesar deste fato não ser restritivo para produtos de curta vida de prateleira, seria inviável conduzir um estudo completo para bens de consumo de longa validade porque novos lançamentos no mercado seriam significativamente atrasados. Além disso, uma quantidade considerável de recursos seria despendida sem nenhum retorno imediato. Assim, a condução de estudos acelerados de validade é prática comum na indústria de bens de consumo.

Estudos acelerados consistem em submeter o produto a alguma condição de estocagem mais severa do que aquela comumente encontrada no mercado de modo que a degeneração de seus parâmetros de qualidade ocorra mais rapidamente que na condição ambiente. As principais formas de aceleração dos estudos de validade consistem em submeter o produto a [3-7, 37-45]:

1. temperaturas relativamente elevadas;
2. atmosferas com variação de umidade relativa;
3. ciclos de temperatura;
4. fontes de radiação, normalmente ultravioleta;
5. transições de fases;
6. atmosferas modificadas.

A utilização de fontes de radiação e atmosferas modificadas (utilizando gases oxidantes ou inertes, dependendo dos objetivos do teste) é menos comum que as demais. São usualmente empregadas a produtos sensíveis à oxidação (onde gás oxigênio e/ou radiação ultravioleta são geralmente utilizados), como óleos e alguns polímeros [3-7, 41]. Transições de fases também não são muito comuns em estudos de validade, sendo quase

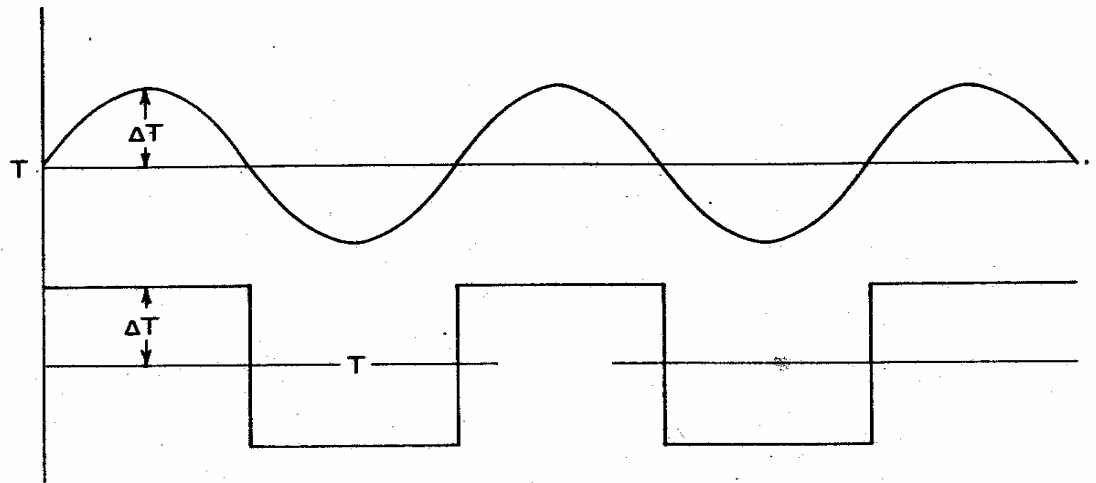
que exclusivamente dedicadas ao estudo de produtos cristalinos ou a alimentos congelados [3-7, 33, 40].

#### ***4.1.1 – Estocagem em Condições de Temperatura Elevadas***

A estocagem a temperaturas elevadas constitui o principal modo de aceleração da degradação de produtos, principalmente porque é de fácil implementação. O princípio deste modo de aceleração vem da observação de que a velocidade da maioria das reações de degradação dos produtos aumenta com o incremento da temperatura de acordo com a equação de Arrhenius (Equação 1.42) [3-5, 7, 35-38].

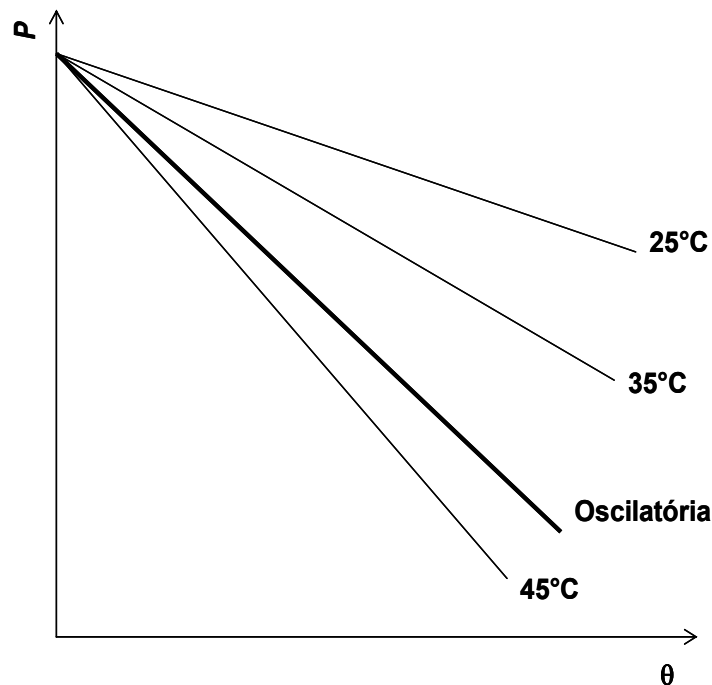
Em estudos de validade, a energia de ativação define a sensibilidade térmica de um parâmetro de qualidade do produto: quanto maior a energia de ativação, mais sensível é a propriedade em estudo a alterações de temperatura [3-5, 35-38]. Esta informação é de extrema relevância, principalmente quando o bem de consumo será comercializado em vários mercados, com condições climáticas distintas, ou será submetido a condições mais drásticas de temperatura durante o transporte. Nestes casos, medidas de preservação adequadas, como transporte ou estocagem sob refrigeração, podem ser adotadas [6, 7, 38].

Em alguns casos pode-se estocar o produto sob condições controladas de flutuação de temperatura. Câmaras climáticas modernas podem ser programadas com ciclos de temperatura que simulam com fidelidade considerável o perfil de temperatura ao qual o produto estará submetido durante sua distribuição e estocagem. Muitos destes perfis podem ser aproximados por funções de onda quadradas ou senoidais (Figura 1.9) e o prazo de validade final do produto pode ser calculado usando as Equações fornecidas por Labuza [3, 5, 36].



**Figura 1.9** – Oscilações de temperatura por funções de onda senoidal e quadrada.

Todavia, a utilização de oscilação de temperatura é restrita pois determinou-se que, na maioria dos estudos práticos, os resultados de estudos que utilizam este modo de aceleração são equivalentes àqueles obtidos tomando-se a média ponderada dos intervalos mínimo e máximo de temperatura no estudo (Figura 1.10). Além disso, o prazo de validade prático de bens de consumo é geralmente menor que aquele determinado matematicamente em estudos de validade de modo a acomodar fontes de variação não inclusas nos estudos, como transporte, irradiação solar, etc [3-7, 37-45].



**Figura 1.10** – Comparação entre as velocidades de reação de produtos estocados em condições constantes e com oscilação de temperatura.

Existem estudos de validade onde a condição de estocagem não é controlada - isto é, amostras do produto são estocadas em condições ambiente – e, nestes casos, ocorrem variações não-sistemáticas de temperatura. Um protocolo de estudo prevendo estas considerações, desenvolvido pelo USDA (*United States Department of Agriculture*) em 1960 e referido como TTT (Tolerância Tempo-Temperatura), tem de fato por objetivo simular condições reais de transporte, manuseio e estocagem do produto durante sua distribuição e utilização pelo consumidor [7, 38, 39, 46].

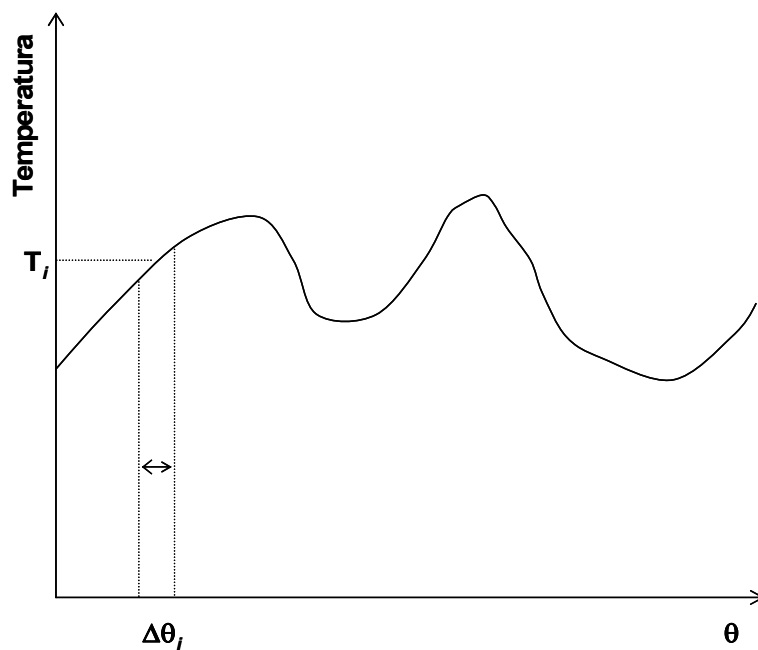
Nos estudos de validade sob condições não-controladas faz-se necessário que o histórico de temperatura onde as amostras do produto permaneceram estocadas seja medido com uma frequência adequada durante todo o estudo de validade. O histórico é então dividido em intervalos pequenos de tempo de modo que a temperatura média possa ser calculada de maneira representativa (Figura 1.11). Utilizando os gráficos cinéticos ou de validade é possível utilizar as equações 1.49 para parâmetros que seguem cinéticas de pseudo-ordem zero e 1.50 para propriedades com cinética de pseudo-primeira ordem [3, 7, 38]:

$$P_{\theta} - P_{\theta_0} = \sum_i k_i \Delta \theta_i, \quad \text{para processos de ordem pseudo-zero} \quad (1.49)$$

$$P_{\theta} = P_{\theta_0} \exp\left(\sum_i k_i \Delta \theta_i\right), \quad \text{para processos de pseudo-primeira ordem} \quad (1.50)$$

onde as constantes de velocidade para cada temperatura média são calculadas utilizando-se os intervalos de tempo correspondentes nos gráficos cinéticos e  $i$  representa índices relativos aos intervalos de tempo.





**Figura 1.11** – Exemplo de histórico de temperatura de um produto estocado em condição ambiente não-controlada.

#### ***4.1.2 – Estudos de Validade com Variação da Umidade Relativa de Estocagem***

O entendimento da velocidade com que um produto absorve vapor de água é de extrema relevância para uma grande gama de bens de consumo sensíveis à umidade excessiva do ambiente, como é o caso de produtos higroscópicos secos como sabões em pó, temperos, comprimidos, pastilhas, etc. Nestes casos o ganho de água pode levar à degradação microbiológica, perda de textura, amolecimento ou endurecimento e agregação de produtos em pó. Por outro lado, a perda de água para o ambiente também pode trazer alterações reológicas indesejadas como endurecimento, perda de textura ou tornar o produto frágil e quebradiço. Além disso, existem regulamentações específicas quanto à perda máxima de peso permitida para vários produtos.

Bens de consumo sensíveis à umidade devem ser preservados através da utilização de embalagens com baixa permeabilidade à água, determinada através da WTR – *Water Transmission Rate*. Dados de WTR podem ser obtidos dos fornecedores das embalagens ou através de experimentos [3-5, 38, 48 51, 52].

A perda ou ganho de água depende da concentração efetiva de água nos produtos, mais comumente conhecida como atividade de água<sup>1</sup> ( $a_w$ ).

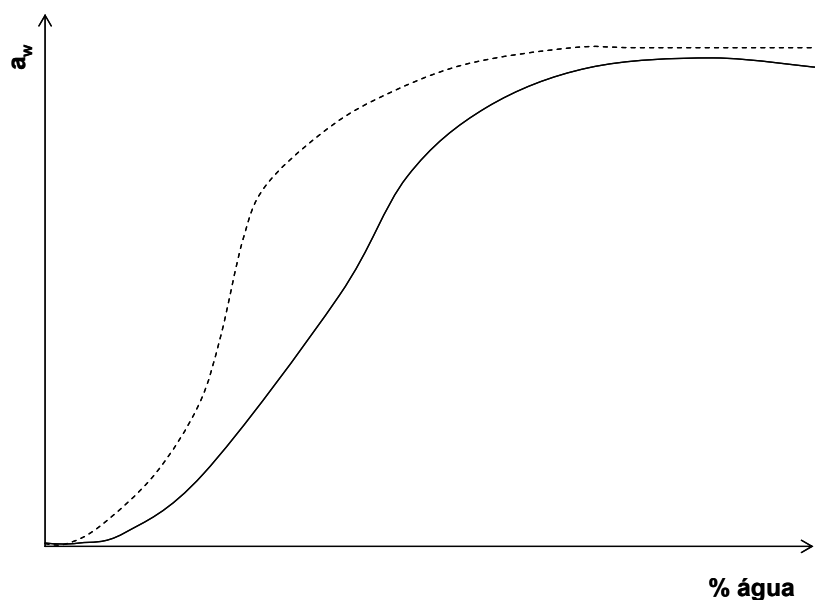
$$a_w = \frac{p_{\text{H}_2\text{O}}}{p_{\text{H}_2\text{O}}^*} \quad (1.50)$$

onde  $p_{\text{H}_2\text{O}}$  é a pressão parcial da água no produto e  $p_{\text{H}_2\text{O}}^*$  corresponde à pressão de vapor da água pura na mesma temperatura do produto – valores de pressão de vapor de água, ou mais convenientemente fugacidade da água a várias temperaturas, podem ser encontrados em diversos *handbooks* [49, 50].

---

1 - A atividade de água também é popularmente conhecida como “água livre”.

A Figura 1.12 mostra uma isoterma típica da atividade de água em função da concentração de água do produto. Nota-se que uma grande quantidade de água deve ser removida antes que a atividade de água do sistema seja reduzida significativamente. Entretanto, existem maneiras mais eficientes de reduzir a atividade de água sem necessariamente reduzir a quantidade de água no produto pela adição de compostos altamente higroscópicos, como sal, açúcar, óxido de titânio, amidos (nativos e modificados) e gomas [3-5, 7].



**Figura 1.12** – Isothermas típicas de absorção de água. A linha sólida representa um produto menos higroscópico que aquele representado pela linha tracejada.

Labuza *et al.* [3, 51] determinaram que a lei de velocidade para perda ou ganho de umidade em produtos, a uma dada temperatura, é dada por:

$$\left(\frac{dw}{dt}\right)_T = \frac{k}{x} A(\rho_{\text{amb}} - \rho_{\text{prod}}) \quad (1.51)$$

em que  $(dw/dt)_T$  é a taxa de perda ou ganho de umidade à uma temperatura  $T$  constante;  $\rho_{\text{amb}}$  e  $\rho_{\text{prod}}$  correspondem, respectivamente, às pressões de vapor do ambiente e do produto;  $A$  é a área da embalagem e  $k/x$  é a permeabilidade do filme à passagem de moléculas de água (WTR). Como  $x$  é a espessura do filme em unidades métricas, a constante de velocidade  $k$  é expressa nas seguintes unidades:

$$k = \frac{g_{\text{H}_2\text{O}} \cdot \text{espessura}}{\text{tempo} \cdot \text{área} \cdot \text{pressão}} \quad (1.52)$$

Como no caso da variação de temperatura, o estudo de validade de produtos sensíveis à quantidade de água também pode ser acelerado aumentando-se a umidade relativa de estocagem. Para tal, é necessário que se determinem:

1. A isoterma de perda ou ganho de água (Figura 1.14);
2. A permeabilidade da embalagem à água (WTR);
3. A relação entre a área da embalagem e o peso seco do produto ( $m_s$ ) em  $m^2/g$  sólidos;
4. Os valores inicial ( $m_i$ ) e crítico ( $m_c$ ) de umidade do produto;
5. A umidade relativa e a temperatura de estocagem do produto, com os quais a quantidade de água  $m_e$  que o produto teria se não tivesse embalagem pode ser calculada.

Labuza *et al.* determinaram que a isoterma de perda ou ganho de água pode ser aproximada por uma função linear se a quantidade de água no produto for relativamente baixa [3, 51, 53].

$$m_{wt} = \beta a_w + m_0 \quad (1.53)$$

em que  $m_{wt}$  é a quantidade de água por grama de sólidos;  $\beta$  é o coeficiente angular da isoterma de perda ou ganho de água na região de baixa umidade;  $a_w$  é a atividade de água e  $m_0$  é a quantidade de água por grama de sólidos quando  $a_w = 0$  [3, 51, 53].

Considerando a aproximação da Equação 1.53, a concentração de água no produto em qualquer tempo  $\theta$  será dada por:

$$\ln\left(\frac{m_e - m_i}{m_e - m}\right) = \frac{k.A.p_0}{x.m_s.\beta} \times \theta, \text{ para absorção de água} \quad (1.54)$$

$$-\ln\left(\frac{m - m_e}{m_i - m_e}\right) = \frac{k.A.p_0}{x.m_s.\beta} \times \theta, \text{ para perda de água} \quad (1.55)$$

As Equações 1.54 e 1.55 mostram que gráficos logarítmicos em função do tempo fornecerão linhas retas com coeficientes angulares iguais a  $(k.A.p_0)/(x.m_s.\beta)$ , através dos quais as constantes de velocidade  $k$  podem ser determinadas.

## 4.2 – Relações entre Diferentes Condições de Estocagem e Cálculo da Validade

Quaisquer que sejam as condições de aceleração, faz-se necessário que se estabeleçam relações entre as condições severas de estocagem e aquela na qual o produto será realmente comercializado. O parâmetro que estabelece a correlação entre as velocidades de degradação da propriedade  $P$  nas condições acelerada e de mercado é definido como [37, 35, 39]:

$$\alpha_{CA,CM} = \frac{v_{CA}}{v_{CM}} = \frac{k_{CA}P^n}{k_{CM}P^m} \quad (1.56)$$

onde  $\alpha_{CA,CM}$  é o *fator de aceleração*,  $v_{CA}$  e  $v_{CM}$  são, respectivamente, as velocidades de reação para as condições acelerada ( $CA$ ) e de mercado ( $CM$ ) e  $m$  e  $n$  são as ordens dos processos de degradação em cada condição de estocagem. Quando os processos de degradação nas diferentes condições de estocagem seguem cinética de mesma ordem,  $\alpha_{CA,CM}$  é constante e a Equação 1.56 pode ser re-escrita, de maneira simplificada, como:

$$\alpha_{CA,CM} = \frac{k_{CA}}{k_{CM}} \quad (1.57)$$

Labuza definiu um fator de aceleração denominado  $Q_{10}$  [3, 36, 39]:

$$Q_{10} = \frac{v_{T+10^{\circ}C}}{v_T} \quad (1.58)$$

que é um caso particular da Equação 1.57.

Para estudos de validade acelerados através de estocagem sob temperaturas elevadas, o fator de aceleração relaciona-se com a energia de ativação de acordo com a expressão:

$$\alpha_{CA,CM} = e^{-\frac{Ea}{R}\left(\frac{1}{T_{CA}} - \frac{1}{T_{CM}}\right)} \quad (1.59)$$

No caso do  $Q_{10}$ , a relação com a energia de ativação é dada por:

$$\log Q_{10} = \frac{2.19E_a}{T(T + 10)} \quad (1.60)$$

Ainda quanto à análise de dados para estudos acelerados de validade, existem publicações que apresentam a suposta regra-geral pela qual a velocidade das reações dobra a cada 10°C de incremento da temperatura de estocagem. Numa representação matemática, esta regra afirma que  $\alpha_{T,T+10^\circ\text{C}} = 2$ , para qualquer propriedade, qualquer que seja o produto. Esta regra é algumas vezes utilizada para definir o tempo necessário para condução dos estudos acelerados de validade e inclusive para tomada de decisão sobre o lançamento de produtos. Infelizmente esta relação provou-se inválida para uma grande gama de estudos de validade, segundo os quais  $\alpha_{T,T+10^\circ\text{C}}$  pode variar entre valores próximos de 1 até 40 - ou seja, as propriedades podem ser praticamente independentes ou extremamente sensíveis à temperatura [7, 36, 39, 47]! Aparentemente esta regra vem da observação empírica de que uma grande gama de reações químicas elementares em solução apresenta energia de ativação variando entre 30 e 50 kJ.mol<sup>-1</sup>, e apenas nesta estreita faixa de variação de energia de ativação a relação  $\alpha_{T,T+10^\circ\text{C}} \cong 2$  seria válida [17, 18].

Infelizmente não existem relações teóricas bem estabelecidas relacionando as constantes de velocidade de absorção de água em função da umidade relativa de estocagem, sendo estas dependentes das condições de estudo. Para uma grande faixa de concentração de água no produto, esta relação é linear (vide Figura 1.14), mas na maioria dos casos é determinada empiricamente.

Uma vez que os fatores de aceleração das condições acelerada e de mercado são conhecidos, o prazo de validade final do produto pode então ser calculado de acordo com a expressão [3-5, 7, 35-42]:

$$Val_{CM} = Val_{CA} \times \alpha_{CA,CM} \quad (1.61)$$

onde  $Val_{CM}$  e  $Val_{CA}$  correspondem, respectivamente, aos prazos de validade do produto nas condições de mercado e acelerada.

### 4.3 - Testes de Estabilidade de Produtos

Existem situações em que estudos de validade formais podem ser dispensados, como quando pequenas alterações são realizadas em produtos já consagrados no mercado. Podem-se citar como exemplos a substituição de compostos inertes na formulação, redução da quantidade de um ingrediente ou alterações em componentes presentes em quantidades muito reduzidas na formulação ou outras alterações de produto cujos impactos já são relativamente bem conhecidos. Nestes casos estudos mais curtos e objetivos, denominados *testes de estabilidade*, podem ser realizados.

Os testes de estabilidade diferem substancialmente dos estudos de validade no sentido em que têm por objetivo **checar** se pequenas alterações nos produtos (formulação, embalagem, processo) terão impacto significativo no prazo de validade estipulado. Usualmente um número limitado de parâmetros de qualidade é monitorado por um prazo definido, em condições-padrão de estocagem. Se nenhuma alteração significativa for detectada nestes parâmetros em um prazo limite previamente estipulado, conclui-se que o protótipo é estável, ou seja, possui a mesma validade do produto original, nas condições do teste. Deste modo, testes de estabilidade são *qualitativos* e *comparativos* por definição. Como uma grande quantidade de informação *a priori* deve estar disponível para que testes de estabilidade sejam executados com segurança, a realização de estudos de validade formais fazem-se necessários para estipular a relação entre as condições de estocagem. É importante salientar que, apesar de serem geralmente mais rápidos e baratos que os estudos de validade, sua natureza qualitativa faz com que os testes de estabilidade forneçam uma quantidade limitada de informações causais, tornando difícil a adoção de medidas para estender o prazo de validade do novo produto.

### 4.4 – Condução Prática de Estudos de Validade

Para que os objetivos dos testes de validade sejam atingidos de maneira eficiente é necessário não apenas que se conheçam as premissas teóricas da técnica, mas também as restrições logísticas práticas envolvidas em sua execução. Antes de iniciar o estudo os pesquisadores devem checar como a quantidade necessária de amostras será coletada, em que formato de embalagem, se os recursos laboratoriais estarão disponíveis nas datas previstas de amostragem, etc.

A Tabela 1.1 traz as principais etapas e atividades dos estudos práticos de validade de produtos e pode servir como guia na condução de estudos efetivos. A fase de planejamento é a mais importante em estudos de validade, pois decisões equivocadas na definição de quaisquer dos quesitos podem levar a conclusões errôneas ou a resultados inconclusivos.

**Tabela 1.1 – Principais etapas dos estudos de validade.**

Etapas	Atividades
Planejamento	Definição de: <ul style="list-style-type: none"> <li>– condições de estocagem</li> <li>– embalagens</li> <li>– quantidade de amostras</li> <li>– tempo total de estudo,</li> <li>– frequência de amostragem</li> <li>– parâmetros a monitorar</li> <li>– métodos de análise</li> </ul>
Estocagem	Monitoramento da temperatura, umidade relativa, e quaisquer outros parâmetros relevantes
Caracterização	Realização de análises químicas, físicas, sensoriais...
Análise de Dados	Construção de gráficos cinéticos Determinação das ordens das reações Cálculo das constantes de velocidade Cálculo dos fatores de aceleração Determinação da validade efetiva do produto
Conclusões e Ações	O prazo de validade atende às expectativas? É necessário aumentar o nível de preservação do produto? Como configurar futuros protocolos de estabilidade?

As propriedades a monitorar são geralmente definidas com base no conhecimento técnico do produto, incluindo seus ingredientes constituintes, condições de processamento, características das embalagens, condições de transporte e armazenamento, etc. Alguns produtos possuem regulamentações próprias que exigem o monitoramento de propriedades específicas, o teor de ativo em fármacos constituindo o caso mais óbvio. Além disso, certamente técnicas analíticas robustas e validadas, com precisão e exatidão adequadas, devem estar disponíveis para medições durante os estudos.



O tempo total de estudo é geralmente definido através de uma estimativa inicial da validade do produto. Tal estimativa leva em consideração o conhecimento técnico da formulação e processo, bem como o prazo de validade de produtos similares existentes no mercado.

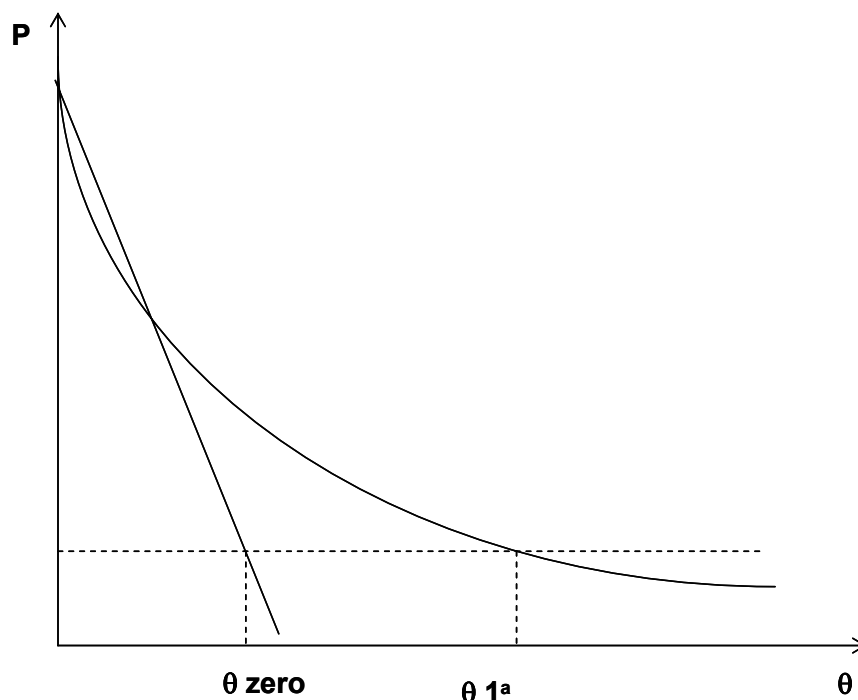
Quanto maior a ordem da reação mais rapidamente a propriedade  $P$  se degradará com o tempo. Deste modo, a correta estimativa da ordem da propriedade  $P$  é de fundamental importância em estudos de validade de produtos uma vez que é através dela que a frequência de análises será estipulada. Assim, caso se conheça *a priori* que uma propriedade de interesse segue cinética de ordem pseudo-zero as avaliações podem ser igualmente espaçadas no tempo. Por outro lado, as análises devem ser realizadas com maior frequência no início do estudo se for estimado que a propriedade segue cinética de pseudo-primeira ou pseudo-segunda ordem [3-5, 36, 39, 54, 55]. Caso a ordem da cinética de alteração da propriedade no tempo não seja conhecida, é usual que se realizem análises químicas, físicas ou sensoriais com maior frequência no início do estudo.

Constitui boa prática de experimentação em estudos de validade armazenar uma quantidade de amostras superior àquela que seria efetivamente necessária para realizar todas as avaliações analíticas nas frequências estipuladas. Esta precaução tem por objetivo garantir que haverá amostras suficientes para repetir quaisquer ensaios dúbios, para que novos parâmetros sejam analisados em caso de necessidade ou para prolongar o estudo caso os parâmetros de qualidade alvo não alcancem níveis considerados inaceitáveis durante o prazo inicialmente estipulado do estudo [3, 4, 36, 54, 55].

As condições de estocagem são usualmente definidas de modo a conseguir-se a resposta do prazo de validade do produto no menor tempo possível; em outras palavras, condições de estocagem que forneçam o maior fator de aceleração ( $\alpha$ ) possível são escolhidas. A decisão leva em consideração a sensibilidade do produto à temperatura e umidade, mas limitações logísticas, como disponibilidade de equipamentos e prazos de lançamento do produto, também são levadas em consideração. Entretanto, não é aconselhável que as condições de estocagem em um estudo acelerado de validade sejam drásticas a ponto de mudar a ordem das reações com relação à condição de mercado, pois condições extremas podem alterar o mecanismo das reações de degradação e levar ao desenvolvimento de características que efetivamente não ocorreriam em situações mais brandas.

A definição dos limites aceitáveis para cada parâmetro, ou *notas de corte*, constitui uma das mais difíceis etapas em estudos de validade. Apesar das notas de corte para alguns parâmetros serem definidos por legislação específica, como peso líquido declarado, teor de compostos ativos em fármacos, nutrientes em alimentos, etc., é normalmente função dos pesquisadores definir quando o produto atingiu estados inaceitáveis de qualidade. Para bens de consumo o desafio torna-se ainda maior, pois os consumidores podem rejeitar um produto baseados apenas na aparência, textura, sabor, odor ou uma combinação de vários. Para estes produtos, a realização de pesquisas de mercado, a fim de se determinar as características mínimas de qualidade do produto, são de extrema valia.

Em estudos em tempo real (não-acelerados), a validade efetiva do produto é definida com base na avaliação de gráficos cinéticos construídos utilizando-se dados de variação das propriedades que definem sua qualidade na temperatura de mercado. Em estudos acelerados, por outro lado, decisões prévias são tomadas com base nos valores das constantes de aceleração ( $\alpha$ ) e validadas *a posteriori*. Nestes casos é usual tomar-se a ordem da reação de degradação do produto na condição acelerada mais drástica, pois, como mostrado na Figura 1.13, a distinção entre cinéticas de, por exemplo, pseudo-zero e pseudo-primeira ordem torna-se difícil quando a propriedade em estudo ainda não sofreu alterações significativas, o que geralmente ocorre para bens de consumo de longa duração estocadas por curtos períodos de tempo à temperatura ambiente. Deste modo, mesmo que em uma dada condição acelerada seja observada cinética de pseudo-primeira ordem enquanto que para a condição de mercado seja observada uma cinética de ordem pseudo-zero, assume-se que a condição ambiente também segue cinética de pseudo-primeira ordem. Entretanto, esta conclusão deve ser validada através da extensão do estudo, ao menos para a condição de mercado, uma vez que a verificação de que esta premissa é falha pode levar a previsões significativamente errôneas quanto à validade real do produto.



**Figura 1.13** – Diferenças entre os prazos de validade estimados para cinéticas de pseudo-zero e pseudo-primeira ordem. Note que no início do estudo, quando poucos pontos estão disponíveis, é possível obter um bom ajuste linear aos dados para cinética de ordem pseudo-zero.

O estudo de validade deve prosseguir até que o produto atinja um nível inaceitável de qualidade em um ou mais parâmetros, em pelo menos uma condição acelerada, dependendo dos objetivos do teste. Só então extrapolações podem ser realizadas utilizando-se os fatores de aceleração determinados para cada condição de estocagem, e o prazo de validade efetivo assim determinado. É importante salientar, entretanto, que estudos de validade acelerados devem ser validados através da monitoração dos parâmetros de qualidade de amostras estocadas nas condições do mercado consumidor por todo o prazo de validade que será efetivamente praticado comercialmente.

## **5 – Conclusões**

A realização de estudos de validade é indispensável para o lançamento de novos bens de consumo no mercado. Os produtos devem manter suas características inalteradas no momento do uso pelos consumidores, cada vez mais exigentes por de qualidade e performance superior.

Apesar de demandarem quantidades significativas de recursos e tempo, estudos de validade são fontes inestimáveis de informação sobre o comportamento dos produtos no sentido em que fornecem modelos matemáticos preditivos capazes de levar a planos de ação para aumentar o nível de preservação do produto, quer seja através de sua formulação, embalagem ou forma de distribuição.

Entretanto, os resultados de estudos de validade não são absolutos no sentido em que são representativos apenas das condições dos testes. Deste modo, é prática comum na indústria declarar nas embalagens um prazo de validade menor que aquele tecnicamente determinado de modo a acomodar fontes de variação não incluídas nos estudos, tais como oscilações de temperatura e transporte.

## 6 – Bibliografia

1. Agência Nacional de Vigilância Sanitária (ANVISA), Resolução - RDC nº 259, de 20 de setembro de 2002: Regulamento Técnico para Rotulagem de Alimentos Embalados.
2. Agência Nacional de Vigilância Sanitária (ANVISA), Resolução – RDC nº 79 de 28 de agosto de 2000: Estabelece a Definição e Classificação de Produtos de Higiene Pessoal, Cosméticos e Perfumes e Outros com Abrangência neste Contexto.
3. Labuza TP, Shelf-Life Dating of Foods, Food & Nutrition Press, Westport, 1982.
4. Kilcast D, Subramanian P, The Stability and Shelf-Life of Food, CRC Press, New York, 2000.
5. Steele R (ed.), Understanding and Measuring the Shelf-Life of Food, CRC Press, Boca Raton, 2004, pp. 3-67, 128-140, 317-354.
6. Banker GS, Rhodes CT, Modern Pharmaceutics, 4<sup>th</sup> Ed., Marcel Dekker, New York, 2002, pp. 215-258.
7. Fennema OW, Food Chemistry, 3<sup>rd</sup> Ed., CRC Press, New York, 1996, pp. 1013-1032.
8. Robinson CJ, Malhotra MK, Defining the concept of supply chain quality management and its relevance to academic and industrial practice, *Int. J. Production Economics*, **2005**, 96, 315-337.
9. Tomat GM, Prices, product differentiation and quality measurement: a comparison between hedonic and matched model methods, *Res. Economics*, **2006**, 60, 54-68.
10. Fowler FJ, Survey Research Methods, 4<sup>th</sup> Ed., Sage Pub. Inc., London, 2008.
11. Deshpande R, Zaltman G, Factors affecting the use of market research information: a path analysis, *J. Marketing Res.*, **1982**, 19, 14-31.
12. McQuarrie DA, Simon JD, Physical Chemistry: a Molecular Approach, University Science Books, Sausalito, 1997, pp. 1047-1122.
13. Atkins PW, Physical Chemistry, 6<sup>th</sup> Ed., Oxford University Press, Oxford, 1998, pp. 760-843.
14. Laidler KJ, Chemical Kinetics, 3<sup>rd</sup> Ed., Benjamin Cummings, New York, 1997, pp. 30-65.
15. House JE, Principles of Chemical Kinetics, 2<sup>nd</sup> Ed., Academic Press, Burlington, 2007.
16. Houston PL, Chemical Kinetics and Reaction Dynamics, Dover Pub., New York, 2006.
17. Moelwyn-Huges EA, The Kinetics of Reaction in Solution, Clarendon Press, Oxford, 1933, pp. 1 – 169.
18. Connors KA, Chemical Kinetics: the Study of Reaction Rates in Solution, VCH, New York 1990, pp. 1-29, 59-77, 187-244.
19. Ramos AF, Miller JA, Klippenstein SJ, Truhlar DG, Modeling the kinetics of bimolecular reactions, *Chem. Rev.*, **2006**, 106, 4518-4584.

20. Wechsler JA, Kurtzst WH, Characterization of arithmetic deoxyribonucleic acid synthesis at restrictive temperature in a dnaE mutant of Escherichia coli K-12, *J. Bacteriol.*, **1979**, 138, 201-206.
21. Du ZL, Xue YY, Chu CW, BA-CA-CU-O compositions superconducting at up to 126K and methods for making the same, **WO1998/022409**.
22. Sabatie J, Choplin L, Paul F, Monsan P, The effect of synthesis temperature on the rheological properties of native dextran, *J. Biotech. Letters*, **2005**, 8, 425-430.
23. Rocco AM, de Paoli MA, Effect of the temperature on the electrochemical synthesis of polypyrrole dodecylsulphate, *J. Braz. Chem. Soc.*, **1993**, 4, 97-101.
24. Liang CH, Hwang CS, Effect of synthesis temperature on the characteristics of hydrous manganese oxide deposited using the hydrothermal electrochemical method, *Jpn. J. App. Phys.*, **2008**, 47, 1662-1666.
25. Arrhenius SA, Über die reaktionsgeschwindigkeit bei der inversion von rohrzucker durch säuren, *Phys. Chem.*, **1889**, 4, 226-248.
26. Tolman RC, Statistical Mechanics with Applications to Physics and Chemistry, Chemical Catalog Co., New York, 1927, pp. 260-270.
27. Truhlar DG, Interpretation of the activation energy, *J. Chem. Educ.*, **1978**, 55, 309-311.
28. Menzinger M, Wolfgang R, The meaning and use of the Arrhenius activation energy, *Angewandte Chemie Int.*, **2003**, 8, 438-444.
29. Laidler K, Unconventional uses of the Arrhenius Law, *J. Chem. Educ.*, **1972**, 343-344.
30. Freitas MA, Costa JC, Shelf life determination using sensory evaluation scores: a general Weibull modeling approach, *Computers Industrial Eng.*, **2006**, 51, 652-670
31. Hough G, Workshop summary: Sensory shelf-life testing: Introduction, *Food Quality and Preference*, **2006**, 17, 640-645.
32. Farner EH, Bloomfield GF, Sudralingam A, Sutton DA, The course and mechanism of autoxidation reactions in olefinic and polyolefinic substances, including rubber, *Trans. Faraday Soc.*, **1942**, 38, 348-356.
33. Corbo, MR, del Nobile MA, Sinigaglia M, A novel approach for calculating shelf life of minimally processed vegetables, *Int. J. Food Microbiol.*, **2006**, 106, 69-73
34. Zwietering MH, Jongenburger FM, Roubouts M, van't Riet K., Modelling of the bacterial growth curve. *App. Environmental Microbiol*, **1990**, 56, 1875-1881
35. Magari RT, Murphy KP, Fernandez T, Accelerated stability model for predicting shelf-life, *J. Clin. Lab. Anal.*, **2002**, 16, 221-226.
36. Labuza TP, Schmidl MK, Accelerated shelf-life testing of foods, *Food Technol.*, **1985**, 39, 57-64.
37. Watermann KC, Adami RC, Accelerated aging: prediction of chemical stability of pharmaceuticals, *Int. J. Pharmaceuticals*, **2005**, 293, 101-125.
38. Taub IA, Sing RP (eds.), Food Storage Stability, CRC Press, New York, 1997, pp. 245-387.

39. Corradini MG, Peleg M, Shelf-life estimation from accelerated storage data, *Trends Food Sci. Technol.*, **2007**, 18, 37-47.
40. Reid DS, Kotte K, Kilmartin P, Young M, A new method for accelerated shelf-life prediction for frozen foods, *J. Sci. Food Agric*, **2003**, 83, 1018–1021.
41. van Dyck SMO, Verleyen T, Dooghe W, Teunckens AE, Adams C, Free radical generation assays: new methodology for accelerated oxidation studies at low temperature in complex food matrices, *J. Agric. Food Chem.*, **2005**, 53, 887-892.
42. Siripatrawan U, Jantawat P, A novel method for shelf life prediction of a packaged moisture sensitive snack using multilayer perceptron neural network, *Expert Sys. Applications*, **2008**, 34, 1562–1567
43. Riva M, Fessas D, Schiraldi A, Isothermal calorimetry approach to evaluate shelf life of foods, *Thermochimica Acta*, **2001**, 370, 73-81.
44. Waterman KC, Carella AJ, Gumkowski MJ, Lukulay P, MacDonald BC, Improved protocol and data analysis for accelerated shelf-life estimation of solid dosage forms, *Pharmaceutical Res.*, **2007**, 24, 780-790.
45. McMeekin TA, Ross T, Shelf life prediction: status and future possibilities, *Food Microbiol.*, **1996**, 33, 65-83.
46. USDA, Conference on Food Quality, USDA Agric. Res. Service, Albany, 1960.
47. Pedro AMK, Ferreira MMC, Multivariate accelerated shelf-life testing: a novel approach for determining the shelf-life of foods, *J. Chemometrics*, **2006**, 20, 76-83.
48. Connors KA, Amidon GL, Stella VJ, Chemical Stability of Pharmaceuticals: a handbook for pharmacists, 2<sup>nd</sup>. Ed., Wiley Interscience, New York, 1996, pp. 3-62, 82-150.
49. Lide DR (ed.), CRC Handbook of Chemistry and Physics, 89<sup>th</sup> Ed., CRC Press, London, 2008.
50. Speight J (ed.), Lange's Handbook of Chemistry, 16<sup>th</sup> Ed., McGraw Hill, New York, 2006.
51. Labuza TP, Mizrahi S, Karel M, Mathematical models for the optimization of flexible film packaging of foods storage, *Trans. ASAE*, **1972**, 15, 150-169.
52. Solovyov SE, Goldman AY, Mass Transport and Reactive Barriers in Packaging: Theory, Application and Design, DEStech Pub. Inc, New York, 2007, pp. 325-382.
53. Mizrahi S, Karel M, Accelerated stability tests of moisture sensitive products in permeable packages by programming rate of moisture content increase, *J. Food Sci.*, **1977**, 42, 958-972.
54. Chow SC, Statistical Design and Analysis of Stability Studies, Chapman & Hall, Boca Raton, 2007.
55. Mazzo DJ (ed.), International Stability Testing, Interpharm Press, Buffalo Grove, 1999.

## **CAPÍTULO II**

### **O MÉTODO MULTIVARIADO ACELERADO PARA DETERMINAÇÃO DO PRAZO DE VALIDADE DE PRODUTOS (*MULTIVARIATE ACCELERATED SHELF-LIFE TEST - MASLT*)**



## 1 – Introdução

A robustez da teoria para determinação do prazo de validade de produtos foi extensamente demonstrada na literatura através de sua aplicação prática a uma grande gama de bens de consumo. Todavia, conforme apresentado no Capítulo I deste trabalho, o Teste Acelerado para Determinação do Prazo de Validade (ASLT – *Accelerated Shelf-Life Test*) foi concebido de modo a estudar os mecanismos de degradação dos produtos de maneira isolada; isto é, cada propriedade deve ter seus parâmetros cinéticos determinados individualmente. Esta característica traz desvantagens, em especial em estudos de bens de consumo onde a monitoração de várias características de qualidade faz-se necessária. Nestes casos, importantes correlações entre diferentes propriedades podem deixar de ser detectadas pelos pesquisadores. Correlações entre propriedades de diferentes naturezas – química, sensorial e física – são de extrema relevância na indústria de bens de consumo, pois têm potencial para revelar formas de controlar a degradação do produto (por exemplo, pode-se alterar o processo de fabricação para impedir a redução no teor de um nutriente) e/ou orientar o desenvolvimento de novos produtos de modo que tenham a máxima aceitação pelos consumidores.

A natureza univariada da teoria também torna impraticável a utilização direta de dados oriundos de técnicas analíticas instrumentais para a determinação do prazo de validade de produtos complexos. Técnicas como espectroscopia, cromatografia e calorimetria são capazes de revelar não apenas o comportamento das propriedades-chave do produto, mas também possíveis processos não esperados para uma determinada classe de produto. Deste modo, a utilização de técnicas quimiométricas multivariadas, em conjunto com a teoria cinética, tem grande potencial de aplicação para determinação da validade de produtos.

De fato, devido a estas limitações dos métodos convencionais, técnicas quimiométricas já foram empregadas no estudo de validade de produtos. Gomez e colaboradores utilizaram PCA (*Partial Least Squares*) e LDA (*Linear Discriminant Analysis*), em conjunto com narizes eletrônicos, no estudo de tomates e laranjas asiáticas [1, 2]. Estratégia similar foi utilizada por Zhu e colaboradores para estudar a estabilidade de aromas em formulações de fármacos [3]. Casale *et al.* utilizaram espectroscopia no infravermelho próximo (NIR) para estudar a degradação de amostras de vinagre [4]. Giardina, Cattaneo *et al.* utilizaram a mesma técnica para estudar a degradação de leite e queijos [5, 6] e Boknaes *et al.* para

estimar a estabilidade de produtos de peixes [7]. Wold *et al.* utilizaram dados sensoriais e análise discriminante para estudar fotoxidação em queijos [8]. Bodechi e colaboradores utilizaram PCA e espectroscopia FTIR (*Fourier Transform Infrared*) para estudar a degradação de utensílios médicos à base de PVC [9]. Chou *et al.* desenvolveu técnica estatística para a determinação da validade de fármacos utilizando propriedades discretas [10]. Existe ainda uma miríade de publicações relatando o uso de propriedades determinadas com o auxílio de quimiometria para estimação da validade de produtos, principalmente *in natura* [11-17].

Outros autores desenvolveram técnicas baseadas em estatística multivariada para modelar a cinética de reações químicas. Carvalho, Sanches, Watoon e Brereton desenvolveram estratégia multivariada baseada em PLS e compararam *soft* e *hard modelling* para determinação de parâmetros cinéticos de uma reação química [18]. Tsong e colaboradores desenvolveram técnicas para lidar com várias propriedades utilizando Análise de Covariância (ANCOVA) [19]. Widjaja utilizou *Multivariate Curve Resolution* (MCR) para determinar a estabilidade de fármacos utilizando microscopia no infravermelho próximo [20]. Dyrstad e colaboradores desenvolveram estratégia para amostragem e validação do prazo de validade em *batches* de produtos farmacêuticos [21]. Entretanto, nenhum trabalho anterior uniu a quimiometria (ou *soft modelling*) com técnicas cinéticas formais para a determinação direta dos modos de degradação e do prazo de validade de produtos (*hard modelling*), conforme proposto neste trabalho.

O *Multivariate Accelerated Shelf-Life Test* (MASLT) - cujas premissas, características e propriedades serão apresentadas neste Capítulo - reúne as vantagens dos métodos quimiométricos de análise de dados multivariados com a capacidade de modelagem e previsão dos métodos cinéticos e da teoria clássica para determinação do prazo de validade de produtos, descritos no Capítulo I deste trabalho. O método se propõe a auxiliar na compreensão dos mecanismos de degradação de produtos complexos, bem como nas possíveis relações entre diversas propriedades que definem a qualidade dos bens de consumo. Além disso, permite que métodos instrumentais sejam utilizados de maneira direta na determinação do prazo de validade para compreensão do comportamento de degradação dos produtos, ao invés dos estudos anteriormente reportados onde as técnicas instrumentais foram utilizadas de maneira indireta para a determinação de propriedades de interesse dos bens de consumo.

## 2 - Objetivos

O objetivo deste capítulo é apresentar o *Multivariate Accelerated Shelf-Life Test* como uma alternativa aos métodos convencionais univariados para determinação do prazo de validade de produtos, bem como suas premissas, vantagens e limitações.

## 3 – Análise de Componentes Principais

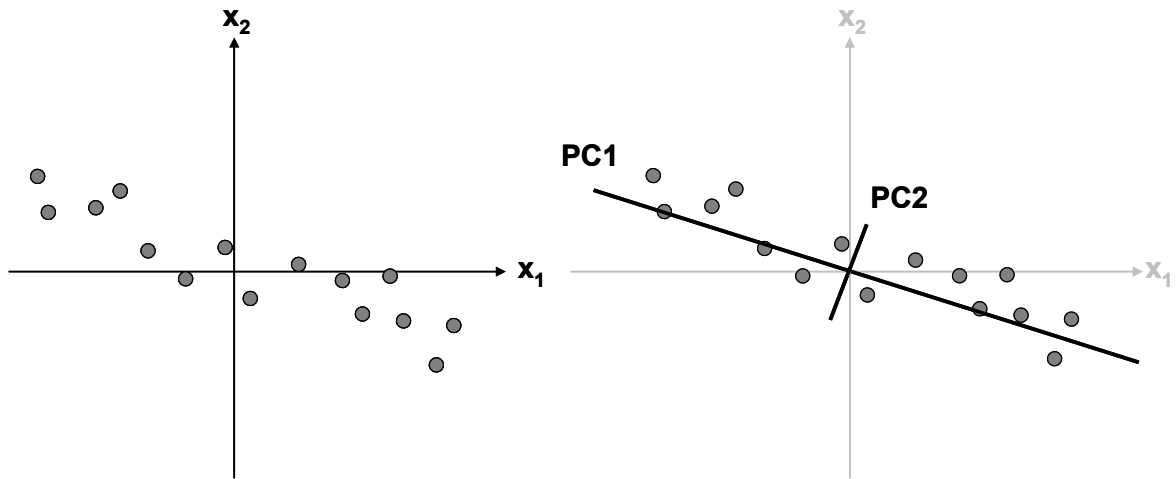
Um dos métodos quimiométricos mais empregados na análise multivariada de dados é a Análise de Componentes Principais (PCA – *Principal Component Analysis*). Seus conceitos foram introduzidos por Karl Pearson, em 1901 [22], mas as primeiras aplicações em química surgiram no final da década de 1960 quando Kowalski e Isenhour publicaram uma série de artigos na revista *Analytical Chemistry* [23-25]. Atualmente é largamente empregada em análise exploratória de dados e na construção de modelos de qualificação ou quantificação [26-29].

A Análise de Componentes Principais também pode ser utilizada como uma forma de revelar a estrutura interna do conjunto de dados de maneira que sua variância intrínseca seja mais bem explicada e visualizada. Quando as Componentes Principais (PCs) revelam algum comportamento físico ou químico do conjunto de dados, são também denominadas como variáveis latentes [26, 27, 29].

Geometricamente, a PCA é definida como uma transformação linear ortogonal que cria um novo sistema de coordenadas onde os dados originais serão projetados. Cada eixo do novo sistema de coordenadas é construído através de combinações lineares das propriedades originais e é denominado Componente Principal (PC – *Principal Component*). Como o novo sistema que descreve os dados em geral tem menor dimensionalidade (isto é, um número menor de eixos) que o sistema original, a PCA constitui ferramenta largamente empregada para compressão de dados [26-29].

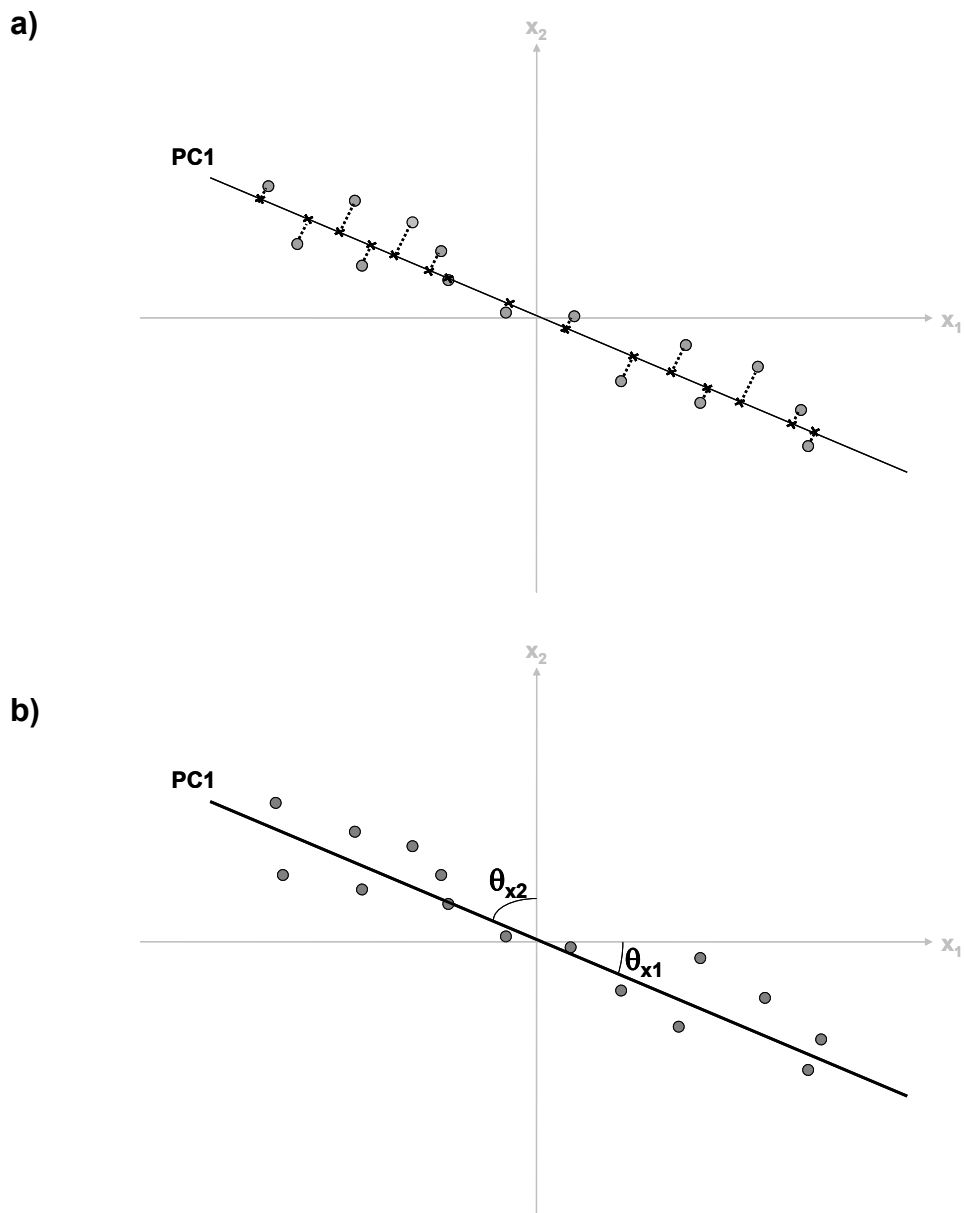
A Figura 2.1 traz um conjunto de dados simples contendo duas variáveis originais,  $x_1$  e  $x_2$ . Por definição, a primeira componente principal é traçada seguindo o eixo capaz de descrever a maior quantidade possível da variância do conjunto de dados. Em seguida, traça-se o eixo definido pela segunda componente principal que, neste exemplo, descreve o restante da variância do conjunto de dados original. Observa-se que todas as PCs devem ser linearmente independentes, ou seja, ortogonais entre si. Além disso, o número

máximo de componentes principais deve ser igual ao número de objetos (linhas) ou variáveis (colunas) no conjunto de dados original, o que for menor [26-30].



**Figura 2.1** – Representação geométrica da aplicação da PCA para um sistema com 2 variáveis. Neste exemplo, 95% da variância no conjunto de dados original pode ser descrita por PC1.

Os valores escalares das projeções ortogonais dos pontos originais sobre cada componente principal (cruzes sobre o eixo definido por PC1 na Figura 2.2a) são denominados *scores*, enquanto que os *loadings* representam as projeções dos vetores que definem cada PC nos eixos originais do sistema de coordenadas ( $\cos\theta_i$ , Figura 2.2b). A PCA é também largamente utilizada como método de análise exploratória de dados, pois Gráficos de *loadings* e *scores* revelam padrões característicos do comportamento das amostras em função de um conjunto de variáveis que dificilmente seriam reconhecidos pela observação de valores tabelados [26-30].



**Figura 2.2** – Representação geométrica dos *scores* e *loadings* para PC1.

No exemplo da Figura 2.1 acima, o padrão de distribuição dos dados sugere que, com exceção de erro experimental, ambas as variáveis originais descrevem um único fenômeno, que pode ser representado unicamente por PC1. Portanto, pode-se também descrever a PCA como uma técnica capaz de separar informação química relevante do ruído experimental [26-30]<sup>1</sup>.

1 – cabe salientar que esta separação não é perfeita. De fato, Faber *et al.* [31, 32] argumenta que todas as PCs trarão contribuições do erro experimental em sua construção.

Para que interpretações e conclusões corretas sejam obtidas através da análise, é importante compreender que a PCA baseia-se em duas premissas principais:

1. linearidade: assume-se que o conjunto de dados original é uma combinação linear de uma dada base de vetores. Existem outros métodos não-lineares, como o *kernel PCA*, que foram desenvolvidos sem assumir linearidade [26-30, 33-35].
2. relevância da média e da covariância: a redução da dimensionalidade do conjunto de dados na PCA é obtida pela retenção das características que contribuem significativamente para a variância, o que é conseguido ignorando-se as componentes de ordens mais altas e retendo as de ordens mais baixas. Assume-se que estas componentes de ordens baixas geralmente contém a maior parte da informação relevante do conjunto de dados. Entretanto, dependendo da aplicação esta premissa pode não ser verdadeira e, nestes casos, pode-se lançar mão de seleção de variáveis específicas do conjunto de dados original ou realizar análise de componentes principais ponderada (*wPCA*, *Weighted Principal Component Analysis*) [26-30, 36].

### 3.1 - O Modelo Matemático da PCA

Seguindo a convenção em álgebra linear, neste trabalho arranjos serão representados por letras maiúsculas sublinhadas em negrito, matrizes por maiúsculas em negrito, vetores por letras minúsculas em negrito e escalares por letras minúsculas em itálico. As dimensões de arranjos, matrizes e vetores serão representados por letras maiúsculas em itálico.

Para evitar que as primeiras PCs descrevam a distância entre o centro do conjunto de dados e a origem do sistema de coordenadas, é usual executar pré-processamento por centralização na média (equação 2.1) na matriz de dados originais de modo que o centro do conjunto de dados coincida com a origem do sistema de coordenadas descrito pelas PCs. Dada uma matriz  $\mathbf{X} \in \mathfrak{R}^{N \times K}$ , em que  $N$  representa o número de objetos e  $K$  o número de variáveis [27-29, 37-40]:

$$x_{c,nk} = x_{nk} - \bar{x}_k \quad (2.1)$$

onde  $x_{c,nk}$  é um elemento da matriz  $\mathbf{X}_c$  cujas colunas foram centradas na média,  $x_{nk}$  é um elemento da matriz  $\mathbf{X}$  de dados originais e  $\bar{x}_k$  é a média dos valores da  $k$ -ésima coluna de  $\mathbf{X}$ .

Outro procedimento de pré-processamento comumente empregado antes da realização da PCA é o auto-escalamento. Este procedimento é principalmente relevante quando a matriz  $\mathbf{X}$  é composta por variáveis com dimensões diferentes. Nesses casos, se o procedimento de escalamento não for realizado, a PCA será orientada não pela variância intrínseca do conjunto de dados, mas pelas diferenças nas dimensões das variáveis – ou seja, variáveis com valores numéricos mais elevados dominarão a construção das PCs. Para evitar que isso ocorra cada coluna de  $\mathbf{X}_c$  é dividida por seu desvio-padrão [27-29, 36-40]:

$$x_{a,nk} = \frac{x_{c,nk}}{s(x_k)} = \frac{x_{nk} - \bar{x}_k}{s(x_k)} \quad (2.2)$$

onde  $x_{a,nk}$  é um elemento da matriz  $\mathbf{X}_a$  de dados auto-escalados e  $s(x_k)$  é o desvio-padrão dos elementos da  $k$ -ésima coluna de  $\mathbf{X}$ .

O método PCA decompõe a matriz pré-processada  $\mathbf{X}^1$  em duas matrizes distintas,  $\mathbf{T}$  e  $\mathbf{L}$  [26-32, 36-40]:

$$\mathbf{X} = \mathbf{TL}^T \quad (2.3)$$

em que  $\mathbf{T}$  é a matriz de *scores* e  $\mathbf{L}$  é a matriz de *loadings*.

A matriz de *scores*,  $\mathbf{T}$ , traz a projeção dos pontos definidos pelas linhas de  $\mathbf{X}$  no novo espaço definido pelos eixos da componente principal enquanto que a matriz de *loadings*,  $\mathbf{L}$ , descreve os pesos de cada variável original para a construção das componentes principais. Geometricamente, descrevem os co-senos das projeções dos vetores definidos pelas variáveis originais em cada eixo definido pelo novo espaço construído pela PCA (vide Figura 2.2 acima). Cabe salientar que uma variável pode possuir um peso considerável para uma dada PC, mas desprezível para outra [26-30, 36-38].

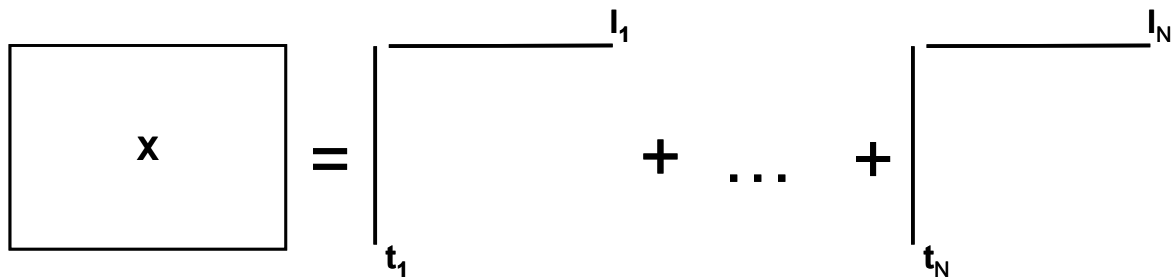
---

1 – a matriz  $\mathbf{X}$  pode ser pré-processada ou não. Para maior clareza, os sub-índices que designam pré-processamentos foram negligenciados.

Uma componente principal é então definida pelo produto vetorial de uma dada coluna de  $\mathbf{T}$  pela coluna de  $\mathbf{L}$  correspondente (Figura 2.3):

$$\mathbf{X} = \sum_{i=1}^{\min(N,K)} \mathbf{t}_i \mathbf{l}_i^T \quad (2.4)$$

em que  $\mathbf{t}$  e  $\mathbf{l}$  são colunas das matrizes de *scores* e *loadings*<sup>1</sup>.



**Figura 2.3** – Representação da PCA como a somatória dos produtos vetoriais de cada PC.

Uma característica importante das matrizes  $\mathbf{T}$  e  $\mathbf{L}$  é que os vetores definidos por suas colunas são ortogonais<sup>1</sup> entre si e, deste modo, não existe colinearidade entre as componentes principais. Além disso, suas colunas são ordenadas em ordem decrescente de quantidade de variância explicada por cada componente principal. Assim, para reduzir a dimensionalidade do sistema, o número de componentes principais ( $R$ ) a manter no modelo pode ser selecionado pela comparação entre estas variâncias e o erro experimental típico do conjunto de dados (Figura 2.4). Componentes principais que descrevem pouca variância geralmente trazem apenas ruído ou informação não-relevante sobre o sistema e podem ser desprezadas. Neste caso, as expressões 2.3 e 2.4 podem ser re-escritas como [26-30, 36-38]:

1 – os vetores da matriz de *scores*  $\mathbf{T}$  são ortogonais no sentido geométrico, isto é, o co-seno entre os vetores definidos por suas colunas é igual a zero. Algebricamente a matriz de *scores* não é ortogonal, uma vez que  $\mathbf{T}^T \mathbf{T} = \mathbf{D}$ , onde  $\mathbf{D}$  é uma matriz diagonal diferente da matriz identidade,  $\mathbf{I}$ . Maiores detalhes são fornecidos na seção 3.1 a seguir.

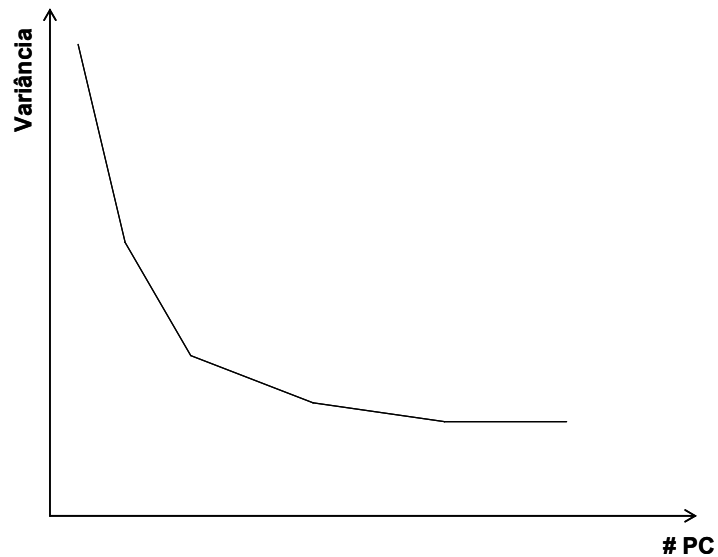


$$\hat{\mathbf{X}} = \mathbf{T}_R \mathbf{L}_R^T \quad (2.5)$$

$$\mathbf{X} = \mathbf{T}_R \mathbf{L}_R^T + \mathbf{E} \quad (2.6)$$

$$\mathbf{X} = \sum_{i=1}^R \mathbf{t}_i \mathbf{l}_i^T + \mathbf{E} \quad (2.7)$$

onde  $\hat{\mathbf{X}}$  é a matriz de dados estimada pelas matrizes  $\mathbf{T}_R$  e  $\mathbf{L}_R$ , que por sua vez representam as matrizes de *scores* e *loadings* truncadas na  $R$ -ésima coluna.



**Figura 2.4** – Gráfico *scree* ou de variância, comumente utilizado para selecionar o número de PCs a manter na PCA.

Deste modo, a PCA constitui ferramenta extensivamente utilizada em análise multivariada de dados, pois com uma única decomposição matricial consegue-se reduzir o número de variáveis (normalmente,  $R \ll K$ ), eliminar a colinearidade em  $\mathbf{X}$  e reduzir o ruído experimental incluso no modelo [26-30, 36-38].

É importante salientar que as soluções da PCA não são únicas, pois os *loadings* e *scores* são sujeitos à ambigüidade rotacional: se os vetores da matriz de *scores* forem girados através de uma matriz de rotação  $\mathbf{Z}$  (quadrada e não-singular), a PCA fornecerá exatamente o mesmo ajuste se a matriz de *loadings* for girada no sentido contrário. Por exemplo, o ajuste do modelo  $\mathbf{X} = \mathbf{T}\mathbf{L}^T$  será idêntico ao fornecido por  $\mathbf{X} = \mathbf{T}\mathbf{Z}^{-1}\mathbf{Z}\mathbf{L}^T = \tilde{\mathbf{T}}\tilde{\mathbf{L}}^T$ . Esta ambigüidade traz complicações em alguns casos, como na análise estrutural e interpretação de *loadings* espectrais, mas considerações teóricas e práticas podem ser utilizadas para definir matrizes de rotação apropriadas, fornecendo assim uma solução PCA única [26-30, 36,37].

### 3.2 – Algoritmos para Realização da PCA

Matrizes podem ser decompostas como o produto de duas ou mais matrizes diferentes, mas que possuem propriedades interessantes. Esta estratégia é utilizada em uma série de aplicações em álgebra linear, incluindo a resolução do problema dos quadrados mínimos [41, 42].

A decomposição em valores singulares é aplicada em vários algoritmos quimiométricos. Esta decomposição transforma uma matriz  $\mathbf{X} \in \mathfrak{R}^{N \times K}$  no produto de três matrizes,  $\mathbf{U}$ ,  $\mathbf{S}$  e  $\mathbf{V}$ :

$$\mathbf{X} = \mathbf{USV}^T \quad (2.8)$$

em que  $\mathbf{U} \in \mathfrak{R}^{N \times N}$ ,  $\mathbf{S} \in \mathfrak{R}^{N \times K}$  e  $\mathbf{V} \in \mathfrak{R}^{K \times K}$ .

As matrizes  $\mathbf{U}$  e  $\mathbf{V}$  são ortogonais. Suas colunas trazem, respectivamente, os autovetores dos produtos cruzados  $\mathbf{XX}^T$  e  $\mathbf{X}^T\mathbf{X}$ , que são proporcionais às matrizes de covariância dos objetos e das variáveis, respectivamente.  $\mathbf{S}$  é diagonal e traz os valores singulares destas matrizes, ordenados em ordem decrescente, sendo que os valores singulares relacionam-se com os autovalores ( $\lambda$ ) da matriz  $\mathbf{X}$  [43-47]:

$$s_k = \sqrt{\lambda_k} \quad (2.9)$$

Um excelente algoritmo para realização da decomposição SVD é fornecido por Golub e Van Loan [42]. Entretanto, a maioria dos pacotes em álgebra linear já possui rotinas otimizadas para execução desta decomposição [43, 44].

As matrizes  $\mathbf{T}$  de *scores* e  $\mathbf{L}$  de *loadings* são obtidas, a partir da decomposição SVD, da seguinte forma:

$$\mathbf{T} = \mathbf{US} \quad (2.10)$$

$$\mathbf{L} = \mathbf{V} \quad (2.11)$$

E, de acordo com as equações 2.9 e 2.10:

$$\lambda_k = \sum_{n=1}^N t_{nk}^2 = \mathbf{t}_k^T \mathbf{t}_k \quad (2.12)$$

em que  $\lambda_k$  é o autovalor da  $k$ -ésima coluna de  $\mathbf{T}$ .

Outro método comumente utilizado para obtenção de *loadings* e *scores* é o algoritmo NIPALS desenvolvido por Hermann Wold em 1966 [48]. Trata-se de um procedimento iterativo que extrai uma componente principal por vez. O algoritmo segue as seguintes etapas:

1. Toma-se como estimativa inicial dos primeiros *scores*,  $\mathbf{t}_1$ , a coluna de  $\mathbf{X}$  com a maior somatória quadrática;
2. os *loadings* ( $\mathbf{l}_1$ ) são estimados, por quadrados mínimos, para estes *scores*;
3. Normalizam-se os *loadings* dividindo  $\mathbf{l}_1$  por  $\|\mathbf{l}_1\|$ ;
4. Novos *scores*  $\mathbf{t}_1^n$  são então calculados a partir de  $\mathbf{l}_1$ :  $\mathbf{t}_1^n = \mathbf{X}\mathbf{l}_1^T$ ;
5. As etapas 2 a 4 são repetidas até que haja convergência dos *scores*;
6. Calcula-se a matriz  $\mathbf{X}_1$ , recomposta pela primeira componente principal:  $\mathbf{X}_1 = \mathbf{t}_1\mathbf{l}_1^T$ ;
7. Calcula-se uma matriz de resíduos:  $\mathbf{E}_1 = \mathbf{X} - \mathbf{X}_1$ ;
8. Repete-se o processo para extração de todas as  $R$  componentes principais, substituindo  $\mathbf{X}$  por  $\mathbf{E}_n$ , para  $1 \leq n \leq R$ .

A principal vantagem deste algoritmo é que torna possível determinar *a priori* o número de componentes principais ( $R$ ) a extrair de  $\mathbf{X}$ , economizando recursos computacionais. Apesar dos computadores modernos possuírem capacidade computacional considerável, esta característica ainda é vantajosa quando o número de variáveis ou comprimentos de onda da matriz  $\mathbf{X}$  é muito grande [36, 38, 48].

## 4 – Multivariate Accelerated Shelf-Life Test (MASLT)

### 4.1 – Propriedades e Premissas

O MASLT baseia-se na compressão do espaço definido pelas variáveis originais através de uma PCA, de onde as informações relativas aos padrões de degradação do produto no tempo serão extraídas. Os *scores* de cada componente principal são utilizados como propriedades  $P$  (vide equação 1.43, capítulo I) no cálculo dos parâmetros cinéticos, enquanto que os *loadings* são utilizados para interpretar quais variáveis tiveram influência significativa na degradação do produto (Figura 2.5).

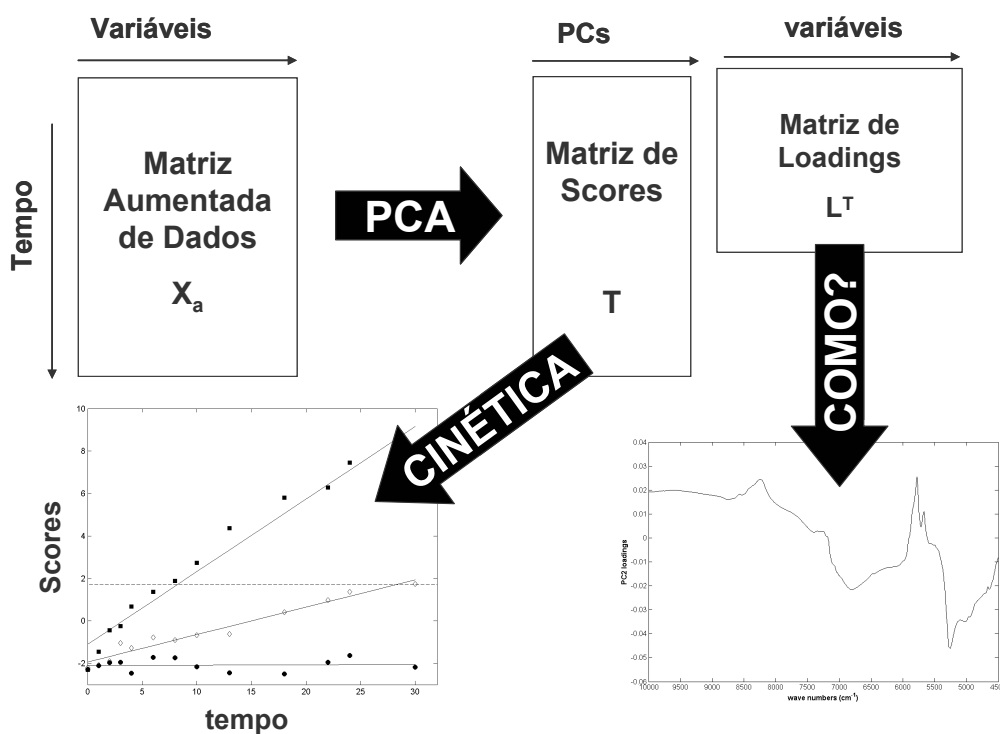


Figure 2.5 – Representação pictórica do MASLT.

A principal premissa do MASLT é que as reações de degradação do produto constituem a principal fonte de variação do conjunto de dados: de fato, uma vez que as amostras sejam idênticas no momento da armazenagem e que as condições de estocagem sejam controladas adequadamente, a PCA será orientada por fenômenos relacionados ao tempo.

Outra característica importante do MASLT vem da natureza da PCA. Como as componentes principais são extraídas em ordem decrescente de variância explicada,

espera-se que as primeiras PCs tragam informação relativa à degradação do produto, enquanto que as PCs subseqüentes devem explicar ruído ou processos não relativos ao tempo. Assim, a natureza multivariada do MASLT fornece ferramentas poderosas para diagnosticar possíveis falhas na execução do estudo de validade: os procedimentos de estocagem, amostragem e aquisição de dados devem ser cuidadosamente revistos caso se determine que as primeiras PCs não trazem informação relativa ao tempo.

Outra premissa do MASLT é que a PCA é capaz de acomodar processos não-lineares nos *scores* das PCs relacionadas com o tempo [26-30, 38]. Esta característica é especialmente importante para estudos de validade onde reações de ordens pseudo-primeira ou superiores são dominantes. A forma como estas não-linearidades estão sendo acomodadas pela análise de componentes principais pode ser determinada através da análise dos gráficos de *scores vs. tempo*.

Além disso, como a PCA é teoricamente o esquema de transformação linear ótimo, em termos de quadrados mínimos, para comprimir um conjunto de vetores de muitas dimensões em um conjunto de vetores com menor dimensionalidade [26, 38], os parâmetros de validade multivariados podem ser interpretados como médias ponderadas – pelos *loadings* - daqueles obtidos pelo cálculo de cada propriedade individualmente.

É importante salientar que amostras apresentando valores iniciais ( $\theta = 0$ ) elevados (em módulo) de *scores* não apresentarão necessariamente variação temporal significativa: os mesmos valores de *scores* podem ser determinados para estas amostras durante todo o estudo de validade, o que é bastante provável para amostras estocadas a baixas temperaturas, onde as reações de degradação estariam virtualmente paradas.

Pode ser necessário executar pré-processamentos no conjunto de dados antes da execução do MASLT, especialmente quando as medidas das variáveis originais apresentam-se em escalas diferentes. Alguns pré-processamentos, especialmente aqueles envolvendo re-escalamento, podem trazer alguns efeitos negativos ao estudo multivariado no sentido em que propriedades que não apresentaram variação significativa no tempo podem ser enfatizadas, enquanto que aquelas que apresentaram uma grande variação podem ser minimizadas. Como estas características são intrínsecas de quaisquer procedimentos de escalamento [39, 40], devem-se adotar medidas rigorosas de controle das condições de estocagem, bem como seguir criteriosamente os procedimentos analíticos no momento da aquisição de dados, a fim de evitar estes efeitos negativos.

Os dados devem ser organizados de maneira apropriada para que as informações relativas ao tempo e às condições de estocagem sejam mantidas na PCA. Esta forma de organização será descrita em maiores detalhes na seção 4.2 deste capítulo, mas cabe salientar que é necessária para que a PCA conserve as informações relativas ao tempo e às condições de estocagem no mesmo espaço multi-dimensional, condição especialmente relevante quando a aplicação de pré-processamentos faz-se necessária. Uma alternativa seria realizar análises de componentes principais individuais para cada condição de estocagem. Entretanto, este procedimento projetaria os dados em espaços diferentes uma vez que se espera que condições de estocagem distintas possuam variações intrínsecas características diferentes entre si – condições aceleradas apresentam maior variância que condições de mercado ou refrigeradas, por exemplo.

A principal desvantagem da execução da PCA nos dados arranjados conforme proposto no algoritmo do MASLT é que as relações entre as condições de estocagem são perdidas. O MASLT contorna esta desvantagem ao aliar a teoria cinética à análise de componentes principais: estas relações são recuperadas através da análise das constantes de velocidade e das energias de ativação.

## 4.2 – Algoritmo do MASLT

O MASLT é executado de acordo com as seguintes etapas:

1. Colete os dados em um arranjo  $\underline{\mathbf{X}} \in \Re^{N \times K \times C}$ , onde  $N$  é o número de amostragens no tempo por condição de estocagem,  $K$  é o número de variáveis monitoradas no estudo de validade e  $C$  denota o número de condições de estocagem;
2. Desdobre o arranjo  $\underline{\mathbf{X}}$  nas condições de estocagem para formar a matriz aumentada  $\mathbf{X} \in \Re^{CN \times K}$  (Figura 2.6);
3. Se necessário, aplique procedimentos de pré-processamento na matriz  $\mathbf{X}$ , obtendo  $\mathbf{X}_a$ ;
4. Realize uma PCA na matriz  $\mathbf{X}_a$  de modo a obter as matrizes de *scores* ( $\mathbf{T}$ ) e *loadings* ( $\mathbf{L}$ ), bem como a tabela de variância explicada por cada PC. Selecione as primeiras  $R$  PCs que descrevem a maior quantidade de variância do conjunto de dados originais;
5. Separe a matriz  $\mathbf{T} \in \Re^{CN \times R}$  em  $C$  matrizes  $\mathbf{T}_c \in \Re^{N \times R}$  distintas;

6. Construa gráficos de *scores* vs. tempo para as  $R$  PCs que capturam a maior quantidade de variância e identifique as  $A$  que têm relação temporal (Figura 2.7);
7. Para cada PC relacionada com o tempo, identifique as ordens de reação e determine os parâmetros cinéticos multivariados ( $k^m$ ,  $Ea^m$  e  $\alpha_{T+\delta T, T}^m$ ) utilizando os *scores* como propriedades nas Equações 1.42 a 1.46 e 1.58 (vide Capítulo I);
8. Construa gráficos de *loadings* das PCs relacionadas com o tempo e tente relacionar seu comportamento com fenômenos de degradação do produto;

Se o perfil (físico-químico, sensorial, espectroscópico, etc.) de uma amostra limítrofe estiver disponível, use os *loadings* para calcular os parâmetros multivariados de corte das  $A$  PCs relacionadas com o tempo, como segue:

9. Concatene as características da amostra limítrofe no vetor  $\mathbf{x}^T \in \Re^{1 \times K}$ ;
10. Execute a mesma rotina de pré-processamento realizada na etapa 3 sobre o vetor  $\mathbf{x}$ ;
11. Use a matriz de *loadings* para calcular os *scores* de corte para cada PC:

$$\mathbf{t}_{crit}^T = \mathbf{x}_a^T \mathbf{L} \quad (2.12)$$

12. onde  $\mathbf{t}_{crit}^T$  é o vetor de *scores* críticos ou de corte,  $\mathbf{x}_a^T$  é um vetor-linha e  $\mathbf{L} \in \Re^{K \times A}$  é a matriz de *loadings* obtida no passo 4;
13. finalmente, calcule a validade estimada do produto projetando o *score* de corte da condição acelerada na condição de mercado:

$$Val_{CM} = Val_{CA} \times \alpha_{CA,CM} \quad (2.13)$$

onde  $Val_{CM}$  e  $Val_{CA}$  correspondem, respectivamente, aos prazos de validade do produto nas condições de mercado e acelerada e  $\alpha_{CA,CM}$  é o fator de aceleração entre estas condições (vide equações 1.57 e 1.58, Capítulo I).

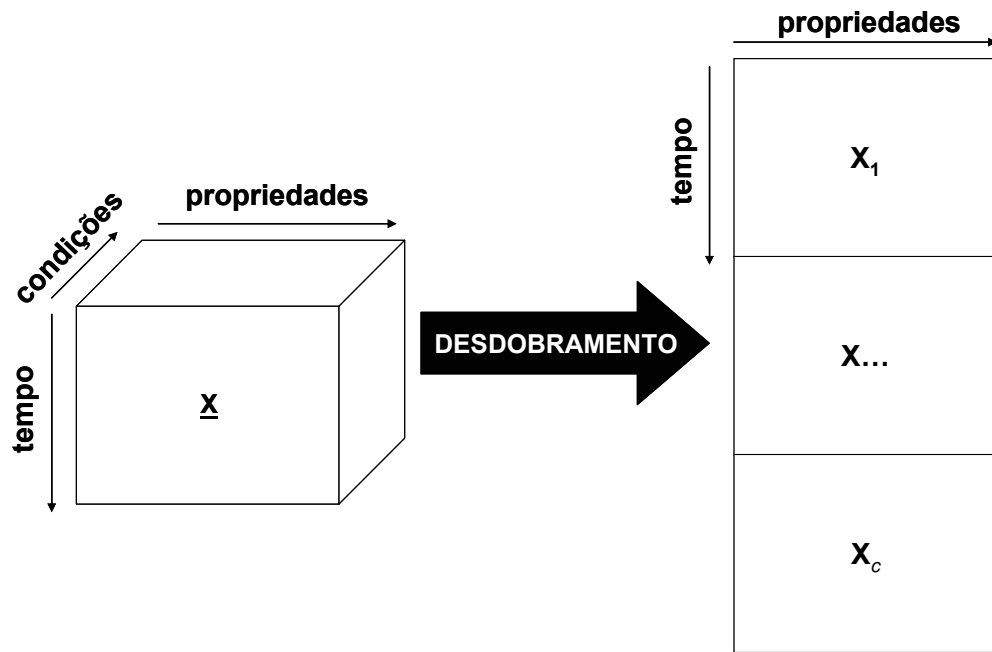


Figura 2.6 - Procedimento de desdobramento do arranjo de dados  $X$ .

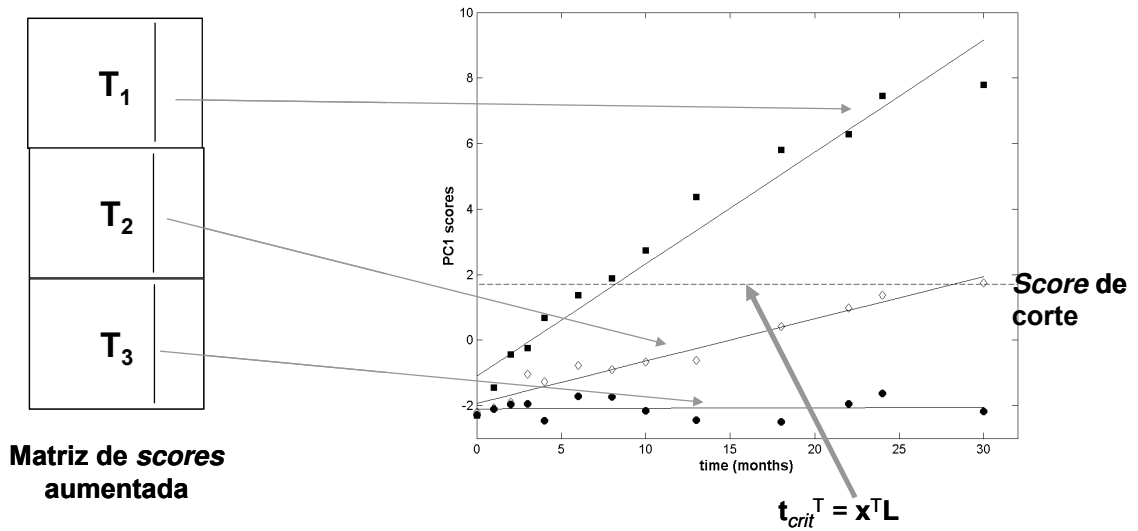


Figura 2.7 – Gráfico de *scores* vs. tempo para seleção das PCs relevantes para estudo da degradação do produto. Também são mostradas as inclinações das retas para as diferentes condições, bem como o *score* crítico.



## 5 – Conclusões

O MASLT introduz um avanço considerável na teoria clássica para determinação da validade de produtos no sentido em que fornece ferramentas multivariadas para um entendimento integrado dos modos de degradação de produtos, permitindo assim unir informações oriundas de fontes multidisciplinares, como físico-química, sensoriais e instrumentais.

Correlações entre os diferentes parâmetros que definem as características de qualidade dos produtos são fornecidas através da análise dos *loadings*. Além disso, possui ferramentas de diagnóstico de falhas na condução dos estudos de validade de produtos.

O método multivariado traz um algoritmo auto-consistente no sentido em que as conclusões são oriundas unicamente do conjunto de dados coletado durante os estudos de validade, não sendo necessário nenhum conhecimento *a priori*.

Ademais, o MASLT torna possível a utilização direta, como será demonstrado adiante, de técnicas instrumentais não-específicas, como cromatografia, espectroscopia e calorimetria, que permitem que se obtenha uma visão completa dos modos de degradação dos bens de consumo uma vez que são capazes de detectar quaisquer transformações ocorrendo nos produtos.

## 6 – Bibliografia

1. Gomez AH, Wang J, Hu G, Pereira AG, Monitoring storage shelf-life of tomato using electronic nose technique, *J. Food Eng.*, **2008**, 85, 625-631.
2. Gomes AH, Wang J, Hu G, Pereira AG, Discrimination of storage shelf-life for mandarin by electronic nose technique, *LWT – Food Sci. Tech.*, **2007**, 40, 681-689.
3. Zhu L, Seburg RA, Tsai E, Puech S, Mifsud JC, Flavor analysis in a pharmaceutical oral solution formulation using an electronic nose, *J. Pharma. Biomed. Anal.*, **2004**, 34, 453-461.
4. Casale M, Abajo MJS, Saiz JMG, Pizarro C, Forina M, Study of the aging and oxidation processes of vinegar samples from different origins during storage by near-infrared spectroscopy, *Analytica Chimica Acta*, **2006**, 557, 360-366.
5. Giardina C, Cattaneo TMP, Barzaghi S, Study of modifications in delactosated milk during shelf-life by NIR and FT-IR spectroscopy, *Milchwissenschaft-Milk Sci. Int.*, **2003**, 58, 363-366.
6. Cattaneo TMP, Giardina C, Sinelli N, Riva M, Giangiacomo R, Application of FT-NIR and FT-IR spectroscopy to study the shelf-life of crescenza cheese, *Int. Dairy J.*, **2005**, 15, 693–700.
7. Boknaes N, Jensen KN, Andresen CM, Martens H, Freshness assessment of thawed and chilled cod fillets packed in modified atmosphere using near-infrared spectroscopy, *Lebensmittel-Wissenschaft und Technologie*, **2002**, 35, 628-634.
8. Wold JP, Veberg A, Lundby F, Nilsen NA, Moan J, Influence of storage time and color of light on photooxidation in cheese: a study based on sensory analysis and fluorescence spectroscopy, *Int. Dairy J.*, **2006**, 16, 1218-1226.
9. Bodecchi LM, Cocchi M, Malagoli M, Manfredini M, Marchetti A, Application of infrared spectroscopy and multivariate quality control methods in PVC manufacturing, *Analytica Chimica Acta*, **2005**, 554, 207-217.
10. Chou SC, Shau J, Stability analysis with discrete responses, *J. Biopharm. Stat.*, **2003**, 13, 451-462.
11. Pedro AMK, Ferreira MMC, Simultaneously calibrating solids, sugars and acidity of tomato products using PLS2 and NIR spectroscopy, *Analytica Chimica Acta*, **2007**, 595, 221-227.
12. Quevedo R, Mendoza F, Aguilera JM, Chanona J, Gutierrez-Lopez G, Determination of senescent spotting in banana (*Musa Cavendish*) using fractal texture Fourier image analysis, *J. Food Eng.*, **2008**, 84, 509-515.
13. Moros J, Garrigues S, Guardia M, Evaluation of nutritional parameters in infant formulas and powdered milk by Raman spectroscopy, *Analytica Chimica Acta*, **2007**, 593, 30-38.
14. Pedro AMK, Ferreira MMC, Nondestructive determination of solids and carotenoids in tomato products by near-infrared spectroscopy and multivariate calibration, *Anal. Chem.*, **2005**, 770, 2505-2511.

15. Raatikainen O, Reinikainen V, Minkkinen P, Ritvanen T, Muje P, Pursianien J, Multivariate modeling of fish freshness index based on ion mobility spectrometry measurements, *Analytica Chimica Acta*, **2005**, 544, 128-134.
16. Dyrstad K, Vaggeland J, Thomassen C, A multivariate method to predict water vapour diffusion rate through polypropylene packaging, *Int. J. Pharma.*, **1999**, 188, 105-109.
17. Papadima SN, Arvanitoyannis I, Bloukas JG, Fournitzis GC, Chemometrics model for describing Greek traditional sausages, *Meat Sci.*, **1999**, 51, 271-277.
18. Carvalho AR, Sanchez MdN, Wattoo J, Brereton RG, Comparison of PLS and kinetic models for a second-order reaction as monitored using ultraviolet-visible and mid-infrared spectroscopies, *Talanta*, **2006**, 68, 1190-1200.
19. Tsong YI, Chen WJ, Chen CW, ANCOVA approach for shelf-life analysis of stability study of multiple factor designs, *J. Biopharma. Stat.*, **2003**, 13, 375-393.
20. Widjaja E, Lim GH, Chow PS, Tan S, Multivariate data analysis as a tool to investigate the reaction kinetics of intramolecular cyclization of enalapril maleate studied by isothermal and non-isothermal IR microscopy, *European J. Pharma. Sci.*, **2007**, 32, 349-356.
21. Dyrstad K, Thomassen C, Elvindvik K, An opportunistic stability strategy: simulation with real data, *Int. J. Phama.*, **1999**, 188, 97-104.
22. Pearson K, On lines and planes of closest fit to systems of points in space, *Philosophical Magazine*, **1901**, 2, 559-572.
23. Jurs PC, Kowalski BR, Isenhour TL, Reilly CN, Computerized learning machines applied to chemical problems, *Anal. Chem.*, **1969**, 41, 690-695.
24. Jurs PC, Kowalski BR, Isenhour TL, Reilly CN, Computerized learning machines applied to chemical problems: multicategory pattern classification by least squares, **1969**, 41, *Anal. Chem.*, **1969**, 695-700.
25. Jurs PC, Kowalski BR, Isenhour TL, Reilly CN, Computerized learning machines applied to chemical problems: Interpretation of infrared spectrometry data, **1969**, 41, *Anal. Chem.*, **1969**, 1949-1953.
26. Malinowski ER, Factor Analysis in Chemistry, John Wiley & Sons, New York, 1991.
27. Kowalki BR, Chemometrics: Mathematics and Statistics in Chemistry, D. Reidel, Dordrecht, 1984.
28. Beebe KR, Pell RJ, Seasholtz MB, Chemometrics: a Practical Guide, John Wiley & Sons, New York, 1998.
29. Jolliffe IT, Principal Component Analysis, 2<sup>nd</sup> Ed., Springer, New York, 2002.
30. Wold S, Esbensen K, Geladi P, Principal component analysis: a tutorial, *Chemometrics Intel. Lab. Sys.*, **1987**, 2, 37-52.
31. Faber NM, Meinders MJ, Geladi P, Sjöström M, Buydens LMC, Kateman G, Random error bias in principal component analysis. Part I. derivation of theoretical predictions, *Analytica Chimica Acta*, **1995**, 304, 257-271.

32. Faber NM, Meinders MJ, Geladi P, Sjöström M, Buydens LMC, Kateman G, Random error bias in principal component analysis. Part II. Application of theoretical predictions to multivariate problems, *Analytica Chimica Acta*, **1995**, 304, 273-283..
33. Schölkopf B, Smola AJ, Müller KR, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, **1998**, 10, 1299-1319.
34. Aizerman MA, Braverman EM, Rozono'er LI, Theoretical foundations of the potential method in pattern recognition learning, *Automation Rem. Control*, **1964**, 25, 821-837.
35. Kiers HAL, Weighted least squares fitting using ordinary least squares algorithms, *Psychometrika*, **1997**, 62, 251-266.
36. Martens H, Naes T, Multivariate Calibration, John Wiley & Sons, Chichester, 1993.
37. Ferreira MMC, Antunes AM, Melo MS, Volpe PLO, Quimiometria I: Calibração multivariada, um tutorial, *Quím. Nova*, **1999**, 22, 724-731.
38. Wold H, Multivariate Analysis, Academic Press, New York, 1966.
39. Bro R, Smilde AK, Centering and scaling in component analysis, *J. Chemometrics*, **2003**, 17, 16-33.
40. Kvalheim OM, Brakstad F, Liang Y, Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise, *Anal Chem.*, **1994**, 66, 43-51
41. Bjork A, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.
42. Golub G, van Loan CF; Matrix Computations; 2<sup>nd</sup> Ed., Johns Hopkins Press, London 1989.
43. Hanselman D, Littlefield B, Matlab 6 – Curso Completo, Prentice Hall, São Paulo, 2003.
44. INRIA Meta 2 Project/ENPC Cergrene, Introduction to Scilab, INRIA, 2003.
45. Strang G, Introduction to Linear Algebra, Wellesley-Cambridge Press, Wellesley 1993.
46. Strang G, Linear Algebra and its Applications, 3<sup>rd</sup> Ed., Harcourt Brace Jovanovich, Pub, San Diego, 1988.
47. Noble B, Daniel JW, Applied Linear Algebra, Prentice Hall, Englewood Cliffs, 1988.
48. Wold H, Estimation of principal components and related models by iterative least squares, in Multivariate Analysis (ed. Krishnaiah PR), Academic Press, New York, 1966, pp. 391-420.

## **CAPÍTULO III**

### **O MASLT APLICADO A DADOS FÍSICO-QUÍMICOS E SENSORIAIS:**

#### **PRODUTO CONCENTRADO DE TOMATE**

## 1 – Introdução [1]

O tomate tem papel de destaque no mercado agrícola mundial. É o segundo vegetal mais consumido globalmente, sendo a batata o primeiro [2]. Mais de 120 milhões de toneladas de tomate *in natura* foram produzidos em 2006 [3], das quais cerca de 85% foram destinadas para manufatura de produtos industrializados como concentrados e molhos de tomate [4].

O fruto é majoritariamente composto por sólidos de tomate – solúveis e insolúveis – e água. Sólidos solúveis são açúcares, como frutose e sacarose, e sais, principalmente KCl. Os sólidos insolúveis, por sua vez, são constituídos por fibras como pectina e celulose. Sólidos totais são a soma dos sólidos solúveis e insolúveis [2].

O valor de mercado do tomate é controlado por parâmetros de qualidade como aparência (ausência de defeitos visuais), ausência de doenças, cor e tamanho dos frutos. Mas como industrialmente os sólidos de tomate (solúveis ou totais) ditam o rendimento fabril – quanto mais alto seu valor menos água necessita ser evaporada e menos energia gasta nos processos de concentração – este parâmetro influencia consideravelmente o preço final pago aos produtores. Usualmente, excluindo-se pele e sementes, o tomate apresenta entre 4,5 e 8,5% de sólidos de tomate, dependendo da variedade, do solo e das condições climáticas [2, 5].

O tomate também possui micro-nutrientes e antioxidantes importantes, como os carotenóides licopeno e  $\beta$ -caroteno [6, 7]. Estes compostos possuem várias insaturações em suas estruturas químicas e, assim, estão associados à prevenção do envelhecimento e de danos oxidativos aos tecidos pela supressão de radicais livres formados em processos metabólicos do organismo [8, 9]. Enquanto uma molécula de  $\beta$ -caroteno dá origem a duas moléculas de vitamina A, que possui reconhecida ação antioxidante, pesquisas recentes têm demonstrado que o licopeno pode prevenir doenças cardíacas, arteriosclerose e alguns tipos de câncer como os de próstata, pulmão e intestino [10-12]. Além disso, outros benefícios à saúde associados ao licopeno, especialmente em sinergia com outros antioxidantes também presentes no tomate, como vitaminas C e E, estão sob investigação [13, 14].

A presença de micronutrientes, assim como outros atributos de qualidade como cor, sabor, aparência e textura, são significativamente valorizados pelos consumidores e, assim, é responsabilidade da indústria de alimentos garantir que seus produtos processados à base

de tomate entreguem estes benefícios durante todo o período de validade. Dentre estes, cor e sabor devem ser especialmente monitorados uma vez que têm papel significativo na preferência dos consumidores. Por exemplo, produtos com coloração levemente marrom ou com sabor amargo certamente serão rejeitados mesmo que apresentem altos teores de nutrientes [2].

As indústrias em geral adotam principalmente três estratégias para manter a qualidade de seus produtos durante o período de validade: aplicam processamentos adequados, evitando ao máximo a degradação do produto durante a manufatura; incluem conservantes nas formulações e/ou utilizam embalagens apropriadas, evitando que o material entre em contato com agentes externos danosos como oxigênio, luz e microorganismos [15-18].

Com relação à manufatura, os produtos de tomate passam por etapas de processo a altas temperaturas para desativar enzimas que degradam a pectina, diminuindo sua viscosidade e textura; de concentração, que permite estocagem de uma quantidade equivalente maior de frutos em um mesmo espaço disponível; e de esterilização para eliminar microorganismos que podem degradar o produto rapidamente, principalmente fungos, bolores, leveduras e bactérias. Estes processos devem, todavia, ser otimizados a fim de evitar que afetem outras características relevantes do produto. Por exemplo, a aplicação de tratamento térmico em excesso pode destruir as vitaminas e carotenóides do tomate, resultando em produtos de coloração marrom e sabor amargo [2, 6, 19-25].

Alguns alimentos recebem conservantes a fim de diminuir a taxa de perda de atributos de qualidade relevantes. Por exemplo, certos produtos de tomate são preservados com ácido cítrico ou fosfórico, pois a maioria dos microorganismos não sobrevive em pH baixo. Em outras formulações, vitamina C é adicionada em excesso com função antioxidante, a fim de preservar outros micronutrientes. Devido à presença de carotenóides em altas concentrações, também é importante evitar, através do uso de embalagens adequadas, que os produtos atomatados entrem em contato direto com a luz [7, 18, 24, 25].

Mas os alimentos à base de tomate não se degradam exclusivamente por ação de microorganismos ou oxidação de micronutrientes. Aminoácidos presentes no fruto podem reagir com açúcares redutores formando intermediários que, além de poderem conferir sabor indesejável, podem formar polímeros nitrogenados (melanoidinas) e dar origem a pigmentos que conferem coloração marrom e sabor amargo ao produto. A esta seqüência de reações dá-se o nome de escurecimento não-enzimático, ou reação de Maillard [6, 26-

28]. Apesar de ser desejada em alguns alimentos por produzir seus aromas característicos [26, 28], a reação de Maillard deve ser evitada em produtos a base de tomate. A velocidade da reação é extremamente dependente da temperatura e da atividade de água do produto e, também por este motivo, os processos térmicos e de concentração devem ser otimizados [29].

## **2 - Objetivo**

O objetivo deste capítulo é demonstrar a mecânica do MASLT e como suas premissas influenciam os estudos de prazo de validade multivariados. Para isto, utilizaram-se dados físico-químicos e sensoriais coletados de um produto derivado de tomate cuja resina da embalagem foi modificada, pois a natureza dos dados permite comparações com os métodos clássicos univariados de determinação da validade de produtos.

## **3 – Parte Experimental**

### **3.1 – Preparo e Estocagem das Amostras**

Um produto simples concentrado de tomate com 18 NTSS (*Natural Tomato Soluble Solids*) foi preparado em planta-piloto a partir de polpa de tomate duplo-concentrada com 29 NTSS obtida durante a safra brasileira de tomate. Sal e açúcar foram adicionados de acordo com a fórmula típica deste produto no mercado brasileiro. O produto foi aquecido a 120 °C e em seguida transferido para 150 latas de aço recobertas internamente com um verniz epóxi-acrílico experimental. A esterilização foi realizada em água fervente (~ 98 °C) por 15 minutos.

As amostras foram estocadas a 8, 25 e 35 °C em três estufas Marconi MA 035 com controle de temperatura pelo software Flycon v.1.0. A variação máxima de temperatura em cada câmara foi de 1 °C.



### 3.2 – Análises Físico-Químicas e Sensoriais

Quatro parâmetros instrumentais – licopeno,  $\beta$ -caroteno, cor e vitamina C, bem como oito atributos sensoriais – cor visual, sabor doce, sabor salgado, acidez, consistência, sabor de tomate verde, sabor amargo e sabor de tomate passado – foram monitorados durante 30 meses.

Realizou-se Análise Descritiva Quantitativa (ADQ) utilizando 24 indivíduos treinados [30, 31], selecionados de acordo com suas habilidades em descrever e discriminar os diversos atributos de sabor e consistência relacionados com a degradação dos produtos de tomate. Cada pessoa recebeu, em cada sessão, uma amostra para avaliação e, deste modo, a pontuação de cada atributo foi obtida pela média de 24 avaliações. As escalas para os atributos sensoriais variaram de 0 (nenhum) até 10 (muito). Para cor, 0 correspondeu a amarelo-esverdeado, 5 a vermelho escuro e 10, a marrom. Para que a cor, principalmente dos produtos já degradados, não influenciasse a percepção dos demais atributos, as avaliações de sabor foram efetuadas sob luz vermelha, enquanto que a determinação sensorial da cor foi realizada em separado, utilizando luz do dia (D65). Para remover possíveis pontuações anômalas, os dados de ADQ foram analisados através de ANOVA seguida por teste HSD de Tuckey [30, 32].

Como o licopeno,  $\beta$ -caroteno e vitamina C foram calculados em matéria seca (ms), a umidade dos produtos (%) foi determinada em triplicata utilizando um estufa Fanem EV8 a 70 °C sob vácuo (~150 mm Hg de pressão absoluta) obtido através de uma bomba de vácuo Edwards E2M8. Três gramas da amostra foram misturadas com 1 g de terra diatomácea em cápsulas de alumínio e mantidas na estufa até peso constante (~ 4 h).

O licopeno e  $\beta$ -caroteno ( $\text{mg.kg}^{-1}$  ms) foram determinados utilizando um HPLC Shimadzu equipado com um forno de coluna CTO-10A, um injetor automático Sil-10A, bombas de pistão LC-10AD e um detector UV-Vis SPD-10AV a 473 nm. A separação dos compostos foi obtida utilizando-se uma coluna RP18 Zorbax ODS (5  $\mu\text{m}$ , 15  $\times$  0.46 cm) e fase móvel MetOH:THF:H<sub>2</sub>O (67:27:6), em regime isocrático a 1,0  $\text{mL.min}^{-1}$ . Para a extração dos carotenóides da matriz de tomate, aplicou-se o procedimento descrito por Sadler *et al.* [33].

A vitamina C ( $\text{mg.100g}^{-1}$  ms) foi determinada em duplicata utilizando-se o mesmo sistema cromatográfico descrito acima, substituindo-se a coluna por uma Supelco HS C18 (150  $\times$  2.1 mm, 5  $\mu\text{m}$ ) e ajustando o comprimento de onda do detector para 260 nm a fim de se

evitar interferência de outros compostos [34]. A fase móvel utilizada foi  $\text{KH}_2\text{PO}_4$  (pH 2,3), em regime isocrático a  $0,4 \text{ mL}\cdot\text{min}^{-1}$ . Para a extração, seguiu-se o procedimento descrito por Nisperos-Carriedo *et al.* [35].

A cor dos produtos foi medida utilizando-se um colorímetro HunterLab modelo PC2 $\Delta$  com iluminante CIE tipo C e geometria 45°/0°. Calibrações com padrões preto e branco foram efetuadas antes das medições e checagens com padrão vermelho ( $L = 26,18$ ,  $a = 27,3$  e  $b = 12,7$ ) foram realizadas periodicamente. Os valores de  $\Delta E$  foram calculados de acordo com a expressão:

$$\Delta E_{i,T} = \sqrt{(L_{i,T} - L_0)^2 + (b_{i,T} - b_0)^2 + (a_{i,T} - a_0)^2} \quad (3.1)$$

onde  $L_{i,T}$ ,  $a_{i,T}$  e  $b_{i,T}$  correspondem às leituras das amostras no tempo  $i$  e temperatura  $T$ , enquanto que  $L_0$ ,  $a_0$  e  $b_0$  são as leituras do produto no tempo  $\theta = 0$ .

### 3.3 – Análise do Prazo de Validade

Os parâmetros cinéticos univariados foram determinados utilizando a teoria descrita no Capítulo I, enquanto que os parâmetros cinéticos multivariados ( $k^m$ ,  $Ea^m$  e  $\alpha_{35,25}^m$ ) foram calculados através do algoritmo do MASLT descrito no capítulo II.

Os prazos de validade univariados, bem como os critérios de corte multivariados, foram calculados utilizando-se os valores constantes na Tabela 3.1 abaixo.

Todos os cálculos foram realizados utilizando o software Matlab 6.1 com rotinas implementadas pelo autor.

**Tabela 3.1** – Perfil mínimo desejável para produtos concentrados de tomate.

<b>Variável</b>	<b>Valor-limite</b>
$\Delta E$	7
Licopeno (mg kg <sup>-1</sup> ms)	1400
$\beta$ -caroteno (mg kg <sup>-1</sup> ms)	40
Vitamina C (mg/100g ms)	170
Doçura	3,4
Salgado	4,5
Azedo	5,5
Consistência	6,5
Sabor de tomate verde	1,5
Gosto amargo	2,0
Tomate Passado	2,5
Cor	6,5

## **4 – Resultados e Discussão**

### **4.1 – Resultados Físico-Químicos e Sensoriais**

A Tabela 3.2 mostra os valores iniciais e finais das 12 variáveis monitoradas durante o estudo de validade dos produtos atomatados, nas três diferentes temperaturas de estocagem. Nota-se que nenhuma das propriedades apresentou variação significativa para as amostras estocadas a 8 °C, corroborando a premissa utilizada em vários estudos de prazo de validade de produtos de longa duração segundo a qual a degradação dos produtos é praticamente nula para amostras mantidas a baixas temperaturas [36-38]. Nestes casos, essas amostras são comumente tomadas como representativas dos produtos frescos, principalmente para realização de comparações de atributos sensoriais com produtos estocados em condições severas de temperatura.

**Tabela 3.2** – Valores iniciais e finais para os parâmetros de concentrado de tomate durante o estudo de validade.

Propriedades	$\theta = 0$	$\theta = 30$ meses		
		8 °C	25 °C	35 °C
Cor Instrumental ( $\Delta E$ )	0,33 $\pm$ 0,02	0,36 $\pm$ 0,03	7,34 $\pm$ 0,07	18,44 $\pm$ 0,09
Licopeno (mg kg <sup>-1</sup> ms)	1650 $\pm$ 216	1648 $\pm$ 152	1434 $\pm$ 105	913 $\pm$ 132
$\beta$ -Caroteno (mg kg <sup>-1</sup> ms)	52 $\pm$ 9	51 $\pm$ 3	47 $\pm$ 7	40 $\pm$ 10
Vitamin C (mg/100g ms)	386 $\pm$ 12	382 $\pm$ 14	138 $\pm$ 9	56 $\pm$ 5
Sabor doce	3,84 $\pm$ 0,16	3,89 $\pm$ 0,22	3,50 $\pm$ 0,28	3,18 $\pm$ 0,17
Sabor salgado	5,04 $\pm$ 0,09	5,12 $\pm$ 0,06	4,88 $\pm$ 0,16	4,72 $\pm$ 0,27
Sabor azedo	6,18 $\pm$ 0,22	6,13 $\pm$ 0,35	4,94 $\pm$ 0,30	3,01 $\pm$ 0,27
Consistência	8,46 $\pm$ 0,13	8,16 $\pm$ 0,19	6,19 $\pm$ 0,20	5,51 $\pm$ 0,13
Sabor de tomate verde	1,12 $\pm$ 0,16	1,09 $\pm$ 0,22	1,11 $\pm$ 0,31	1,30 $\pm$ 0,37
Sabor amargo	1,10 $\pm$ 0,46	1,32 $\pm$ 0,26	1,95 $\pm$ 0,30	2,66 $\pm$ 0,23
Sabor de tomate passado	2,21 $\pm$ 0,25	2,42 $\pm$ 0,19	1,32 $\pm$ 0,13	2,10 $\pm$ 0,18
Cor	4,87 $\pm$ 0,42	4,97 $\pm$ 0,33	6,80 $\pm$ 0,46	10,00 $\pm$ 0,82

**Obs.:** valores expressos como média  $\pm$  desvio padrão.

A cor instrumental ( $\Delta E$ ) aumentou cerca de 20 vezes para amostras estocadas a 25 °C e 55 vezes para aquelas mantidas a 35 °C. Além da medida de cor ser muito rápida e prática, é importante monitorar este parâmetro em estudos de prazo de validade de atomatados porque uma relação linear entre a degradação da cor e a formação de furosina, um subproduto da reação de Maillard, foi previamente reportada na literatura [39].

A vitamina C decaiu em torno de 64% e 85% para amostras estocadas a 25 e 35 °C, respectivamente. Todavia, este resultado era esperado uma vez que estudos anteriores comprovaram que o ácido ascórbico decompõe-se à taxas relativamente elevadas em meios aquosos, mesmo em pH reduzido [6, 40].

Um decréscimo na concentração de carotenóides também foi medido, mas em proporção significativamente inferior àquela observada para vitamina C. Licopeno diminuiu em 13% e 45%, enquanto que o  $\beta$ -caroteno caiu cerca de 10% e 23% para amostras estocadas a 25 e 35 °C, respectivamente.

Com relação às propriedades sensoriais, sabor doce, sabor salgado, sabor de tomate verde e sabor de tomate passado não apresentaram alterações significativas em nenhuma

temperatura de estocagem. As notas para sabor de tomate verde foram baixas porque o produto simples concentrado de tomate foi manufaturado com uma polpa de tomate de entressafra, sendo assim submetido a dois estágios de aquecimento. Apesar disso, não houve desenvolvimento de sabor de tomate passado durante o período de estocagem.

A acidez e o sabor amargo, assim como a consistência sensorial, apresentaram alterações moderadas durante o período de estocagem. A acidez diminuiu conforme a percepção de sabor amargo aumentou. De fato, os avaliadores descreveram o sabor global dos produtos estocados a 35 °C como “sabor de papelão” no final do estudo. A percepção de consistência diminuiu com o tempo devido à degradação da pectina, que dá a estrutura rígida aos produtos concentrados de tomate [6, 41, 42].

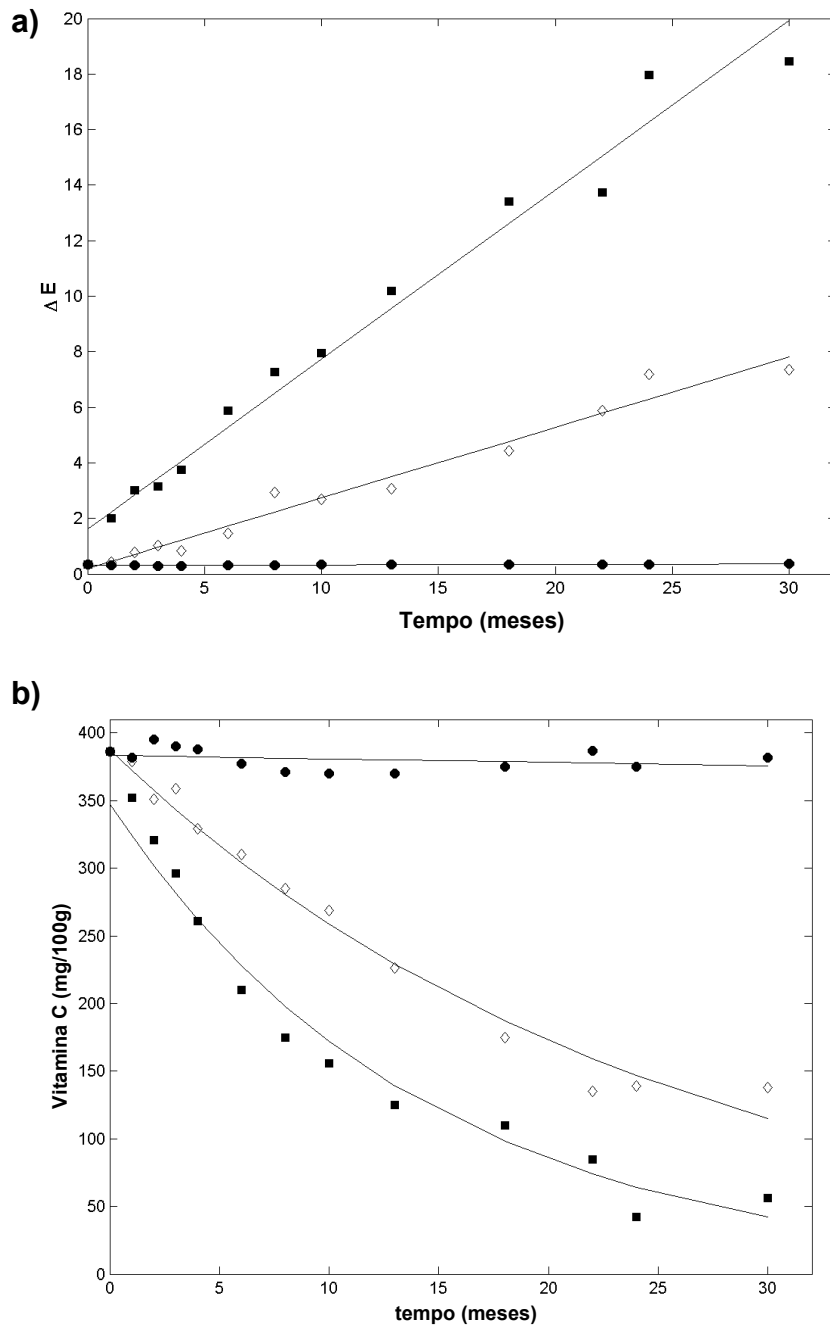
Como observado nas medidas instrumentais, a cor mostrou-se o parâmetro sensorial mais afetado durante o estudo. Variações perceptíveis (apesar de ainda aceitáveis) na cor foram detectadas pelo painel sensorial após 20 e 6 meses de estudo para amostras estocadas a 25 e 35 °C, respectivamente.

A Tabela 3.4 mostra os valores dos parâmetros cinéticos calculados para cada variável individualmente, com exceção de sabor de tomate verde, sabor salgado, e sabor de tomate passado. Observa-se que o licopeno, o  $\beta$ -caroteno, a vitamina C e o sabor azedo (acidez) seguiram cinética de pseudo-primeira ordem, enquanto que as demais propriedades seguiram cinética de ordem pseudo-zero (vide Figura 3.1 para exemplos de gráficos cinéticos). De fato, cinéticas de pseudo-primeira ordem para carotenóides já foram reportadas na literatura [43].

**Tabela 3.3 – Parâmetros cinéticos univariados das propriedades selecionadas via PCA.**

Propriedade	n	Temperatura (°C)	Constante		Energia de Ativação (kJ mol <sup>-1</sup> )
			de Velocidade k	$\alpha_{35,25}$	
$\Delta E$	zero	25	0,25	2,4	165 ± 35
		35	0,61		
Licopeno (mg kg <sup>-1</sup> ms)	1 <sup>a</sup>	25	-0,005	4,0	125 ± 10
		35	-0,020		
$\beta$ -Caroteno (mg kg <sup>-1</sup> ms)	1 <sup>a</sup>	25	-0,003	3,9	70 ± 10
		35	-0,011		
Vitamina C (mg/100g ms)	1 <sup>a</sup>	25	-0,04	1,7	130 ± 30
		35	-0,07		
Sabor doce	zero	25	-0,01	2,0	130 ± 20
		35	-0,02		
Acidez	1 <sup>a</sup>	25	-0,009	2,6	180 ± 30
		35	-0,023		
Consistência	zero	25	-0,07	1,3	105 ± 25
		35	-0,10		
Sabor amargo	zero	25	0,032	1,7	60 ± 5
		35	0,055		
Cor visual	zero	25	0,070	2,7	120 ± 15
		35	0,189		

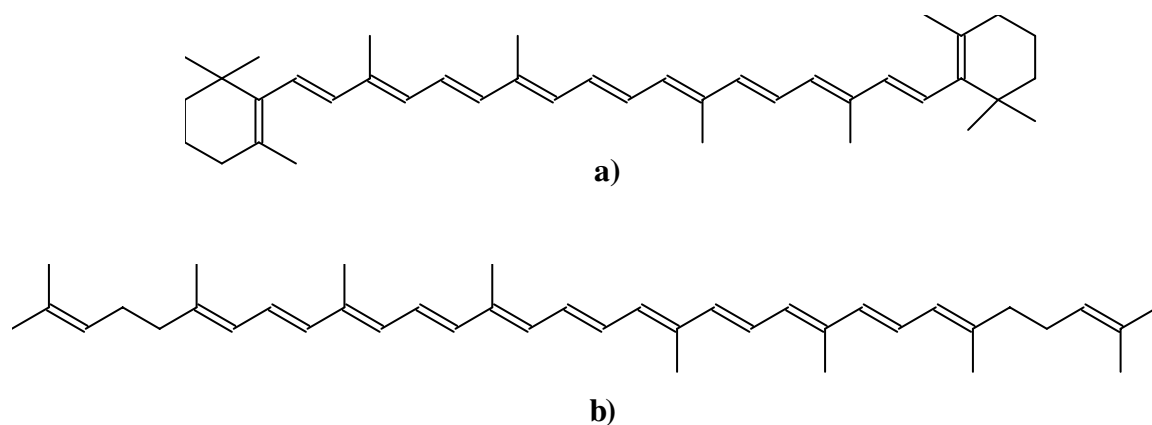
**Obs.:**  $k$  é expresso em  $\text{mês}^{-1}$  para reações de pseudo-primeira ordem e  $[\text{unidade}].\text{mês}^{-1}$  para reações de ordem pseudo-zero.



**Figura 3.1** – Gráficos cinéticos para a) cor e b) vitamina C, que exemplificam cinéticas de pseudo zero e pseudo-primeira ordem, respectivamente. (●) corresponde a amostras estocadas a 8 °C, (◇) a 25 °C e (■) a 35 °C.

As constantes de velocidade para a cor instrumental foram maiores que aquelas determinadas para cor visual, provavelmente porque a medida espectrofotométrica foi mais sensível em detectar variações deste parâmetro. Entretanto, como as energias de ativação apresentaram valores parecidos, a sensibilidade térmica foi similar para ambos os parâmetros, o que possivelmente explica a correlação significativa entre estas duas variáveis.

É interessante notar que, apesar de serem estruturalmente similares, o  $\beta$ -caroteno degradou-se mais lentamente que o licopeno: as constantes de velocidade deste último foram aproximadamente 45% menores que as do primeiro. Por outro lado, o  $\beta$ -caroteno mostrou-se termicamente mais sensível que o licopeno, apresentando uma energia de ativação aproximadamente 50% inferior. Foge ao escopo deste trabalho fornecer uma interpretação mecanística dos processos de degradação de carotenóides. Entretanto, estas diferenças devem-se possivelmente ao fato dos mecanismos de oxidação destes compostos serem ligeiramente diferentes nas condições de processamento e nas embalagens dos produtos de tomate. Por exemplo, uma das rotas de degradação do  $\beta$ -caroteno envolve a abertura de um anel de seis membros insaturado, o mesmo não ocorrendo com o licopeno (Figura 3.2) [6, 7].



**Figura 3.2** – Estruturas dos carotenóides a)  $\beta$ -caroteno e b) licopeno.

O sabor amargo se desenvolveu mais rapidamente que a queda da percepção de acidez e sabor doce pois os intermediários da reação de Maillard que conferem tal sabor formam-se antes dos polímeros que conferem coloração escura ao produto [6, 26-28].



É necessário cuidado na interpretação dos valores de energia de ativação da Tabela 3.4, principalmente no caso dos atributos sensoriais. Deve-se lembrar que, ao contrário das variáveis instrumentais, não há reação química formal envolvida, mas sim a percepção humana a um atributo sensorial. Deste modo, é incorreto afirmar, por exemplo, que o sabor ácido é sensível à temperatura. Na verdade é a percepção do sabor ácido, na forma de pontuações médias da análise descritiva quantitativa, que se mostra consideravelmente dependente da temperatura.

A Tabela 3.4 mostra os prazos de validade calculados para cada propriedade individualmente, utilizando os valores críticos fornecidos na Tabela 3.1 e os fatores de aceleração constantes na Tabela 3.3. O desafio para a determinação unívoca do prazo de validade real do produto consiste em escolher qual atributo é o mais crítico. Esta decisão pode ser tomada com base em pesquisas com consumidores, requisitos legais, informações da literatura ou na experiência prévia. Todavia, como os benefícios funcionais nem sempre são os mais valorizados pelos consumidores não é incomum que, cientistas, gestores e profissionais de *marketing* demorem a entrar em consenso, principalmente quando dados sensoriais estão envolvidos. Por exemplo, caso o menor tempo calculado fosse adotado, o produto teria 20 meses de validade. Entretanto, a experiência mostra que o sabor doce não é, por si só, tão importante para a aceitação do produto, sendo o sabor global, principalmente o balanço acidez/açúcares, mais relevante do ponto de vista dos consumidores [2, 44, 45]. De fato, o prazo de validade praticado para este produto é de 24 meses.

**Tabela 3.4** – Prazos de validade calculados para as principais propriedades do estudo univariado.

Propriedade	Prazo de Validade Calculado (meses)
$\Delta E$	21
Licopeno (mg kg <sup>-1</sup> ms)	35
$\beta$ -Caroteno (mg kg <sup>-1</sup> ms)	95
Vitamina C (mg/100g ms)	51
Sabor doce	<b>20</b>
Acidez	63
Consistência	21
Sabor amargo	26
Cor visual	24

## 4.2 – Aplicação do MASLT a Produtos Concentrados de Tomate

A PCA demonstrou que duas componentes principais capturaram 80% da variância nos dados originais, uma quantidade razoável de informação considerando-se a variabilidade intrínseca das propriedades originais (10 a 15%, vide Tabela 3.2).

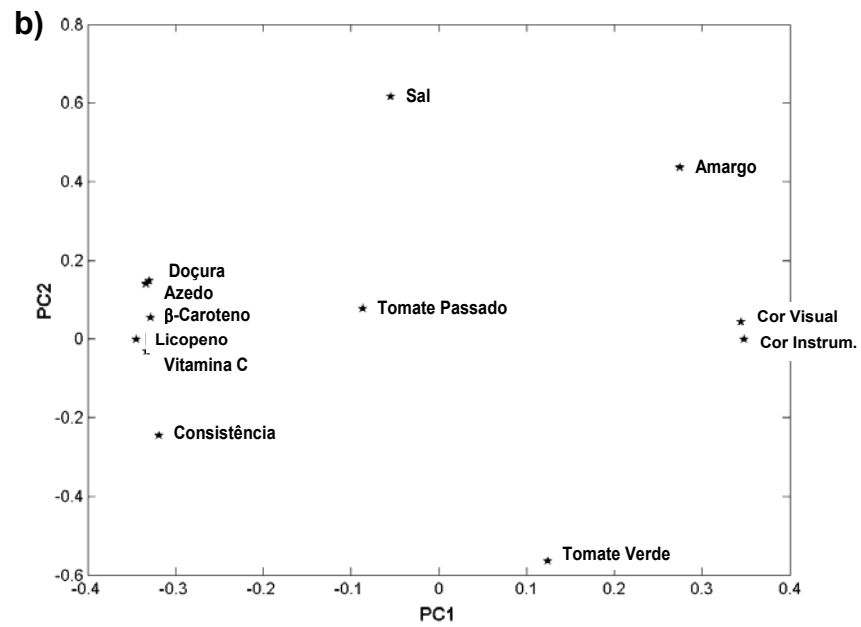
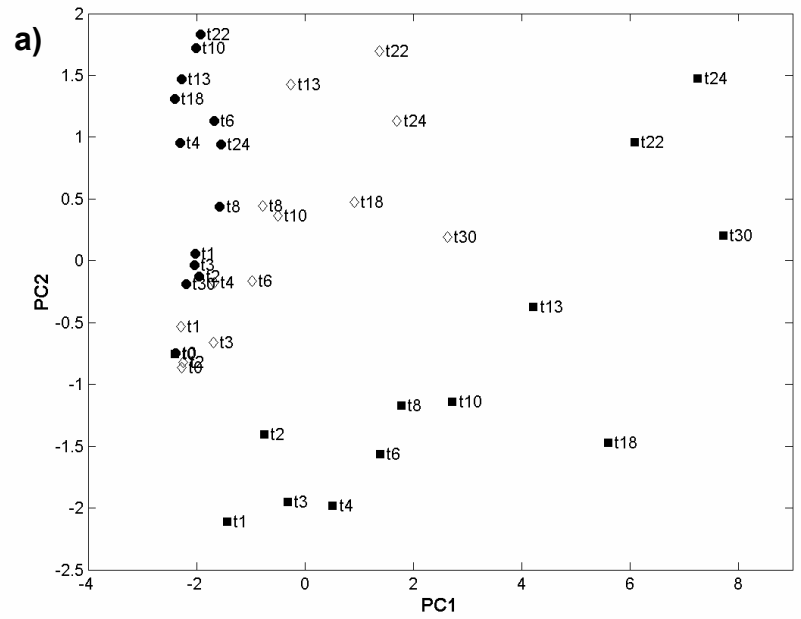
A Figura 3.3a amostra o gráfico de *scores* para as primeiras duas PCs, sendo que as amostras foram rotuladas com seus respectivos tempos de análise ( $t_i$ ) para que se pudesse visualizar a correlação de cada PC com a degradação relativa ao tempo. Nota-se que a projeção dos pontos nos *scores* de PC1 segue relação temporal seqüencial, enquanto que a relação é aleatória quando a projeção é realizada em PC2. Além disso, as amostras estocadas a 8 °C, por exemplo, apresentaram uma pequena variação na PC1, mas esta variação aparentemente não está relacionada com o tempo. Deste modo, antes mesmo de realizar a análise cinética dos *scores*, pode-se formular a hipótese de que apenas PC1 tem relação com a degradação do produto, enquanto PC2 representa majoritariamente informação não relacionada com o tempo.

Apesar de não apresentarem variação temporal, valores absolutos elevados de *scores* foram observados para as amostras mantidas a 8 °C. Este comportamento era esperado uma vez que as amostras foram estocadas sob uma condição extrema de temperatura.

O gráfico de *loadings* na Figura 3.3b revela quais foram os atributos-chave responsáveis pela degradação dos produtos de tomate. Nota-se que valores positivos de *loadings* de PC1 foram obtidos para as variáveis cujos valores numéricos aumentaram no tempo, enquanto que o contrário verificou-se para aquelas cujos valores decresceram<sup>1</sup>. Nota-se também que as propriedades tomate verde, tomate passado e sabor salgado apresentaram contribuição mínima para esta PC, em acordo com os dados originais mostrados na Tabela 3.2. Como estas variáveis não apresentaram variação com o tempo, mas contribuíram significativamente para a construção da PC2, a hipótese de que esta PC contém majoritariamente informação não relacionada com a degradação do produto, sendo constituída por ruído, torna-se ainda mais provável.

---

1 - Devido à ambigüidade rotacional da análise de componentes principais o contrário poderia ocorrer: as propriedades que tiveram seus valores numéricos aumentados poderiam receber *scores* negativos, e vice-versa (vide Capítulo II). Entretanto, aqui é importante ressaltar que a PCA foi capaz de separar as propriedades de maneira lógica.



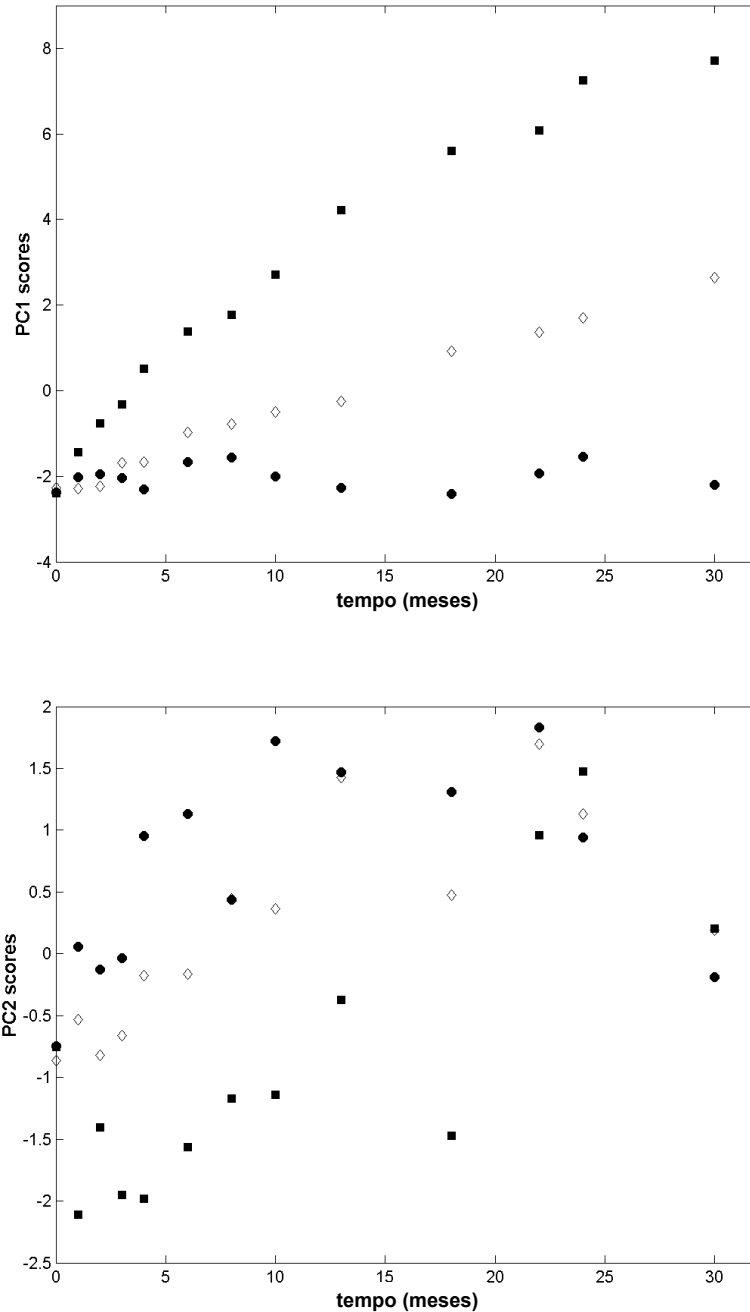
**Figura 3.3** – Gráficos de PC1 (68%) vs. PC2 (12%) para a) *scores* e b) *loadings*. Na figura a), (●) corresponde a amostras mantidas a (●) 8 °C, (◇) a 25 °C e (■) a 35 °C.

Outra vantagem em se utilizarem os gráficos de *loadings* para estudos de validade de produtos é que as correlações entre as propriedades tornam-se visualmente evidentes. Por exemplo, a cor instrumental ( $\Delta E$ ) está significativamente correlacionada com a avaliação visual de cor, assim como  $\beta$ -caroteno e vitamina C (Tabela 3.5). Correlações são especialmente relevantes quando ocorrem entre variáveis instrumentais e sensoriais pois deste modo pode-se, por exemplo, tomar a medida rápida e objetiva de cor e inferir a resposta dos métodos subjetivos, caros e demorados de avaliação sensorial.

**Tabela 3.5** – Coeficientes de correlação entre variáveis.

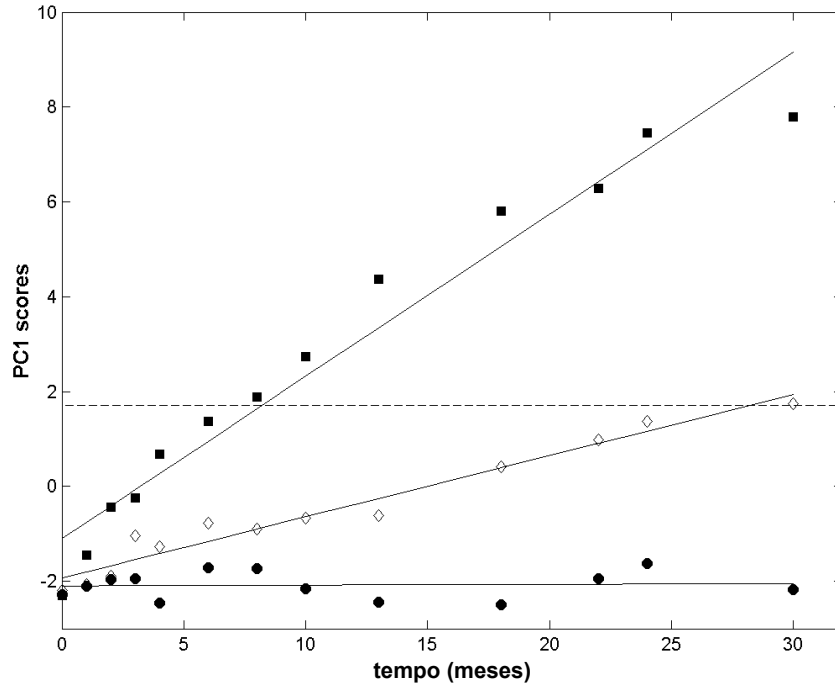
Variáveis	<i>r</i>
$\Delta E$ e cor visual	0,9866
Licopeno e $\beta$ -Caroteno	0,9323
Acidez e Doçura	0,8902
Vitamina C e Licopeno	0,8435

A Figura 3.4 mostra os gráficos de *scores* vs tempo para as duas primeiras PCs. Pode-se observar que apenas a PC1 apresentou uma estrutura clara no tempo enquanto que PC2 não mostrou qualquer relação temporal. Este modo, para estes estudo,  $A = 1$  (vide Capítulo II).



**Figura 3.4** – Comportamento das primeiras duas PCs em função do tempo para amostras estocadas a (●) 8 °C, (◇) 25 °C e (■) 35 °C.

Na Figura 3.5 é mostrado o gráfico da PC1 vs. tempo, juntamente com as regressões por mínimos quadrados para cada temperatura. Equações de cinética de ordem pseudo-zero forneceram ajustes satisfatórios, sendo que as inclinações das curvas correspondem às constantes de velocidade multivariadas ( $k^m$ ) para cada temperatura (Tabela 3.6).



**Figura 3.5** – Gráficos cinéticos multivariados para amostras mantidas a (●) 8 °C, (◆) 25 °C and (■) 35 °C. A linha tracejada representa o critério multivariado de corte.

**Tabela 3.6** – Parâmetros cinéticos multivariados para os produtos concentrados de tomate.

Temperatura (°C)	Constantes de Velocidade $k^m$ (score.mês <sup>-1</sup> )	$\alpha_{35,25}^m$	Energia de Ativação $E_a^m$ (kJ mol <sup>-1</sup> )
8	0,002 ± 0,009		120 ± 15
25	0,129 ± 0,005	2,7	
35	0,34 ± 0,04		

Uma das premissas do MASLT é que os parâmetros multivariados podem ser entendidos como médias ponderadas (pelos *loadings*) dos valores univariados obtidos na análise cinética de todas as propriedades em estudo. Em outras palavras,  $k^m$ ,  $\alpha^m$  e  $E_a^m$  seriam combinações lineares dos valores individuais, para cada propriedade, de  $k$ ,  $\alpha$  e  $E_a$ . Os valores dos parâmetros univariados na Tabela 3.4 foram multiplicados pelos *loadings* da PC1, fornecendo  $\alpha_{35,25}^* = 2,79$  e  $E_a^* = 147$  kJ.mol<sup>-1</sup> (os asteriscos indicam que estes

resultados são médias ponderadas das propriedades originais), em concordância com aqueles obtidos utilizando-se o MASLT.

Conforme demonstrado anteriormente algumas propriedades apresentaram correlação significativa (vide Figura 3.4b) e, deste modo, nem todas seriam necessárias para a determinação do prazo de validade. De fato, os cálculos foram repetidos considerando-se apenas as variáveis licopeno,  $\beta$ -caroteno, acidez, sabor amargo, consistência e cor instrumental. Os resultados foram muito similares àqueles obtidos considerando-se todas as variáveis:  $k^m_8 = 0,001$ ;  $k^m_{25} = 0,138$ ;  $k^m_{35} = 0,299$ ;  $\alpha_{35,25} = 2,3$  e  $E_a = 115 \pm 15 \text{ kJ.mol}^{-1}$ .

A aplicação dos valores de referência constantes na Tabela 3.1 à equação 2.12 (vide Capítulo II) forneceu um valor de *score* crítico de 1,7 para a PC1 (linha tracejada na Figura 3.6), que corresponde a um prazo de validade de 28 meses para amostras estocadas à condição ambiente, aqui considerada de 25 °C. Entretanto, o prazo de validade de produtos industrializados de tomate em latas é, geralmente, de 24 meses. De fato, é prática comum na indústria de alimentos reduzir o prazo de validade declarado nas embalagens de modo a garantir que os consumidores não recebam produtos insatisfatórios devido a erros inerentes de fontes de variação não inclusas no estudo para determinação do prazo de validade de produtos como, por exemplo, efeitos de transporte e oscilações de temperatura [36-38].

Para um prazo de validade estimado de 24 meses, futuros testes multivariados de estabilidade deveriam ser conduzidos por, no mínimo, 10 meses para produtos similares (reveja as definições de estabilidade e prazo de validade no Capítulo I). Temperaturas de estocagem mais elevadas poderiam ser utilizadas para acelerar ainda mais o estudo, mas novos fatores de aceleração teriam que ser determinados.

## 5 – Conclusões

O potencial do MASLT para análise de dados físico-químicos e sensoriais foi demonstrado pela determinação, com sucesso, do prazo de validade real de um produto concentrado de tomate industrializado. Por reduzir o estudo cinético a uma única variável, o método forneceu uma simplificação e redução significativa dos cálculos cinéticos realizados quando comparado ao método univariado. Além disso, forneceu informação de extrema importância sobre os fatores dominantes de degradação do produto de maneira clara e visual. Esta propriedade do método é especialmente interessante quando os resultados devem ser apresentados a pessoas não-técnicas.

As premissas do método também foram demonstradas de maneira clara nesta aplicação prática: sua auto-consistência – no sentido em que as propriedades mais relevantes para a degradação do produto recebem maiores valores de *loadings*; sua habilidade de comprimir conjuntos de dados relativamente grandes, possibilitando que as relações existentes entre as variáveis sejam aproveitadas na análise; e sua capacidade de unir dados de fontes distintas, como sensoriais e analíticos.

O método também simplificou consideravelmente o processo decisório da definição do prazo de validade do produto por ter definido um único critério global de corte. Esta característica do método é de extrema valia uma vez que na indústria, não raras vezes, cientistas, gestores e profissionais da área de *marketing* demoram a concordar critérios de corte para cada propriedade medida.



## 6 – Referências

1. Pedro AK, Ferreira MMC, Multivariate Accelerated Self-Life Testing: a novel approach for determining the shelf-life of foods, *J. Chemometrics*, **2006**, 20, 76-83.
2. Gould WA, Tomato Production, Processing & Technology, 3<sup>rd</sup> ed., CTI Pub, Baltimore, 1992, pp 15-120, 236-349.
3. FAO Statistical Database Online, <http://www.fao.org/corp/statistics.html>, 25.mai.08.
4. USDA Foreign Agricultural Service (FAS), *World Tomato and Tomato Products Situation and Outlook*, [http://www.fas.usda.gov/htp/Hort\\_Circular/2005/08-05/08-01-05%20Tomato%20article.pdf](http://www.fas.usda.gov/htp/Hort_Circular/2005/08-05/08-01-05%20Tomato%20article.pdf), 05.abril.2007.
5. Alencar RCS, O Tomateiro, LTC, São Paulo, 1979, pp. 5-45.
6. Fennema OR, Food Chemistry; 3<sup>rd</sup> Ed., Blakie Academic & Professional, London, 1996, pp. 321-615, 723-767, 1013-1029.
7. Isler O, Carotenoids, Birkhäuser Verlag, Stuttgart, 1971, pp. 1-43, 125-289.
8. Ferguson LR, Prospects for Cancer Prevention, *Mutation Res.* **1999**, 428, 329-338.
9. Lee A, Thurnhan D, Chopra M, Consumption of tomato products with olive oil but not sunflower oil increases the antioxidant activity of plasma, *Free Radical Biol. Med.*, **2000**, 29, 1051 – 1055.
10. Bramley P, Is lycopene beneficial to human health?, *Phytochem.* **2000**, 54, 233-236.
11. Tapiero H, Townsend DM, Tew KD, The role of carotenoids in the prevention of human pathologies, *Biomed. Pharmacother.* **2004**, 58, 100 – 110.
12. Willcox J, Catignani GL, Lazarus S, Tomatoes and cardiovascular health, *Crit. Rev. Food Sci. Nutr.* **2003**, 43, 1-18.
13. Giovannucci E, Tomatoes, tomato based products, lycopene and cancer: review of the epidemiologic literature, *J. Natl. Cancer Inst.* **1999**, 91, 317 – 331.
14. Jain CK, Argawal S, Rao V, The effect of dietary lycopene on bioavailability, tissue distribution, in-vivo antioxidant properties and colonic preneoplasia in rats, *Nutr. Res.* **1999**, 19, 1383-1391.
15. Tucker GS (ed.), Food deterioration and preservation, Wiley-Blackwell, New York, 2007.
16. Krebbers B, Matser AM, Hoogerwerf SW, Moezelaar R, Tomassen MM, van den Berg RW, Combined high-pressure and thermal treatments for processing tomato puree: evaluation of microbial inactivation and quality parameters, *Innovative Food Sci. Emerging Tech.*, **2003**, 4, 377-385.
17. Mathlouthi M, Food Packaging and Preservation: Theory and Practice, Elsevier, Amsterdam, 1986.
18. Russel NJ, Gould GW, Food Preservatives, 2<sup>nd</sup> Ed., Springer, New York, 2003.
19. Barbagallo RN, Chisaro M, Branca F, Spagna G, Pectin Methyl-esterase, polyphenol oxidase and physicochemical properties of typical long-storage cherry tomatoes cultivar under water stress regime, *J. Sci. Food Agric.*, **2008**, 88, 389-396.

20. van Dijk C, Boeriu C, Stolle-Smits T, Tijssens LMM, The firmness of stored tomatoes (*cv. Tradiro*). 2. Kinetic and near infrared models to describe pectin degrading enzymes and firmness loss, *J. Food Eng.*, **2006**, 77, 585-593.
21. Lin H, Qin X, Aizawa K, Inakuma T, Yamauchi R, Kato K, Chemical properties of water-soluble pectins in hot- and cold-break tomato pastes, *Food Chem.*, **2005**, 93, 409-415.
22. Simpson R, Almonacid S, López D, Abakarov A, Optimum design and operation conditions of multiple effect evaporators: tomato paste, *J. Food Eng.*, **2008**, 89, 488-497.
23. Zanoni B, Pagliarini E, Giovanelli G, Lavelli V, Modelling the effects of thermal sterilization on the quality of tomato puree, *J. Food Eng.*, **2003**, 56, 203-206.
24. Goula AM, Adamopoulos KG, Paris CC, Nikas VA, Prediction of lycopene degradation during a drying process of tomato pulp, *J. Food Eng.*, **2006**, 74, 37-46.
25. Shi J, Dai Y, Kakuda Y, Mittal G, Xue SJ, Effect of heating and exposure to light on the stability of lycopene in tomato purée, *Food Control*, **2008**, 19, 514-520.
26. Baynes JW, Monnier VM, Ames JM, Thorpe SR, The Maillard Reaction: Chemistry at the Interface of Nutrition, Aging and Disease, New York Academy of Sciences, New York, 2005.
27. Niquet C, Tessier FJ, Free glutamine as a major precursor of brown products and fluorophores in Maillard reaction systems, *J. Amino Acids*, **2007**, 33, 165-171.
28. van Boekel MAJS, Formation of flavour compounds in the Maillard reaction, *Biotech. Advances*, **2006**, 24, 230-233.
29. Hayes WA, Smith PG, Morris AEJ, The production and quality of tomato concentrates, *Crit. Rev. Food Sci. Nutr.*, **1998**, 38, 537-564.
30. Moskowitz HR, Product Testing and Sensory Evaluation of Foods, Food & Nutrition Press, Westport, 1983.
31. ISO 8587, Sensory Analysis Methodology-Ranking, 2001.
32. Barnett V, Lewis T, Outliers in Statistical Data, John Wiley & Sons, Chichester, 1978.
33. Sadler G, Davis J, Dezman D, Rapid extraction of lycopene and beta-carotene from reconstituted tomato paste and pink grapefruit homogenates, *J. Food Sci.*, **1990**, 55, 1460-1461.
34. Finley JW, Duang E, Resolution of Ascorbic, Dehydroascorbic and diketogulonic acids by paired-ion reversed phase chromatography, *J. Chromatogr.*, **1981**, 207, 449-453.
35. Nisperos-Carriedo MO, Buslig BS, Shaw PE, Simultaneous detection of dehydroascorbic, ascorbic and some organic acids in fruits and vegetables by HPLC, *J. Agric. Food Chem.*, **1992**, 40, 1127-1130.
36. Labuza TP, Shelf-Life Dating of Foods, Food & Nutrition Press, Westport, 1982.
37. Kilcast D, Subramanian P, The Stability and Validate of Food, CRC Press, New York 2000.

38. Magari R.T., Murphy K.P., Fernandez T., Accelerated Stability Model for Predicting Shelf-Life, *J. Clin. Lab. Anal.*, **2002**, 16, 221-226.
39. Wilson RJ, Beezer AE, Mitchell JC, A kinetic study of the oxidation of L-ascorbic acid (vitamin C) in solution using an isothermal microcalorimeter, *Thermochimica Acta*, **1995**, 264, 27-40.
40. Riemer J, Karel M, Shelf-life studies of vitamin C during food storage: prediction of L-ascorbic acid retention in dehydrated tomato juice, *J. Food Proc. Pres.*, **2007**, 1, 293-312.
41. Masson ML, Malvestini M, Haminiuk CWI, Reicher C, Anjos A, Siqueira Jr AA, Effect of process conditions and properties of pectin and apparent viscosity of sterilized and concentrated pulp of tomatoes, *Int. J. Food Ing.*, **2005**, 1, 1-12.
42. Fraeye I, De Roeck A, Duvetter T, Valent I, Hendrickx M, Van Loey A, Influence of pectin properties and processing conditions on thermal pectin degradation, *Food Chem.*, **2007**, 105, 555-563.
43. Bezbradical D, Jelamil AK, Petrovi SD, Iler-Marinkovi S, Na investigation of the influence of solvent on the degradation kinetics of carotenoids in oil extracts of *Calendula officinalis*, *J. Serb. Chem. Soc.*, **2005**, 70, 115-124
44. Jha SN, Matsuoka T, Non-destructive determination of acid-brix ratio of tomato juice using near-infrared spectroscopy, *Int. J. Food Sci. Tech.*, **2004**, 39, 425-430.
45. Pedro AMK, Ferreira MMC, Nondestructive determination of solids and carotenoids in tomato products by near infrared spectroscopy and multivariate calibration, *Anal. Chem.*, **2005**, 77, 2505-2511.

## **CAPÍTULO IV**

# **APLICAÇÃO DO MASLT A DADOS INSTRUMENTAIS: DETERMINAÇÃO DO PRAZO DE VALIDADE DE COSMÉTICOS POR ESPECTROSCOPIA NIR**

## 1 – Introdução [1]

A indústria de cosméticos cresce a taxas sustentáveis em todo o mundo, com ênfase aos países em desenvolvimento como Brasil, China e Índia. Este crescimento deve-se principalmente aos consistentes investimentos que a Indústria Cosmética aplica em Pesquisa e Desenvolvimento de novas tecnologias, especialmente aquelas que possam trazer novos benefícios oriundos de matérias-primas naturais. Entretanto, o desenvolvimento de novos cosméticos demanda uma considerável quantidade de recursos humanos e financeiros. Meses ou anos de pesquisa podem ser necessários até que um novo produto chegue ao mercado, dependendo do nível de complexidade da tecnologia nele aplicada.

Dentre tais pesquisas incluem-se os estudos de prazo de validade. Estes devem ser conduzidos de maneira robusta de modo a atingir os requisitos legais e, mais importante, entregar aos consumidores os benefícios primários e secundários divulgados nas embalagens.

Entre os produtos cosméticos de maior utilização no mercado encontram-se as loções corporais, que têm a função de hidratar ou manter a hidratação natural da pele [2, 3]. Estes produtos são constituídos majoritariamente por uma emulsão óleo-em-água e contém agentes hidratantes - como glicerina, sorbitol, uréia ou aminoácidos livres - e emolientes, os mais comuns sendo lanolina, petrolato e os silicones [2-4]. Enquanto a função primária dos agentes hidratantes é ligar-se às moléculas de água das camadas superiores do *stratum corneum* para manter a pele hidratada, os emolientes têm como principal papel formar uma película sobre a pele, evitando a perda de água por evaporação [2-4]. Algumas formulações de loções corporais trazem ainda ingredientes mais sofisticados, com funções específicas, como protetores solares (óxidos de titânio e zinco, metoxicinamatos e metoxidibenzometano), ácidos glicólicos como os alfa-hidroxi ácidos - que têm função exfoliante - e antioxidantes, como os carotenóides e vitaminas C e E [2, 4-7]. Não raras vezes, estes ingredientes com funções especiais promovem benefícios primários (funcionais) mais relevantes, do ponto de vista do consumidor, que a função primordial de hidratação da loção corporal.

Devido à relevância das características funcionais para os consumidores, os ingredientes promovedores dos benefícios primários de loções corporais são usualmente adicionados às formulações em quantidades suficientes para garantir sua eficácia durante toda a

validade estimada do produto. Para este fim, estudos específicos de eficácia e estabilidade são realizados. Deste modo, as características secundárias dos produtos usualmente constituem o fator limitante do prazo de validade das loções corporais. Dentre estas características incluem-se a intensidade e o caráter da fragrância, a viscosidade do produto e a estabilidade da emulsão [3, 4, 8].

Além disso, devido ao fato de sua qualidade ser dominada por características secundárias, a definição dos critérios de corte de produtos cosméticos (assim como muitos outros) constitui uma das atividades mais complexas em estudos de validade de loções corporais. Quando dados físico-químicos ou sensoriais quantitativos estão disponíveis, esta decisão é usualmente baseada em conhecimento prévio, legislação específica ou pesquisas com consumidores. Entretanto, quando técnicas instrumentais indiretas, como a espectroscopia NIR, são utilizadas, devem-se determinar formas adequadas de relacionar os dados instrumentais com as informações químicas e/ou relativas à aceitação do produto por parte dos consumidores. A tarefa torna-se extremamente complexa na medida em que a instrumentação analítica moderna é capaz de gerar conjuntos de dados com milhares de variáveis (por exemplo, cada comprimento de onda em um espectro é uma variável). O MASLT fornece ferramentas para calcular parâmetros cinéticos utilizando dados instrumentais multivariados a fim de definir o prazo de validade de um novo produto cosmético. Mais importante, o método multivariado permite que os dados instrumentais sejam relacionados com informações relativas à preferência dos consumidores.

Não existem registros recentes quanto ao uso da espectroscopia NIR para determinação do prazo de validade de produtos cosméticos. Todavia, a técnica foi aplicada a alguns estudos na área alimentícia. A utilização da técnica NIR para estudos de validade foi primeiramente descrita por Chen, Huffman e Park [9]. Eles determinaram que o padrão de alteração dos espectros NIR apresentava correlação significativa com a degradação da mioglobina em produtos de frango, constituindo assim uma medida indireta do frescor da carne. Estratégia similar foi posteriormente adotada por Nielsen *et al.* para prever o grau de frescor em peixes [10]. Giardina, Cattaneo e Barzaghi também utilizaram a técnica NIR em um estudo exploratório de desnaturação de proteínas do leite [11]. Sinelli, Barzaghi, Giardina e Cattaneo desenvolveram estudos utilizando espectroscopia NIR para estimar o prazo de validade e as principais causas de degradação em queijos [12, 13]. Mais recentemente, Camps *et al.* utilizaram a espectroscopia NIR para classificar maçãs de acordo com a maturação, durante o prazo de validade [14]. Entretanto, nenhum destes

estudos se assemelha ao MASLT no sentido em que não agregaram modelos de cinética química e, portanto, não determinaram formas de acelerar os estudos de validade.

## 2 - Objetivo

O objetivo deste capítulo é demonstrar como o MASLT pode ser aliado à espectroscopia no infravermelho próximo (NIR) para determinação do prazo de validade de um produto cosmético, dispensando assim a utilização de métodos analíticos tradicionais, usualmente mais caros e demorados que o método espectroscópico.

## 3 – Parte Experimental

Cem frascos de loção corporal foram coletados diretamente da linha de produção e estocados em câmaras climatizadas a 8, 25, 35 e 45°C, com variação de temperatura inferior a 1°C. Espectros NIR e avaliações sensoriais foram realizados nas amostras recém-produzidas e semanalmente por 10 semanas consecutivas, isto é, até  $t = 70$  dias.

Todos os espectros de infravermelho próximo foram adquiridos à temperatura ambiente (entre 20 e 22 °C) utilizando-se um espectrofotômetro Büchi NIRFlex N-500 na região entre 4000 e 10000  $\text{cm}^{-1}$  com resolução nominal de 8  $\text{cm}^{-1}$ , equipado com o módulo NIRFlex Solids. A região compreendida entre 4000 e 4500  $\text{cm}^{-1}$  foi excluída devido à presença de ruído em excesso. Assim, de acordo com o algoritmo do MASLT, para este estudo  $N = 11$  (número de coleta de informações no tempo),  $K = 1376$  (comprimentos de onda) e  $C = 4$  (condições de estocagem).

Avaliações sensoriais quanto ao aspecto (viscosidade e estabilidade da emulsão), cor e fragrância (caráter e intensidade) foram executadas por um painel composto por 10 pessoas treinadas [15, 16]. Cada pessoa recebeu uma amostra para avaliação, à temperatura ambiente (20-22 °C), perfazendo um total de 10 avaliações por atributo em cada período de tempo. As amostras estocadas a 8°C foram tomadas como referência de comparação - isto é, considerou-se que não sofreram alterações com o tempo - e os avaliadores determinaram se havia diferenças significativas entre estas e os produtos estocados nas outras condições seguindo a escala abaixo. Qualquer amostra com pontuação média igual ou superior a 3 foi considerada imprópria.

- 0 = nenhuma diferença;
- 1 = pouca diferença;
- 2 = diferença perceptível;
- 3 = claramente diferente;
- 4 = bastante diferente.

Os cálculos foram realizados utilizando-se o software Matlab v.7.1 com rotinas desenvolvidas pelo autor. O MASLT foi aplicado à matriz aumentada de dados (vide Figura 2.6, Capítulo II) sem nenhum pré-processamento e após pré-processamento por centralização na média, para comparação.

## 4 – Resultados e Discussão

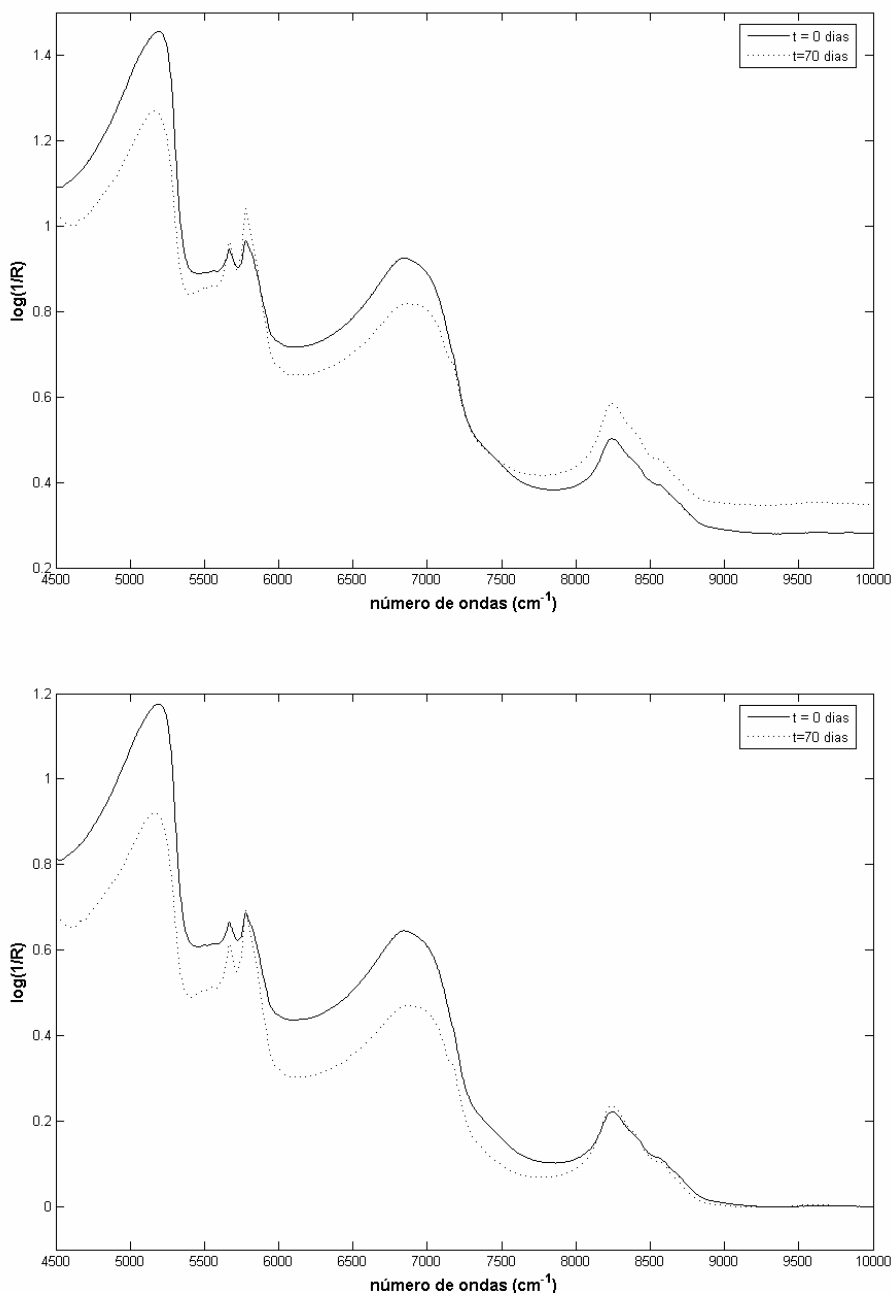
A Figura 4.1a compara os espectros de uma amostra recém-produzida com o de um produto estocado por 70 dias a 45°C. Esta temperatura constitui a condição de estocagem mais severa em estudo, onde as reações de degradação são mais pronunciadamente aceleradas e, portanto, representa de maneira satisfatória o comportamento do produto estocado por longos períodos de tempo à condição ambiente (25°C). Nota-se que o espectro da amostra estocada por 70 dias sofreu alterações significativas em comparação com aquele da amostra recém-produzida e, portanto, a espectroscopia NIR foi sensível em detectar alterações que ocorreram no produto em função do tempo e das condições de estocagem. Entretanto, notou-se também que estas diferenças espectrais devem-se em grande parte a efeitos de linha de base (note a torção nas linhas de base em aproximadamente 7200 cm<sup>-1</sup>).

A Figura 4.1b mostra os mesmos espectros com deslocamento das linhas de base para zero, tomando como referência os números de onda entre 9000 e 10000 cm<sup>-1</sup>. Nota-se uma diferença significativa no padrão de inclinação entre as linhas de base e, portanto, conclui-se que a torção observada na Figura 4.1a deveu-se principalmente à combinação de efeitos ópticos característicos de espectros NIR adquiridos por reflectância difusa [17, 18].

A medida do caminho óptico em espectroscopia por reflectância difusa é geralmente imprecisa e irreprodutível devido ao espalhamento da luz que ocorre na amostra. Este espalhamento da luz em emulsões é principalmente função do tamanho e do formato das

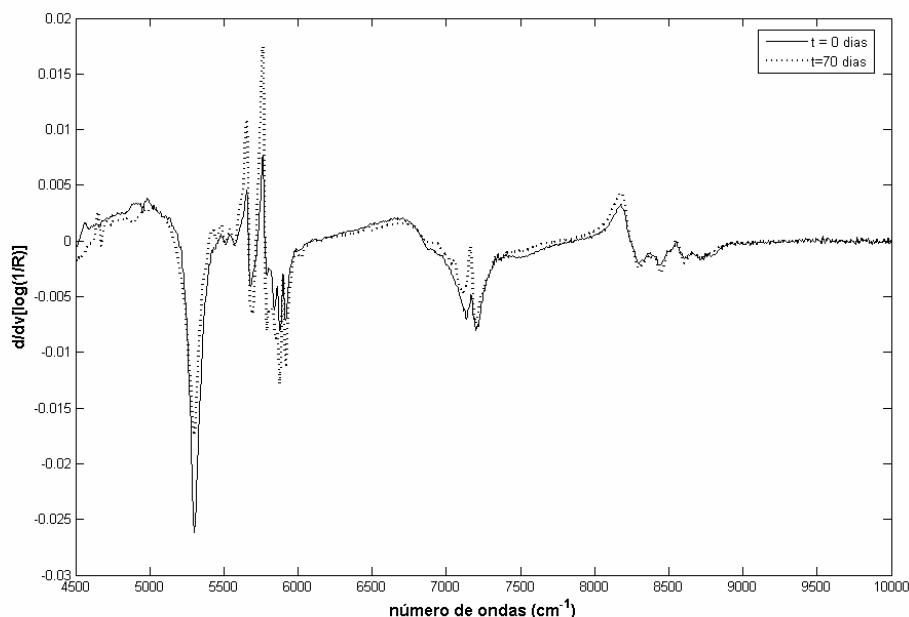


gotículas e/ou partículas, da diferença de índice de refração entre as fases aquosa e orgânica e da densidade de material particulado na amostra: quanto maior o número de partículas e/ou gotículas, quanto menor seu tamanho e quanto maior a diferença entre os índices de refração das fases aquosa e orgânica, maior o espalhamento da luz e, conseqüentemente, menor o caminho óptico na amostra [17, 18].



**Figura 4.1** – Espectros de amostras estocadas a 45°C: a) sem ajuste das linhas de base; b) ajustando as linhas de base, tomando como referência os números de onda entre 9000 e 10000 cm<sup>-1</sup>.

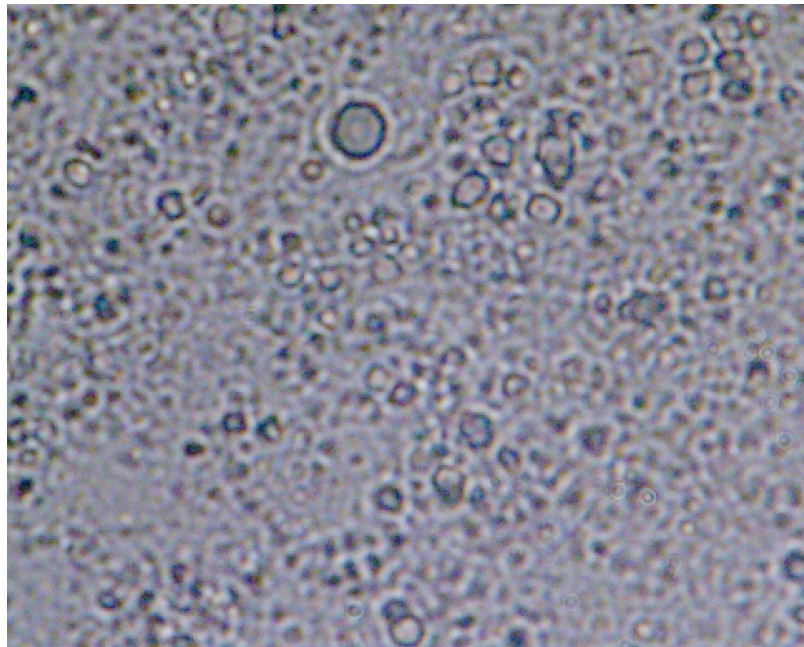
Entretanto, notam-se também efeitos espectroscópicos importantes na figura 4.1b, que podem ser melhor visualizados após eliminação das inclinações e desvios das linhas de base pela aplicação de pré-tratamento dos espectros da Figura 4.1a por derivada-primeira [19, 20]. A Figura 4.2 mostra que as bandas de vibração em torno de 5000 e 7000  $\text{cm}^{-1}$ , típicas do estiramento O-H de água, apresentaram reduções significativas de intensidade, enquanto que o efeito oposto foi observado para as bandas em 5600  $\text{cm}^{-1}$  - típica da vibração C-C - e 8300  $\text{cm}^{-1}$ , banda característica do estiramento C-O de álcoois primários [17]. Todavia, determinou-se que a variação de umidade entre as amostras foi menor que 0,5% e, assim, as diferenças nas intensidades das bandas de absorção não podem ser explicadas unicamente pela perda de água do produto.



**Figura 4.2** – Espectros de amostras estocadas a 45°C após aplicação de pré-processamento por derivada-primeira.

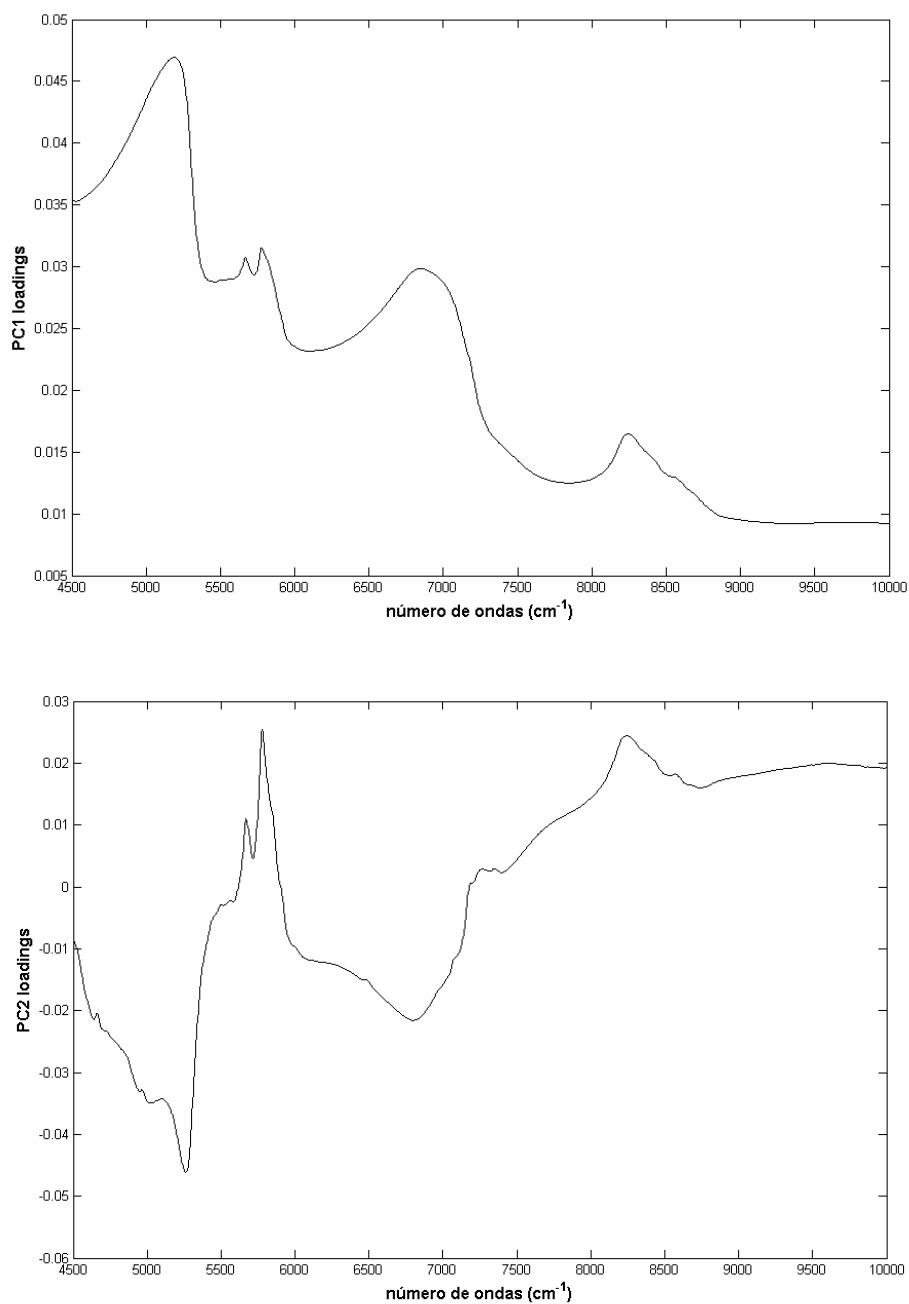
Tanto os fenômenos espectrais quanto as alterações nas linhas de base estão, todavia, provavelmente relacionados à coalescência das gotículas de óleo da emulsão em função do tempo e das condições de estocagem. Como emulsões são sistemas bifásicos estabilizados cineticamente, as gotículas de óleo tendem a se encontrar com o passar do tempo, formando glóbulos de tamanho maiores – caso as amostras fossem deixadas estocadas por tempo suficiente, o processo terminaria com a separação completa das fases, que é o sistema termodinâmico mais estável [8]. De fato, determinou-se que o

tamanho médio das gotículas de óleo alterou-se de 0,5  $\mu\text{m}$  na amostra recém-produzida para 4  $\mu\text{m}$  no produto estocado a 45°C por 70 dias (Figura 4.3). Assim, as gotas de óleo de diâmetros maiores formadas por coalescência alteraram o padrão de espalhamento da luz infravermelha, afetando assim os padrões de inclinação e deslocamento das linhas de base dos espectros NIR. Além disso, a diminuição do espalhamento da luz aumentou o caminho óptico na fase orgânica (óleo), causando a diminuição das bandas de vibração da água e aumentando a intensidade das bandas de vibração C-C.



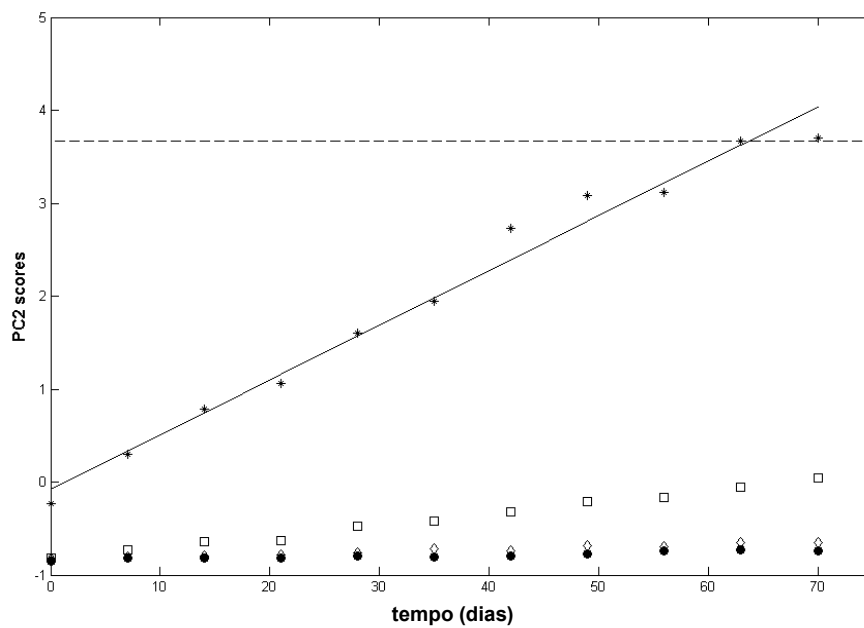
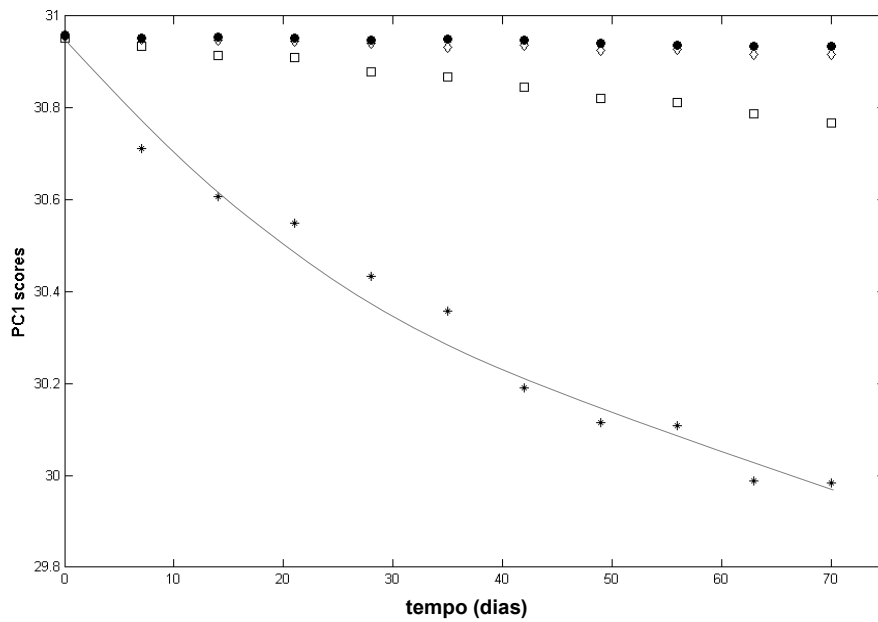
**Figura 4.3** – Micrografia da emulsão da loção corporal após 70 dias de estocagem a 45°C. As gotículas maiores são oriundas de coalescência.

Como o comportamento das linhas de base está relacionado aos processos de degradação do produto, a PCA na matriz aumentada de dados foi realizada primeiramente sem que nenhum procedimento de pré-processamento fosse aplicado aos espectros NIR. A PCA forneceu duas PCs relacionadas com o tempo, cada uma capturando, respectivamente, 98% e 1% da variância total do conjunto de dados. Os gráficos de *loadings* destas duas PCs encontram-se na Figura 4.4. Observa-se que enquanto PC1 foi caracterizada principalmente por fenômenos ocorrendo na fase aquosa, a PC2 caracterizou majoritariamente processos que se passaram na fase orgânica da emulsão. Apesar de este comportamento ser esperado para emulsões, é interessante notar como a espectroscopia NIR foi capaz de capturá-lo, bem como a forma pela qual o MASLT foi eficiente em evidenciá-lo.



**Figura 4.4** – Gráficos de *loadings* para PC1 e PC2.

Os *scores* da PC1 seguiram cinética de pseudo-primeira ordem, enquanto que os da PC2 seguiram cinética de ordem pseudo-zero (Figura 4.5). As constantes de velocidade multivariadas ( $k^m$ ), bem como os fatores de aceleração ( $\alpha_{T,25}^m$ ) e as energias de ativação ( $Ea^m$ ) são mostradas na Tabela 4.1.



**Figura 4.5** – Gráfico cinético dos scores para PC1 e PC2 de amostras estocadas a (●)8, (◇)25, (□)35 e (\*) 45°C. As linhas tracejadas representam os critérios de corte.

**Tabela 4.1** – Parâmetros cinéticos multivariados para PC1 e PC2 para o conjunto de dados espectrais sem aplicação de pré-processamento.

PC1 (98%)			
Condições de Estocagem (°C)	Constantes de velocidade $k^m$ (dia <sup>-1</sup> )	Fatores de Aceleração ( $\alpha_{T,25}^m$ )	Energia de Ativação $E_a^m$ (kJ.mol <sup>-1</sup> )
8	$1,1 \times 10^{-5} \pm 1 \times 10^{-6}$	-	73 ± 8
25	$1,8 \times 10^{-5} \pm 1 \times 10^{-6}$	-	
35	$8,5 \times 10^{-5} \pm 2 \times 10^{-6}$	4,5	
45	$42 \times 10^{-5} \pm 2 \times 10^{-5}$	25	
PC2 (1%)			
Condições de Estocagem (°C)	Constantes de velocidade $k^m$ (score.dia <sup>-1</sup> )	Fatores de Aceleração ( $\alpha_{T,25}^m$ )	Energia de Ativação $E_a^m$ (kJ.mol <sup>-1</sup> )
8	$1,6 \times 10^{-3} \pm 2 \times 10^{-4}$	-	70 ± 9
25	$2,7 \times 10^{-3} \pm 2 \times 10^{-4}$	-	
35	$12 \times 10^{-3} \pm 3 \times 10^{-4}$	4,4	
45	$58 \times 10^{-3} \pm 3 \times 10^{-3}$	22	

**Obs.:** tanto as constantes de velocidade quanto as energias de ativação estão expressas como coeficiente angular ± erro padrão do coeficiente [21].

Nota-se que, apesar de apresentarem reações de degradação com cinéticas de ordens diferentes, os fatores de aceleração e as energias de ativação para ambas as PCs foram similares. Este comportamento deve-se à relação entre os fenômenos ocorrendo nas diferentes fases da emulsão. Por exemplo, a velocidade com a qual os compostos mais voláteis da fragrância deixam o produto depende do tamanho das gotículas da fase orgânica [22, 23], que é por sua vez função do tempo.

É interessante notar que, apesar das loções corporais serem produtos de viscosidade relativamente alta (1200 a 1500 cp a 25°C [24]), as energias de ativação obtidas neste estudo foram típicas de reações ocorrendo em meios líquidos [25, 26].

A Tabela 4.1 mostra também que as constantes de velocidade para os produtos estocados a 8°C não foram nulas ( $p < 0,01$  para ambas as PCs) e, deste modo, as reações de degradação do produto não foram completamente paradas pelo resfriamento. Em termos práticos, entretanto, estas constantes de velocidade foram ainda consideravelmente baixas (cerca de 10 e 50 vezes menores que aquelas obtidas nas condições aceleradas). Deste modo, apesar das análises estatísticas dos dados obtidos através dos espectros de infravermelho terem detectado variação temporal significativa para os produtos estocados a 8°C, esta variação é, na prática, irrisória. Deste modo, estas amostras podem ser utilizadas como referência nas avaliações sensoriais sem grandes conseqüências sobre as conclusões quanto ao prazo de validade do produto.

Os atributos sensoriais não apresentaram qualquer variação significativa para as amostras estocadas a 25 e 35°C uma vez que as pontuações máximas de amostras estocadas por 70 dias foram menores que 2. Entretanto, amostras estocadas a 45°C receberam pontuação média igual a 3 após 63 dias para o parâmetro intensidade da fragrância (Tabela 4.2). Assim, não há evidências da ocorrência de reações ou quaisquer desequilíbrios entre os componentes do perfume em função do tempo, havendo apenas uma diminuição proporcional dos compostos que o constituem em função de volatilização. Como a intensidade da fragrância é um parâmetro chave na aceitação do produto pelos consumidores, tomou-se o espectro desta amostra como parâmetro de corte.

É importante salientar trata-se de uma coincidência deste estudo específico o fato de a amostra estocada por 63 dias ter recebido pontuação média exatamente igual a 3. É usual, em estudos práticos de determinação do prazo de validade de produtos, que a pontuação de uma determinada avaliação exceda o limite aceitável estipulado para um parâmetro específico de qualidade – por exemplo, uma pontuação maior que 3 poderia ser obtida na avaliação após 63 dias de estocagem a 45°C. Neste caso, o espectro equivalente à pontuação-limite poderia ser calculado por interpolação entre aqueles oriundos das amostras estocadas por 56 e 63 dias.

**Tabela 4.2 – Pontuações das avaliações sensoriais para amostras estocadas a 45°C.**

Tempo (dias)	Pontuações Médias				
	Fragrância: Caráter	Fragrância: Intensidade	Cor do Produto	Viscosidade do Produto	Estabilidade da Emulsão (visual)
7	0,6	0,2	0,0	0,1	0,1
14	0,5	0,2	0,2	1,1	0,1
21	0,8	1,2	0,5	1,3	0,1
28	0,8	1,3	1,1	1,5	0,3
35	0,9	1,4	1,2	1,7	0,2
42	1,1	2,0	0,9	1,7	0,3
49	1,1	2,2	1,2	1,9	0,6
56	1,3	2,8	1,5	2,2	0,8
63	1,5	<b>3,0</b>	1,3	2,1	1,1
70	1,8	<b>3,3</b>	2,1	2,3	1,0

Finalmente, estimou-se a validade do produto como sendo de 3 anos e 9 meses. O cálculo (Equação 4.1) foi realizado projetando-se o critério de corte no tempo crítico ( $\theta_{crit,45}$ ) da PC2<sup>1</sup>, utilizando  $\alpha_{45,25} = 22$ :

$$\text{validade}_{25} = \alpha_{45,25} \times \theta_{crit,45} = 22 \times 63 = 1386 \pm 120 \text{ dias} \quad (4.1)$$

onde o erro em dias foi calculado por propagação de erros da estimativa de  $\alpha_{42,25}$ .

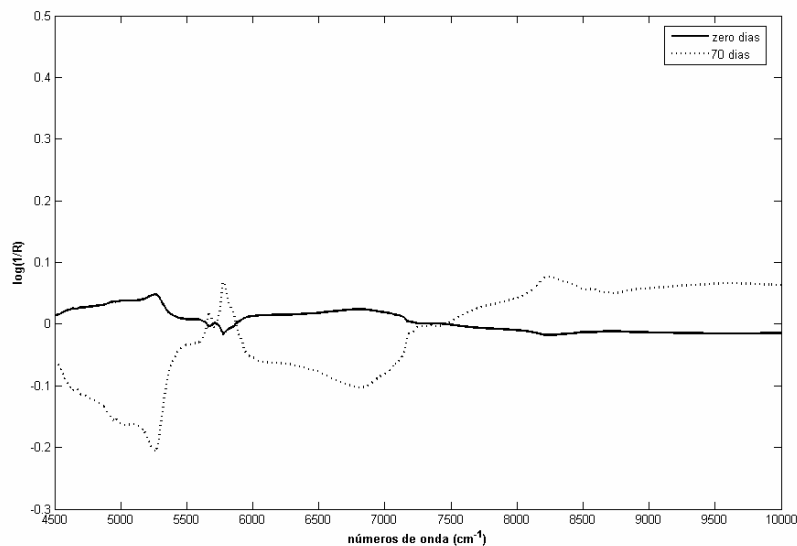
O prazo de validade de produtos similares, determinado em estudos anteriores utilizando técnicas convencionais de análise físico-química e sensorial, é de 3 anos. O resultado obtido neste trabalho pode então ser considerado adequado uma vez que é prática comum na indústria reduzir o prazo declarado nas embalagens dos produtos para acomodar fontes de variação não inclusas nos estudos de validade, como por exemplo efeitos de transporte e oscilações de temperatura.

O MASLT também foi executado utilizando-se os espectros pré-processados por centralização na média (Figura 4.6) para comparação. Ao contrário do caso onde nenhum pré-processamento foi aplicado, apenas uma componente principal capturou virtualmente toda a variância no conjunto de dados (99,9998%). A análise cinética dos *scores* dos dados pré-processados forneceu exatamente os mesmos parâmetros cinéticos obtidos pela regressão dos *scores* da PC2 do conjunto de dados sem aplicação de nenhum pré-processamento. Este resultado não é surpreendente uma vez que os *loadings* da PC1 dos dados pré-processados por centralização na média foram exatamente iguais àqueles obtidos na segunda componente principal do conjunto de dados sem pré-processamento. Assim, o pré-processamento dos dados por centralização na média foi eficiente em eliminar fenômenos de linha de base. Por outro lado, tais fenômenos apresentaram correlação com alterações temporais do produto, relevantes para interpretação de seu comportamento de degradação e, deste modo, conclui-se que quaisquer pré-processamentos devem ser aplicados com cautela para estudos de prazo de validade utilizando não apenas o MASLT, mas métodos de compressão de dados similares à PCA.

---

1 - Como PC2 possuía o menor valor de constante de aceleração, esta componente principal forneceu a estimativa mais conservadora da validade do produto. A estimativa para PC1, cujo  $\alpha_{45,25} = 25$ , forneceria um prazo de 4 anos e 3 meses.





**Figura 4.6** –Espectros de amostras estocadas a 45°C após aplicação de pré-processamento por centralização na média.

## 5 – Conclusões

O MASLT foi eficiente em determinar o prazo de validade de um produto cosmético por vias instrumentais uma vez que forneceu o mesmo prazo obtido com os métodos físico-químicos convencionais. Além disso, forneceu informações relevantes, através dos gráficos de *loadings*, que permitiram interpretar os fenômenos físico-químicos e espectrais que ocorreram no produto em função do tempo.

O método multivariado permitiu que as vantagens da espectroscopia NIR fossem agregadas aos estudos de validade do produto. Dentre estas vantagens podem-se citar velocidade e facilidade de execução das leituras, a eliminação do uso e da manipulação de solventes e reagentes químicos e a dispensa em realizar procedimentos de abertura de amostras.

Entretanto, é importante ressaltar que a técnica NIR raramente possibilita que se determinem as causas diretas da degradação do produto, informações relevantes especialmente quando novas tecnologias são introduzidas. Nestes casos, é necessário que calibrações multivariadas que relacionem as concentrações de ingredientes específicos com os espectros de infravermelho, sejam construídas. Alternativamente, a aquisição de parâmetros físico-químicos formais, em paralelo com dados espectroscópicos e sensoriais, pode ser realizada a fim de fornecer informações valiosas para desenvolvimentos futuros de produtos.

## 6 – Referências

1. Pedro AMK, Ferreira MMC, The Use of NIR Spectroscopy for Determining the Shelf-Life of Products, in press.
2. Leyden JJ, Rawlings AV, Skin Moisturization, Marcel Dekker, New York, 2002, pp. 1-29, 61-80.
3. Rawlings E *et. al.*, Skin Care: Theories and Applications, Allured, 2006, pp. 3-6, 81-96.
4. Schueller R, Romanowski P, Conditioning Agents for Hair and Skin, Marcel Dekker, New York, 1999, pp.
5. Song J, Koch J, Squier M, Skin Blemish Treatment, US Patent 5 843 998, **1998**.
6. Ishii N, Wada K, Futami J, Takama M, Cosmetic Composition, US Patent 6 942 878, **2005**.
7. Tai LHT, Formulating Detergents and Personal Care Products: a Complete Guide to Product Development, AOCS Press, Champaign, 2000, pp, 227-241.
8. Sjöblom J, Emulsions and Emulsion Stability, Marcel Dekker, New York, 1996, pp. 1-39, 237-286, 393-436.
9. Chen YR, Huffman RW, Park B, Changes in the visible/near-infrared of chicken carcasses in storage, *J. Food Process Eng.*, **1996**, 19, 121-134.
10. Nielsen H, Esaiassen M, Heia K, Sigernes F, Visible/near-infrared spectroscopy: a new tool for the evaluation of fish freshness?, *J. Food Sci.*, **2002**, 67, 1821-1826.
11. Giardina C, Cattaneo TMP, Barzaghi S, Study of modifications in delactosated milk during shelf-life by NIR and FT-IR spectroscopy, *Milchwissenschaft-Milk Sci. Int.*, **2003**, 58, 363-366.
12. Cattaneo TMP, Giardina C, Sinelli N, Riva M, Giangiacomo R, Application of FT-NIR and FT-IR spectroscopy to study the shelf-life of crescenza cheese, *Int. Dairy J.*, **2005**, 15, 693–700.
13. Sinelli N, Barzaghi S, Giardina C, Cattaneo TMP, A preliminary study using Fourier transform near infrared spectroscopy to monitor the shelf-life of packed industrial ricotta cheese, *J. Near Infrared Spectrosc.*, **2005**, 13, 293-300.
14. Camps C, Ghillermin P, Mauget JC, Discrimination of storage duration of apples stored in a cooled room and shelf-life by visible-near infrared spectroscopy, *J. Near Infrared Spectrosc.*, **2007**, 15, 169-177.
15. ISO 8587, Sensory Analysis Methodology-Ranking, 2001.

16. Moskowitz HR, Product Testing and Sensory Evaluation of Foods, Food & Nutrition Press, Westport, 1983.
17. Williams P, Norris K (eds.), NIR Technology in the Agricultural and Food Industries, Am. Ass. Cereal Chemists, St, Paul, 1990, pp.1-31.
18. Olinger JM, Griffiths PR, Effects of sample dilution and particle size morphology on diffuse reflecton spectra of carbohydrate systems in the near-infrared and midinfrared 2, durum-wheat, *Appl. Spectrosc.*, **1993**, 47, 695-701.
19. Vandeginste DGM, Massart DL, Buydens LMC, de Jung S, Lewi PJ, Smeyers-Verbeke J, Handbook of Chemometrics Part B, Elsevier, Amsterdam, 1998, pp. 349-370.
20. Savitzky A, Golay MJE, Smoothing plus differentiation of data by simplified least squares procedures, *Anal. Chem.*, **1964**, 36, 1627.
21. Box GEP, Hunter JS, Hunter WG, Statistics for Experimenters, 2<sup>nd</sup> Ed., Wiley Intersciences, Hoboken, 2005, 133-156, 437-482.
22. Theimer ET, Fragrance Chemistry: the Science and Sense of Smell, Academic Press, San Diego, 1982, pp. 575-610.
23. Butler H, Perfumes, Cosmetics and Soaps, 10<sup>th</sup> Ed., Kluwer Academic Pub., Dordrecht, 2000, pp. 505-710.
24. Slavtcheff CS, Gonzales GJ, Mokati MJ, US Patent 5 814 313, **1998**.
25. Moelwyn-Huges EA, The Kinetics of Reaction in Solution, Clavendon Press, Oxford 1933, pp. 1 – 56.
26. Connors KA, Chemical Kinetics: the Study of Reaction Rates in Solution, VCH, New York 1990, pp. 1-29.

## **CAPÍTULO V**

# **COMPARAÇÃO ENTRE O MASLT E O MÉTODO TRI-LINEAR PARAFAC**

## 1 – Introdução

Métodos multi-modos de análise de dados vêm ganhando popularidade em análise química devido ao aumento da capacidade computacional e da complexidade dos dados oriundos da instrumentação analítica moderna: espectros de emissão de fluorescência podem ser adquiridos utilizando-se vários comprimentos de onda de excitação, para diversas amostras; espectros de massa podem ser coletados para diversos picos em um cromatograma; imagens; e perfis de amostras em função do tempo [1-3].

Dados multi-modos são aqueles que podem ser arranjados em um cubo (ou hipercubo) por manterem relações cruzadas em mais de uma via, como é o caso nos testes de validade de produtos: os valores de cada propriedade, para cada amostra, são função do tempo e da condição de estocagem [1-3, 4, 5].

Os principais métodos multi-modos são o PARAFAC (*Parallel Factor Analysis*), Tucker3, *unfolding-PCA* (u-PCA) e a ANOVA multi-modos. Existem vantagens e desvantagens na aplicação de cada método, sendo o mais adequado para cada aplicação usualmente selecionado empiricamente [1-3]. Entretanto, Smilde, Bro e Geladi [1] argumentam que existe uma hierarquia entre os métodos multi-modo: o PARAFAC pode ser considerado uma versão restrita do Tucker3, que por sua vez pode ser considerado uma versão restrita do u-PCA. Kiers [6] demonstrou que u-PCA apresentará sempre um ajuste melhor que o Tucker 3, que por sua vez será melhor, em termos de ajuste, que o PARAFAC. Todavia, este último utiliza um número menor de graus de liberdade e, utilizando este critério, Bro [2] classifica o PARAFAC como mais simples que o Tucker3, que por sua vez seria mais simples que o u-PCA. Assim, qualquer conjunto de dados que possa ser modelado adequadamente pelo PARAFAC poderá ser modelado utilizando-se os outros dois métodos. Todavia, se um modelo PARAFAC for adequado para um dado conjunto de dados, os demais métodos tenderão a utilizar os demais graus de liberdade para modelar ruído ou informação sistemática de maneira redundante e, deste modo, seria preferível utilizar o modelo mais simples [1-3].

Apesar da u-PCA ser geralmente considerada mais simples que o PARAFAC, sua interpretação nem sempre é fácil porque durante o desdobramento dois modos serão misturados, dificultando a interpretação dos *loadings*. Além disso, a principal motivação para a utilização de métodos como PARAFAC não é obter modelos com melhor ajuste, mas sim modelos de interpretação mais direta e robusta [1-3, 7-10].

O método PARAFAC já foi utilizado para estudo de dados temporais. Niazi, Ghasemi e Yazdanipour [11] utilizaram matrizes cinético-espectrofotométricas para determinação de nitrito em água e amostras de carne. De maneira análoga, Silva e Oliveira [12] utilizaram PARAFAC para determinar quantitativamente o teor de metais em águas em concentrações na ordem de micromols e Nahorniak *et al.* [13] para avaliar a degradação de pesticidas e determinar sua concentração em soluções. Reis e colegas [14] realizaram resolução de curvas espectrais (*spectral curve resolution*) para avaliar a luminescência total em tártaro dental humano. Bro e Jakobsen [15] publicaram trabalho comparando PARAFAC e GEMANOVA (*Generalized Multiplicative ANOVA*) para estudar interações entre fatores que resultam na alteração da coloração de carne em função do tempo. Entretanto, existem poucas publicações de estudos dos modos de degradação de produtos utilizando PARAFAC. Bosco e Larrechi [16] publicaram um estudo interessante sobre os efeitos de variáveis de processos sobre a cinética fotocatalítica de fenóis unindo PARAFAC e planejamentos experimentais. Tan e colegas [17] utilizaram a técnica para estudar a degradação de clorofilas presentes em espinafre. Let, Jacobsen e Meyer [18] estudaram a estabilidade de leite enriquecido com óleo de peixe e Christensen *et al.* [19] aplicaram espectroscopia de fluorescência para avaliar a degradação de queijo durante a estocagem. Entretanto, não há publicações onde o PARAFAC tenha sido utilizado para determinação efetiva do prazo de validade de produtos.

## **2 - Objetivos**

Este capítulo tem por objetivo comparar o Teste Acelerado Multivariado para Determinação do Prazo de Validade de produtos (MASLT) com o método tri-linear PARAFAC de maneira a enfatizar as vantagens e desvantagens de cada método.

### 3 – *Parallel Factor Analysis* (PARAFAC)

O PARAFAC é um método de decomposição de matrizes similar ao método bi-linear PCA [20-23]. Originário da psicometria [24, 25], o modelo foi independentemente proposto por Harshman [24] e Carroll *et al.* [25] sob o título de CANDECOMP (*canonical decomposition*).

Diferentemente dos métodos bi-lineares, que apresentam liberdade rotacional, o PARAFAC fornece soluções únicas caso os dados sejam realmente tri-lineares (isto é, descritos pela expressão 5.1 abaixo) [1-3, 7, 8]. Esta propriedade traz uma série de benefícios para determinadas aplicações. Por exemplo, os *loadings* necessitam ser rotacionados adequadamente para refletir espectros de componentes puros em decomposições bi-lineares, o que só é alcançado utilizando-se informação *a priori* [26-28], o que não ocorre no PARAFAC.

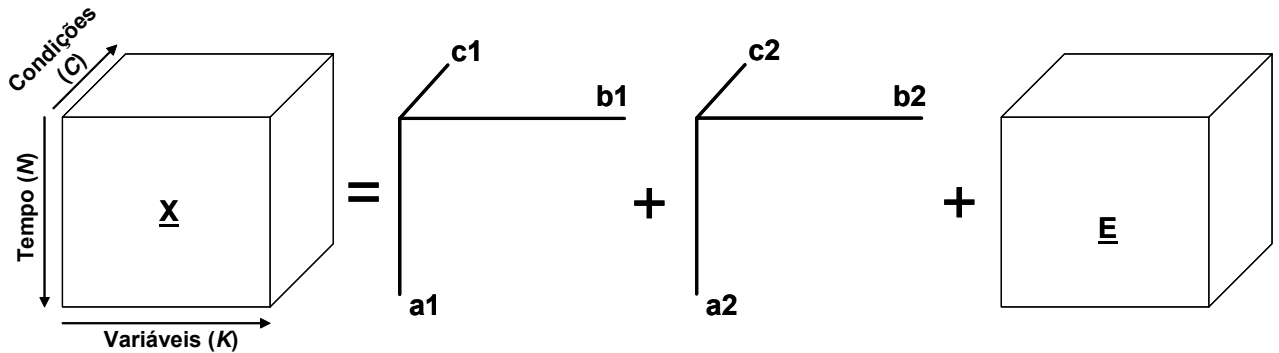
Além disso, enquanto os fatores na PCA são ortogonais, o mesmo não ocorre no PARAFAC. Apesar de restrições de ortogonalidade poderem ser introduzidas no algoritmo de cálculo, esta não é uma característica geral da decomposição [1-3, 29].

Outra propriedade importante dos métodos tri-lineares em geral é que os modelos não são complementares com relação ao número de fatores, como ocorre com os métodos bi-lineares como a PCA. Na análise de componentes principais, um modelo com  $R$  fatores pode ser obtido a partir do modelo com  $R+1$  fatores simplesmente subtraindo-se uma componente principal. O mesmo não ocorre, entretanto, nos métodos tri-lineares e, deste modo, o número de fatores a extrair deve ser determinado *a priori*. Esta característica torna necessária a construção de vários modelos antes que o melhor seja selecionado e, por este motivo, métodos multi-modo são geralmente lentos e consomem uma quantidade significativamente maior de recursos computacionais que os métodos bi-lineares [1-3, 29-31].

#### 3.1 – Representação Matemática do Método PARAFAC

Seguindo a convenção dos modelos PARAFAC, neste trabalho *loadings* do primeiro modo, aqui denominados *loadings* temporais, serão designados pela letra **A** e estocados na matriz **A**. Analogamente, os *loadings* espectrais (ou do segundo modo) na matriz **B** e os referentes às condições de estocagem (terceiro modo), na matriz **C** [1-3, 30, 31, 32].

Considerando-se um cubo de dados  $\underline{\mathbf{X}} \in \mathbb{R}^{N \times K \times C}$ , onde  $N$  é o número de amostragens no tempo,  $K$  é o número de variáveis (ou comprimentos de onda) e  $C$  é o número de condições de estocagem, o modelo PARAFAC é representado como a somatória de componentes tri-lineares (Figura 5.1) ao invés da somatória do produto vetorial de dois componentes, como ocorre na PCA bi-linear (vide Figura 2.3, capítulo II).



**Figura 5.1** – Representação geométrica do PARAFAC.

Um elemento do arranjo tri-linear  $\underline{\mathbf{X}}$  pode ser então representado matematicamente como:

$$x_{nkc} = \sum_{r=1}^R a_{nr} b_{kr} c_{cr} + e_{nkc} \quad (5.1)$$

em que  $x_{nkc}$  representa um elemento do arranjo  $\underline{\mathbf{X}}$ ,  $a_{nr}$ ,  $b_{kr}$  e  $c_{cr}$  representam *loadings* de cada modo e  $e_{nkc}$  representa o resíduo. O modelo também pode ser representado na forma de um produto de tensores:

$$\underline{\mathbf{X}} = \sum_{r=1}^R \mathbf{a}_r \times \mathbf{b}_r \times \mathbf{c}_r \quad (5.2)$$

onde  $\mathbf{a}_r$ ,  $\mathbf{b}_r$  e  $\mathbf{c}_r$  representam a  $r$ -ésima coluna das matrizes de *loadings*  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$ , respectivamente.

Encontrar os *loadings* de um modelo PARAFAC consiste, então, em minimizar a função de perda ( $L$ ):



$$L = \min_{\mathbf{A}, \mathbf{B}, \mathbf{C}} \sum [\underline{\mathbf{X}} - \mathbf{A} \times \mathbf{B} \times \mathbf{C}]^2 \quad (5.3)$$

Existem diversos algoritmos para realizar a decomposição do arranjo  $\underline{\mathbf{X}}$  na análise PARAFAC (isto é, encontrar a solução da Eq. 5.3), o mais comum sendo os quadrados mínimos alternantes (ALS – *Alternating Least Squares*). Apesar de cada algoritmo ser mais aplicável a casos particulares, Tomasi e Bro [30] argumentam que o algoritmo ALS constitui o método de modelagem PARAFAC mais conveniente para a maior parte dos conjuntos de dados tri- e multi-lineares porque tem custo computacional menor, é mais intuitivo, acomoda dados faltantes e geralmente tem convergência mais rápida para a maior parte das aplicações.

No método ALS, o arranjo é desdobrado em um dos modos, para os quais os *loadings* são calculados a partir de uma primeira estimativa (iniciação). O processo é repetido para os três modos até convergência. O algoritmo pode ser representado da seguinte forma [1-3, 29, 31]:

1. Defina o número de componentes,  $R$ ;
2. Estime valores para as matrizes de *loadings*  $\mathbf{B}$  e  $\mathbf{C}$ ;
3. Estime a matriz de *loadings*  $\mathbf{A}$  a partir de  $\underline{\mathbf{X}}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  por quadrados mínimos:

$$\mathbf{A} = \mathbf{XZ}^T(\mathbf{ZZ}^T)^{-1} \quad (5.4)$$

onde  $\mathbf{X}$  é o arranjo desdobrado ( $\mathbf{X} \in \mathfrak{R}^{N \times KC}$ ) e  $\mathbf{Z} = \mathbf{B} \otimes \mathbf{C}$ . O símbolo  $\otimes$  representa o produto de Kronecker que, dadas duas matrizes  $\mathbf{A} \in \mathfrak{R}^{K \times J}$  e  $\mathbf{B} \in \mathfrak{R}^{V \times U}$ , é calculado como [1-3, 33, 34]:

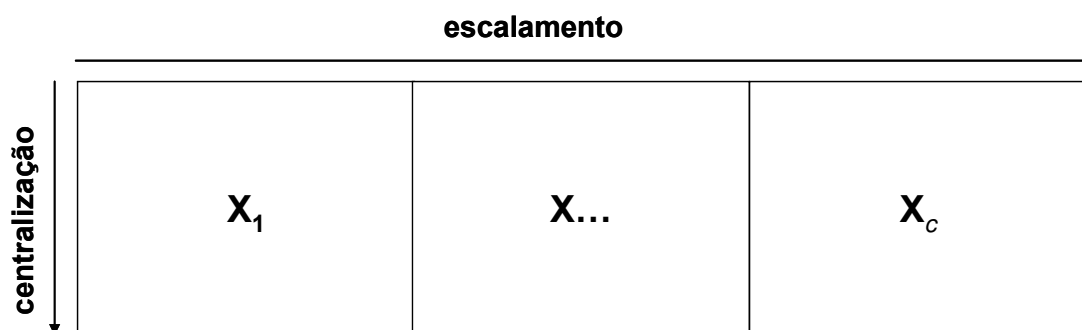
$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1J}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{I1}\mathbf{B} & \cdots & a_{IJ}\mathbf{B} \end{bmatrix} \quad (5.5)$$

4. Estime  $\mathbf{B}$  e  $\mathbf{C}$  de maneira análoga, desdobrando  $\underline{\mathbf{X}}$  em  $\mathbf{X} \in \mathfrak{R}^{K \times NC}$  e fazendo  $\mathbf{Z} = \mathbf{A} \otimes \mathbf{C}$  para estimar os *loadings*  $\mathbf{B}$ ; e desdobrando  $\underline{\mathbf{X}}$  em  $\mathbf{X} \in \mathfrak{R}^{C \times KN}$  e  $\mathbf{Z} = \mathbf{A} \otimes \mathbf{B}$  para estimar os *loadings*  $\mathbf{C}$ ;
5. Repita os passos 2 a 4 até convergência.

Vários métodos de iniciação do método PARAFAC-ALS estão descritos na literatura. Para conjuntos de dados grandes pode-se utilizar a decomposição SVD da matriz  $\mathbf{X}$ , o método GRAM (*Generalized Rank Annihilation Method*) ou matrizes de números aleatórios. A vantagem desta última alternativa é que se o algoritmo for iniciado várias vezes a convergência dar-se-á por caminhos diferentes, evitando soluções em mínimos locais. Esta alternativa, obviamente, não é atraente para conjuntos de dados muito grandes [29-31, 35].

Dependendo da aplicação várias restrições podem ser aplicadas ao algoritmo no momento da extração dos *loadings*. Exemplos são a não-negatividade (especialmente vantajosa para dados espectrais), a ortogonalidade e a uni-modalidade. Cada restrição tem escopo definido e destina-se a aplicações específicas [1-3], mas explicações detalhadas sobre estas técnicas fogem ao escopo deste trabalho.

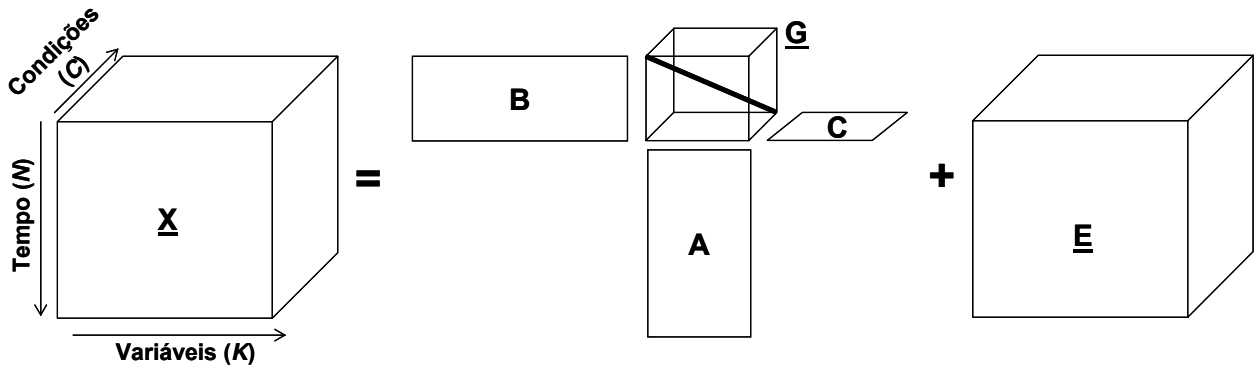
O pré-processamento de conjuntos de dados tri-lineares é mais complexo que nos métodos bi-lineares. Usualmente a centralização na média é realizada dentro do primeiro modo, enquanto que o escalamento pela variância é realizado no segundo modo (Figura 5.2). Não é possível escalar pela variância em mais de um modo porque o escalamento em um modo afeta os demais. É possível, entretanto, escalar mais de um modo para norma unitária através de processos iterativos [36].



**Figura 5.2** – Pré-processamento de um arranjo tri-linear, proposto por Smilde *et al.* [1-3]. A centralização é realizada no primeiro modo, enquanto que o escalamento é realizado no segundo modo.

A seleção do número de fatores no modelo merece atenção especial. Usualmente o número de fatores é determinado através da interpretação física ou química dos *loadings*, mas existem ferramentas de diagnóstico que auxiliam nesta tarefa. Uma dessas ferramentas, largamente empregada em métodos multi-modo, é o diagnóstico de consistência do núcleo (Corcondia – *Core Consistency Diagnostic*). O modelo PARAFAC

mostrado na Figura 5.1 pode ser alternativamente representado como um produto tensorial de matrizes [1-3]:



**Figura 5.3** – Representação do PARAFAC como um produto tensorial de matrizes.

em que  $\underline{G} \in \mathfrak{R}^{R \times R \times R}$  é um arranjo conhecido como núcleo e, no caso do PARAFAC, é super-diagonal – isto é,  $g_{r_A r_B r_C} = 0$  se  $r_A \neq r_B \neq r_C$  e  $g_{r_A r_B r_C} = 1$  se  $r_A = r_B = r_C$ .

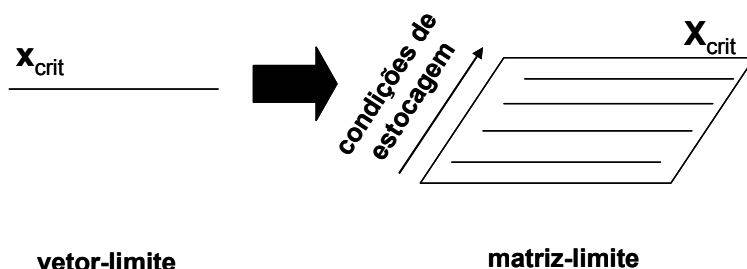
A super-diagonalidade de  $\underline{G}$  é utilizada como ferramenta de diagnóstico no PARAFAC. De fato, a *Core Diagonal Consistency Analysis* (Corcondia) consiste em checar a super-diagonalidade do núcleo: modelos bem ajustados devem apresentar  $g_{r_A r_B r_C} = 0$  se  $r_A \neq r_B \neq r_C$  e  $g_{r_A r_B r_C} = 1$  se  $r_A = r_B = r_C$ . Caso contrário, modelos com números diferentes de fatores devem ser construídos.

## 4 – Experimental

Os mesmos conjuntos de dados expostos nos Capítulos III e IV foram utilizados para comparação entre os métodos MASLT e PARAFAC e, portanto, os procedimentos experimentais para a aquisição de dados de cada aplicação estão descritos nos capítulos correspondentes.

Para os dados de espectroscopia NIR modelos PARAFAC foram construídos sem aplicação de nenhum pré-processamento e após pré-processamento por centralização na média. Para os dados físico-químicos e sensoriais, referentes aos produtos derivados de tomate, utilizaram-se dois procedimentos de pré-processamento: centralização na média no modo das variáveis (B) seguido por escalamento no modo do tempo (A), de acordo

com o procedimento proposto por Smilde, Bro e Geladi (Figura 5.2) [1, 3], bem como auto-escalamento no modo das variáveis proposto pelo MASLT (vide Figura 2.6, Capítulo II). Os modelos PARAFAC foram calculados sem restrições utilizando iniciação com números aleatórios para evitar mínimos locais. O número de fatores foi selecionado utilizando-se o Corcondia e através da interpretação físico-química dos *loadings*. Uma vez que os critérios de corte não guardam relação com nenhuma condição de estocagem específica – de fato, deve existir, para um mesmo perfil-limite (ou perfil crítico), um tempo máximo de validade para cada condição individualmente – para cada conjunto de dados, o vetor com os critérios de corte foi replicado para todas as condições de estocagem de modo a formar a matriz crítica ( $\mathbf{X}_{crit}$ , Figura 5.4). Os *loadings* críticos do modo temporal determinados para cada fator do modelo PARAFAC foram então determinados utilizando a equação 5.4, substituindo  $\mathbf{X}$  por  $\mathbf{X}_{crit}$ . Para os dados físico-químicos, os critérios de corte para cada propriedade são fornecidos na Tabela 3.1 do Capítulo III, enquanto que o espectro-limite do conjunto de dados espectroscópicos foi determinado, como descrito no Capítulo IV, através de avaliação sensorial (vide Tabela 4.2, Capítulo IV).



**Figura 5.4** – Construção de nova amostra e determinação dos *loadings* temporais correspondentes para determinação do prazo de validade utilizando os modelos PARAFAC.

O prazo de validade da condição de estocagem acelerada foi então projetado na temperatura de estocagem ambiente utilizando os fatores de aceleração obtidos através das razões entre os *loadings* das condições de estocagem (C), de acordo com a expressão 1.62 (vide Capítulo I).

Para comparar a quantidade de informação nos modelos PARAFAC e MASLT, utilizou-se a somatória quadrática dos resíduos, conforme expressão:

$$SS_E = \sum (e_{nkc})^2 \quad (5.6)$$

em que  $SS_E$  é a somatória quadrática dos resíduos e  $e_{nkc}$  é um elemento do arranjo  $\underline{E}$ .

Os cálculos foram realizados utilizando-se Matlab v.7.1 com rotinas desenvolvidas pelo autor, com exceção dos algoritmos para determinação dos modelos PARAFAC, onde utilizaram-se rotinas desenvolvidas por Andersson e Bro [37].

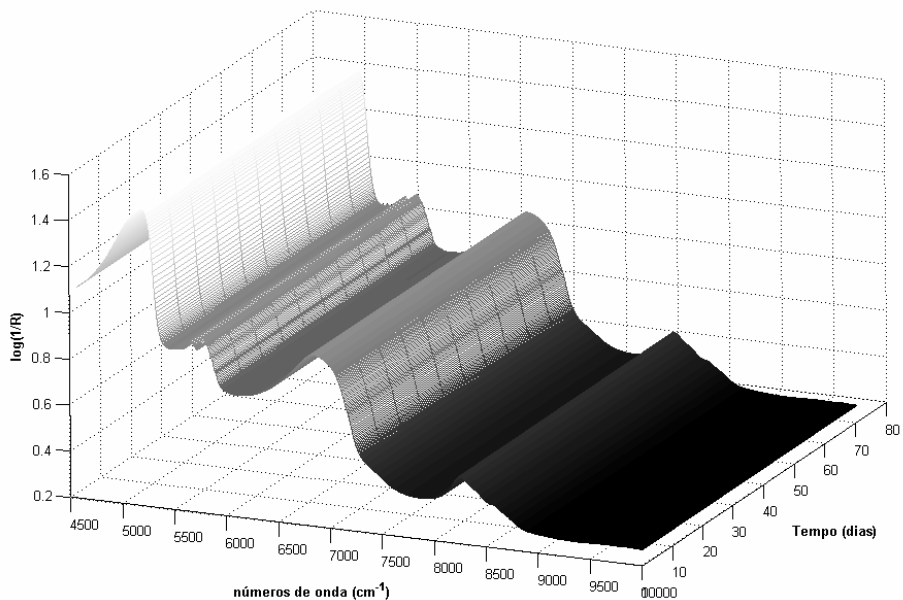
## 5 – Resultados e Discussão

### 5.1 – Dados de Espectroscopia NIR

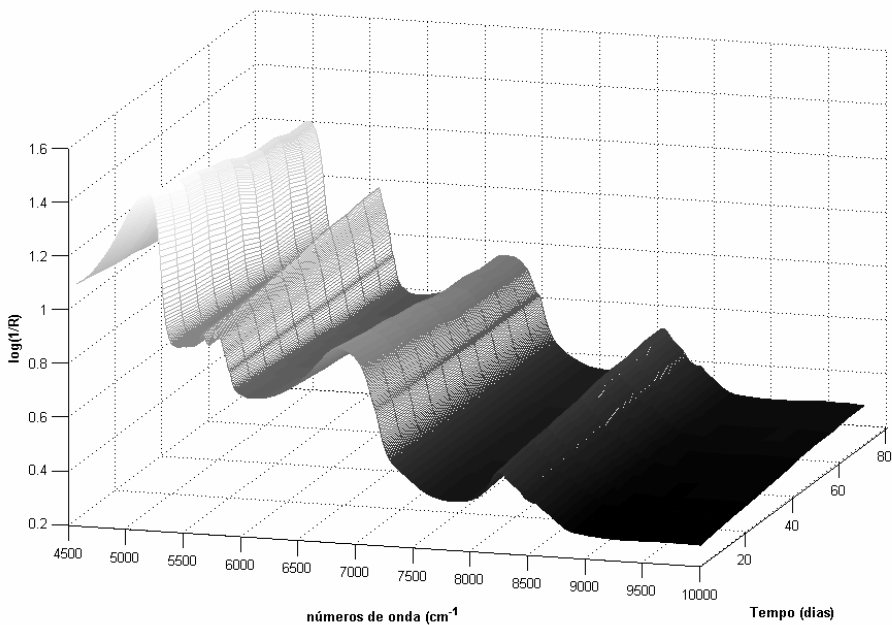
A Figura 5.5 mostra gráficos em três dimensões dos espectros NIR de loções corporais coletados a 25°C e 45°C. Conforme observado no Capítulo IV através dos gráficos em duas dimensões, nota-se que os espectros das amostras estocadas à temperatura ambiente apresentaram pouca variação temporal, enquanto que alterações significativas foram observadas nas bandas localizadas em aproximadamente 5000, 5500 e 7000  $\text{cm}^{-1}$  nos espectros dos produtos estocadas a 45 °C. É possível também observar claramente a alteração das proporções das bandas de vibração O-H e C-C em 5000 e 5500  $\text{cm}^{-1}$ , respectivamente.

Vários modelos PARAFAC foram construídos utilizando-se o cubo de dados construído com os espectros de infravermelho próximo, variando principalmente o número de fatores. A seleção do número de fatores do modelo foi realizada utilizando-se gráficos Corcondia (Figura 5.6) e através da interpretação dos *loadings*. Determinou-se que dois fatores eram suficientes para construção do modelo PARAFAC uma vez que a utilização de 1 fator, apesar de fornecer *loadings* com estrutura relacionada com o tempo e *loadings* no modo espectral com formato típico do espectro médio, não resultou em *loadings* relativos às condições de estocagem coerentes com as diferentes temperaturas. O modelo PARAFAC com 3 fatores, além de fornecer *loadings* sem significado físico, resultou em Corcondia totalmente não-diagonal (note a escala do eixo y na Figura 5.6b).

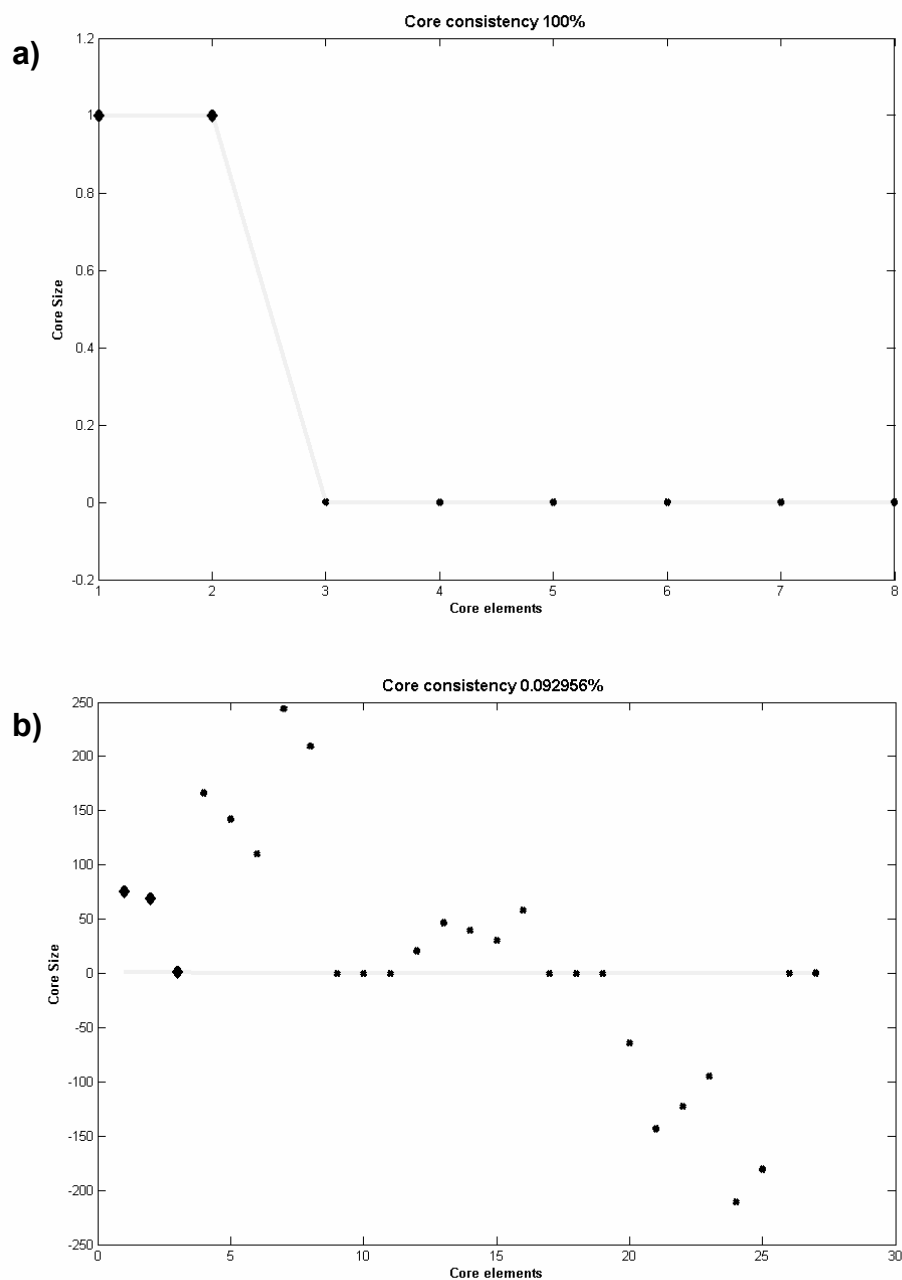
a)



b)

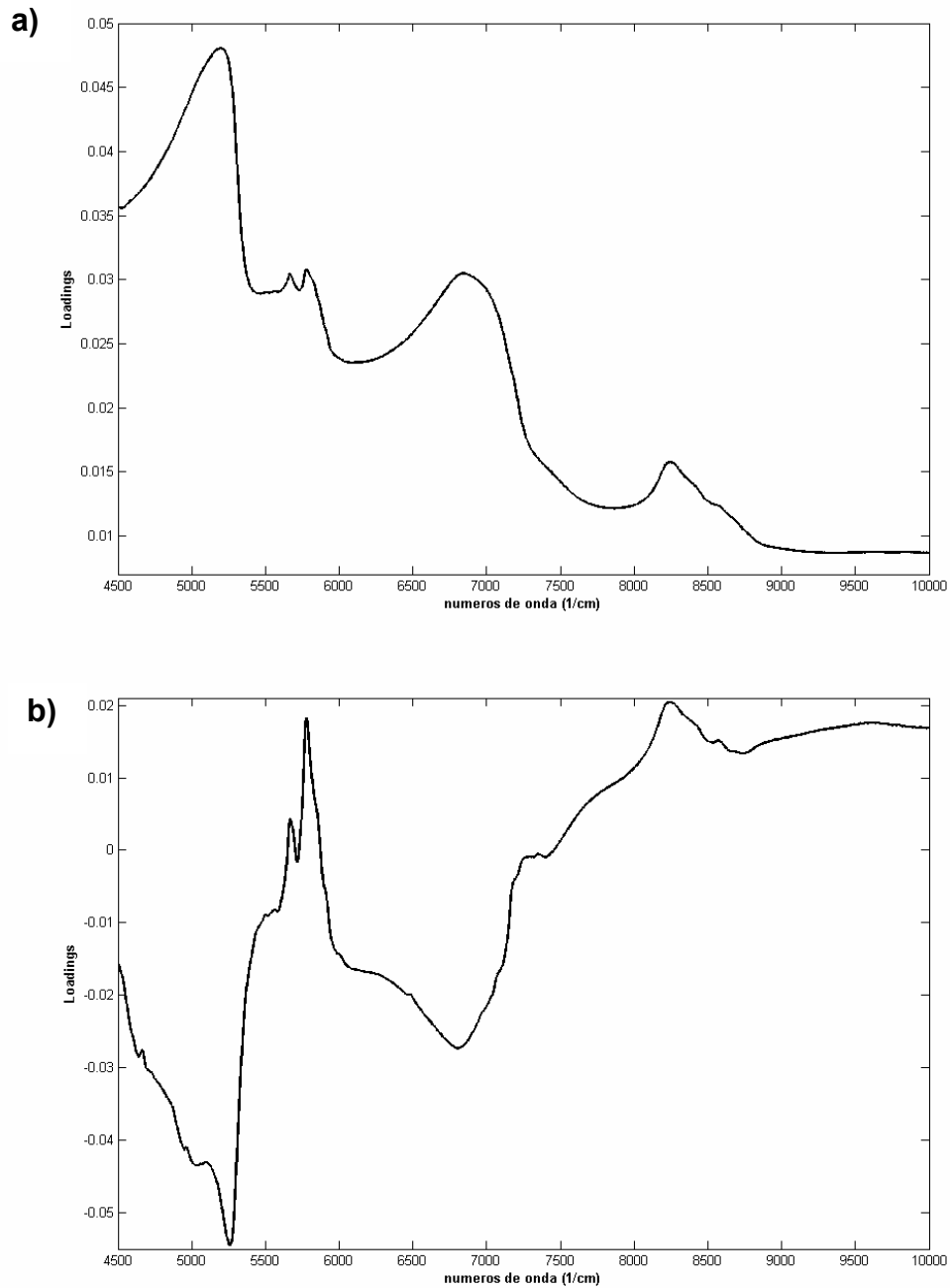


**Figura 5.5** – Representação tridimensional dos espectros NIR em função do tempo para amostras estocadas em a) 25°C e b) 45°C.



**Figura 5.6** – Gráficos Corcondia para modelos PARAFAC com a) 2 fatores e b) 3 fatores, para o conjunto de dados de espectros NIR. Modelos com ajuste satisfatório devem idealmente apresentar valores da super-diagonal (losangos) iguais à 1 e valores fora da diagonal (pontos) iguais a zero.

A Figura 5.7 mostra os gráficos de *loadings* para o modo espectral para um modelo PARAFAC com dois fatores. A semelhança com os *loadings* obtidos com o MASLT (vide Figura 4.4, Capítulo IV) foi considerável, evidenciando que os diferentes métodos de compressão de dados (u-PCA e PARAFAC) modelaram, como esperado, os mesmos fenômenos.

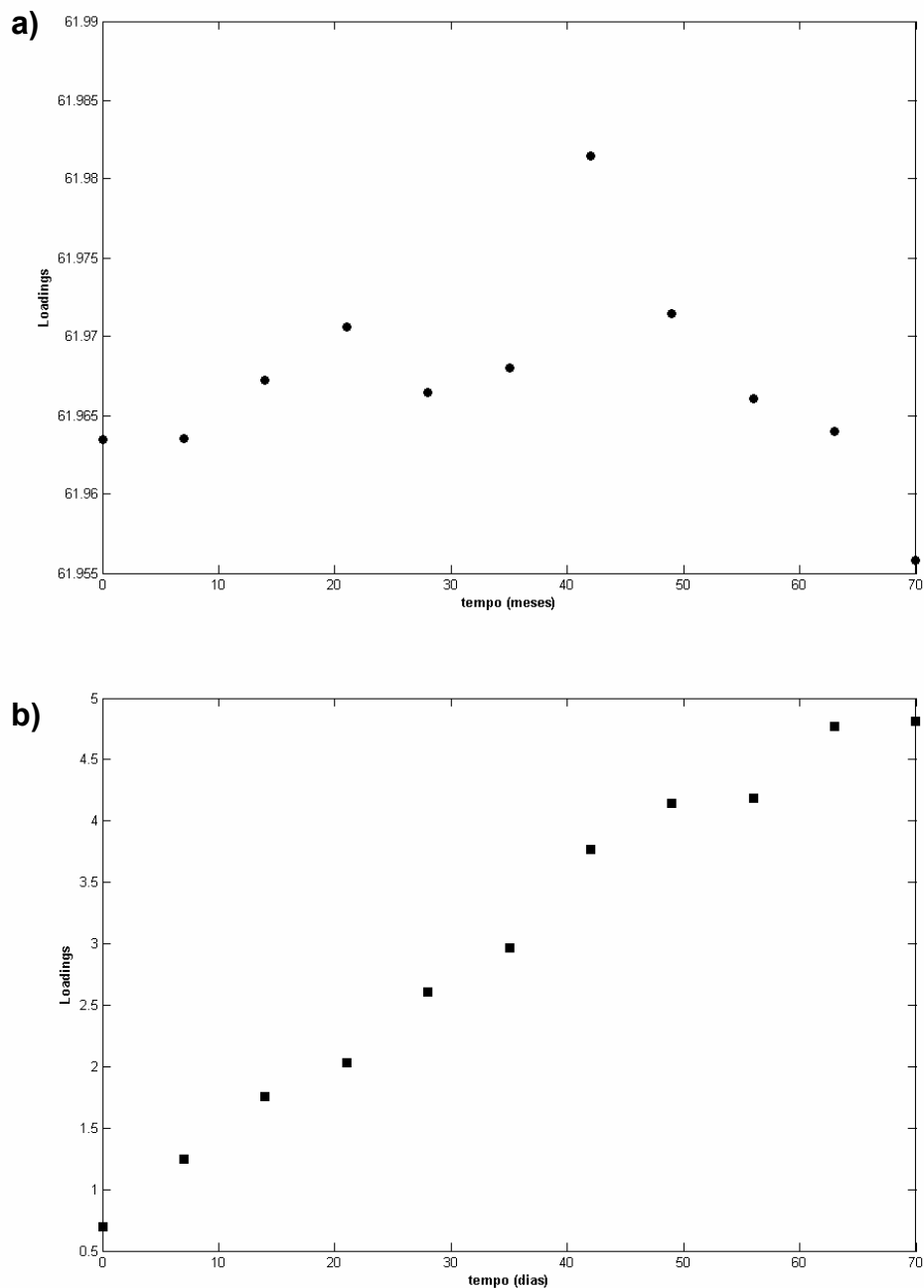


**Figura 5.7** – Gráficos de *loadings* do modo espectral (B) para o conjunto de dados de espectros NIR: a) fator 1 e b) fator 2.

Os *loadings* do modo temporal (Figura 5.8), entretanto, não concordaram com aqueles obtidos através do MASLT (vide Figura 4.4, Capítulo IV). Nota-se que o primeiro fator modelou fenômenos não relativos ao tempo, possivelmente porque foi majoritariamente influenciado pela a projeção dos espectros de todas as condições de estocagem no espectro médio, uma vez que nenhum pré-processamento foi realizado. Sendo assim,



fenômenos de linha de base que, como visto no Capítulo IV, estão relacionados com a alteração do tamanho das partículas da emulsão em função do tempo, não foram explicitamente modelados pelo modelo PARAFAC: ou ficaram ocultos sob outros fenômenos (como a projeção no espectro médio) descritos pelo fator 1, ou sob a alteração das proporções das bandas de vibração C-C e O-H no fator 2.



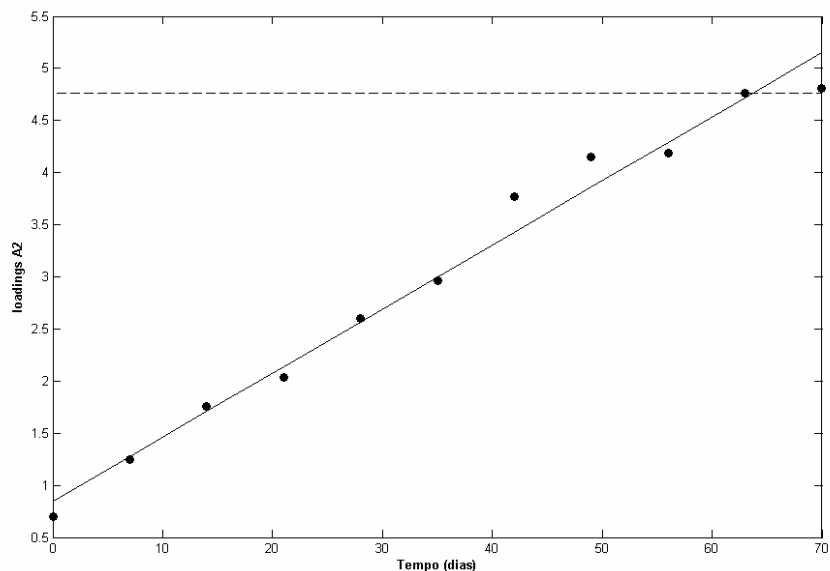
**Figura 5.8** – Gráfico de *loadings* do modo temporal (A) para os dados espectroscópicos: a) fator 1 e b) fator 2.

Os *loadings* do modo das condições de estocagem obtidos com o método PARAFAC (Tabela 5.1) são de difícil interpretação quando comparados com o MASLT, onde podem-se determinar constantes de velocidade de variação dos *loadings* através da utilização da teoria cinética das reações químicas.

**Tabela 5.1** – *Loadings* e parâmetros cinéticos do fator 2 para o modo das condições de estocagem (C).

Temperatura de Estocagem (°C)	<i>Loadings</i> (Fator 2)	Fatores de Aceleração $\alpha_{T,25}$	Energia de Ativação (kJ.mol <sup>-1</sup> )
8	0,0364		
25	0,0531		
35	0,1765	3,3	60 ± 15
45	0,9822	18,5	

A Tabela 4.5 (vide Capítulo IV) mostrou que as constantes de velocidade da primeira componente principal (que apresentou cinética de pseudo-primeira ordem) possuíam magnitude da ordem de  $10^{-5}$  e as da segunda PC (para a qual uma cinética de pseudo-ordem zero foi observada),  $10^{-3}$ , consideravelmente diferentes dos *loadings* das condições de estocagem obtidos através do PARAFAC. Entretanto os fatores de aceleração (proporções entre os *loadings* das condições de estocagem), bem como a energia de ativação, foram comparáveis com aqueles obtidos pelo MASLT ( $\alpha_{35,25}^m = 4,4$ ;  $\alpha_{45,25}^m = 22$  para PC2 e  $E_a^m = 70$  kJ.mol<sup>-1</sup>). Pode-se então argumentar que os *loadings* do modo das condições de estocagem do método PARAFAC são proporcionais às constantes de velocidade obtidas com o MASLT. Todavia, os *loadings* deste último são de interpretação mais direta e intuitiva porque estão suportados pela teoria cinética das reações químicas. A projeção da matriz-limite forneceu *loading* crítico correspondente à amostra estocada por 63 dias à 45 °C (Figura 5.9) e a aplicação do fator de aceleração  $\alpha_{45,25}^m = 18,5$  forneceu um prazo de validade igual a 1165 dias, ou 3 anos e 3 meses. Este prazo foi ligeiramente inferior ao determinado através do MASLT, de 3 anos e 9 meses.



**Figura 5.9** – Gráfico cinético dos *loadings* temporais para o segundo fator do modelo PARAFAC aplicado aos dados de espectroscopia NIR.

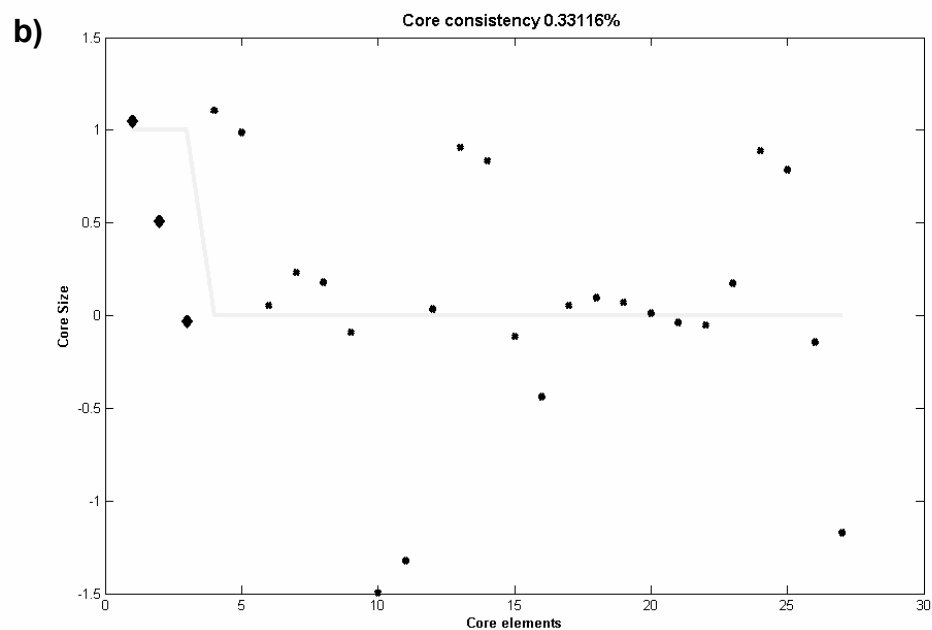
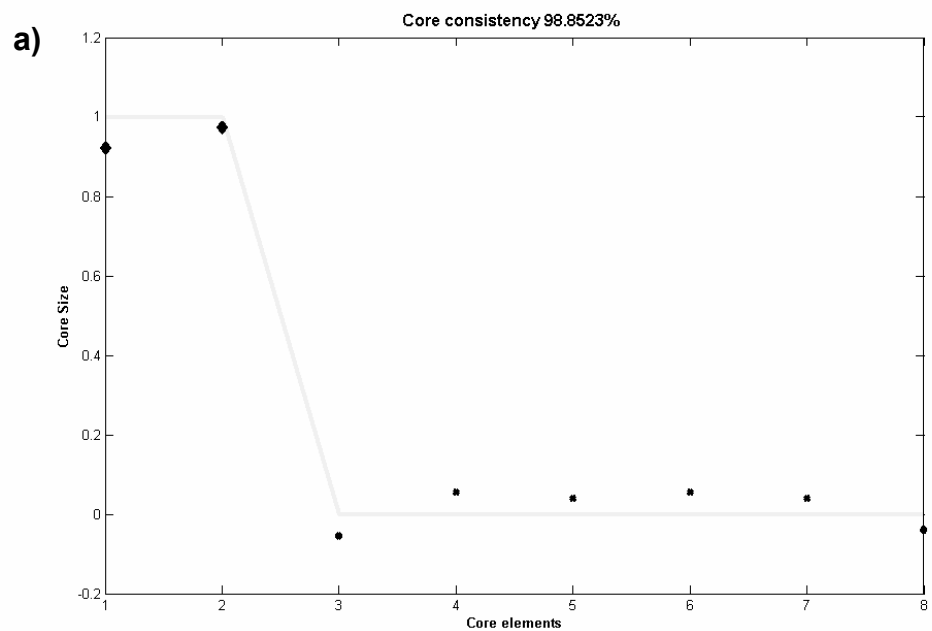
Modelos PARAFAC também foram construídos após aplicação de pré-processamento por centralização na média. Conforme observado com a aplicação do MASLT, um único fator mostrou-se suficiente para modelar os dados cinéticos. Os *loadings* dos modos temporal (A) e espectral (B) foram similares àqueles obtidos sem aplicação de pré-processamento (Figuras 5.7b e 5.8b). Entretanto, os *loadings* das condições de estocagem não apresentaram qualquer semelhança com aqueles obtidos através do modelo PARAFAC do conjunto de dados sem pré-processamentos, uma vez que valores negativos foram obtidos (Tabela 5.2). A construção de um modelo PARAFAC com restrição por não-negatividade no modo das condições de estocagem (C) também não forneceu *loadings* das condições de estocagem satisfatórios. Deste modo, não foi possível determinar o prazo de validade do produto utilizando o conjunto de dados espectrais com aplicação de pré-processamento.

**Tabela 5.2** – *Loadings* do modo das condições de estocagem para o conjunto de dados espectroscópicos após aplicação de pré-processamento por centralização na média.

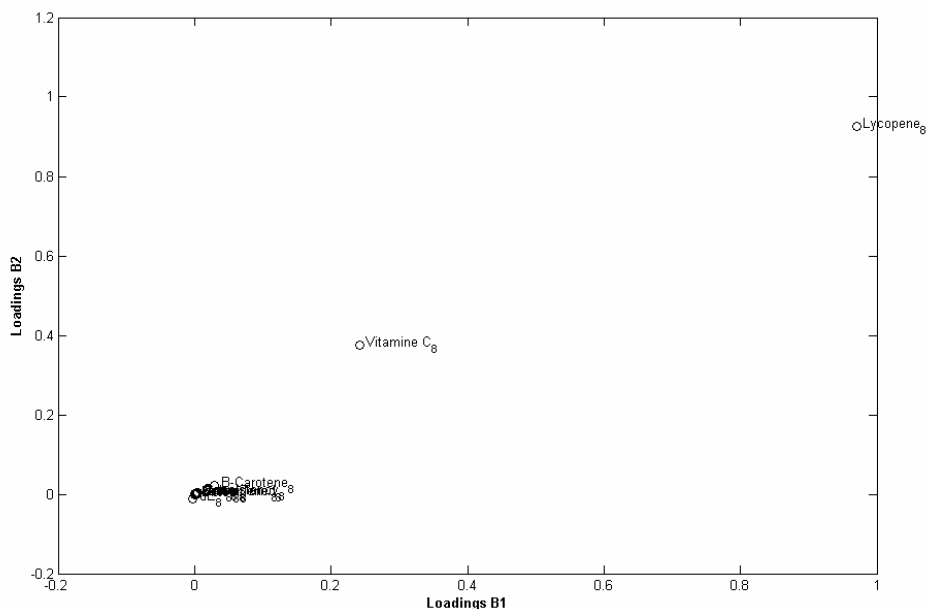
Temperatura de Estocagem (°C)	<i>Loadings</i> (C) sem restrições	<i>Loadings</i> (C) com restrições de não-negatividade
8	-0,2738	0
25	-0,2526	0
35	-0,0960	0
45	0,9231	1

## 5.2 – Dados Físico-Químicos e Sensoriais

O modelo PARAFAC com dois fatores, apesar de não ter apresentado ajuste perfeito dos elementos do núcleo (Figura 5.10), demonstrou ser o mais adequado para descrever o conjunto de dados físico-químicos e sensoriais com aplicação de pré-processamento conforme descrito por Smilde, Bro e Geladi (vide Figura 5.2) [1-3]. Entretanto, como esperado, a Figura 5.11 mostra que, como as variáveis apresentavam escalas diferentes (por exemplo, na ordem de milhar para o licopeno e unidades para os parâmetros sensoriais) e o escalamento não foi realizado no modo das variáveis, a análise foi largamente dominada pelas propriedades apresentando maior valor numérico e não pela variância intrínseca da degradação do produto. Deste modo, os *loadings* do primeiro modo não apresentaram qualquer relação com o tempo porque não modelaram variação temporal. Conclui-se que, apesar desta forma de pré-processamento ser recomendada na literatura, sua aplicabilidade para este conjunto de dados (e para quaisquer outros com a mesma configuração e para a mesma aplicação) torna-se inadequada. De fato, quaisquer procedimentos de pré-processamento devem levar em consideração não apenas a natureza do conjunto de dados bem como as características físicas e químicas do sistema e os objetivos do estudo.



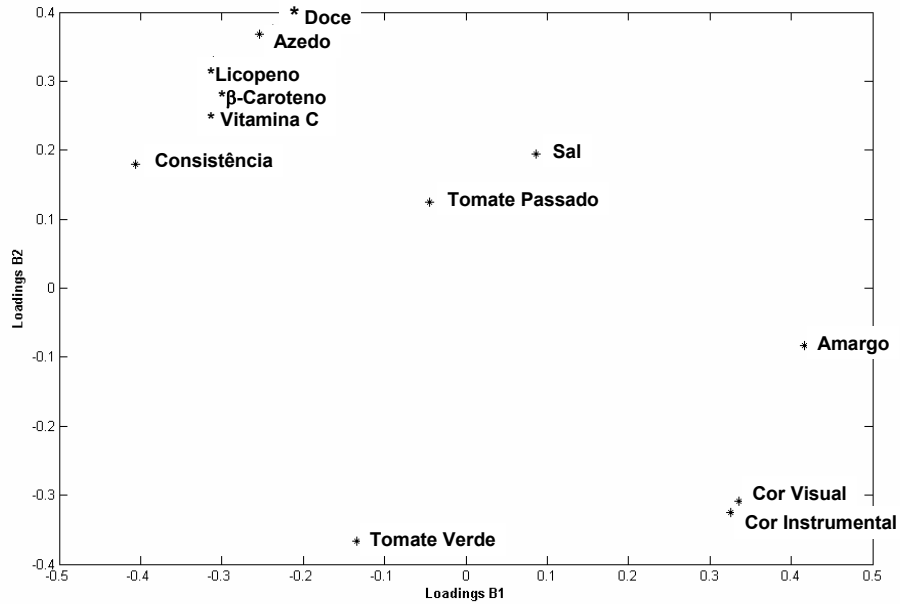
**Figura 5.10** – Gráficos de diagnóstico de consistência do núcleo para modelos PARAFAC com a) 2 fatores e b) 3 fatores, para o conjunto de dados físico-químicos e sensoriais pré-processados de acordo com o procedimento descrito por Smilde, Bro e Geladi. Modelos com ajuste satisfatório devem idealmente apresentar valores da super-diagonal (losangos) iguais à 1 e valores fora da diagonal (pontos) igual a zero.



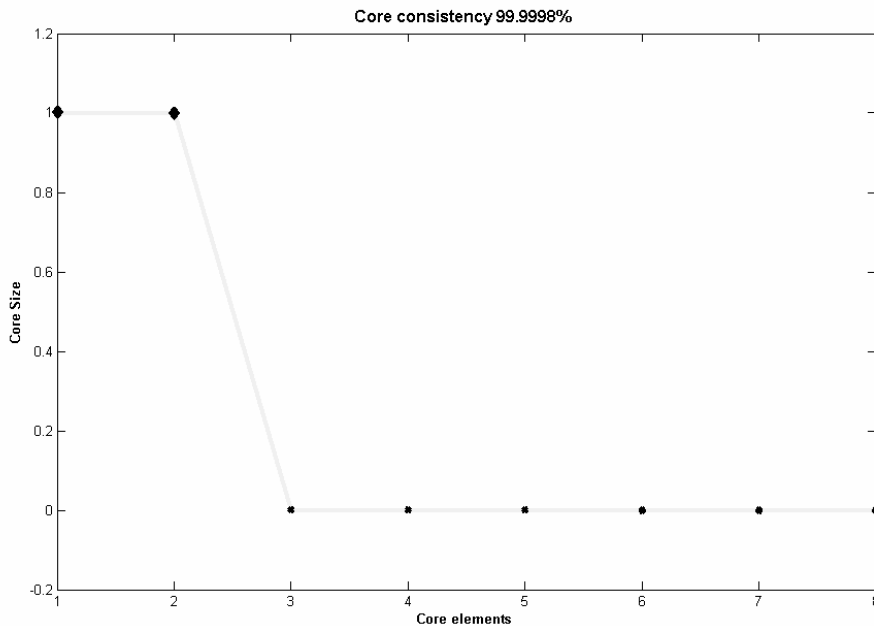
**Figura 5.11** – Gráfico de *loadings* do modo das variáveis (B) para o conjunto de dados físico-químicos e sensoriais para um modelo PARAFAC construído sem pré-processamento de dados.

Após pré-processados por auto-escalamento no modo temporal (vide Figura 3.4, Capítulo III), os *loadings* do modo das variáveis (Figura 5.12) para um modelo PARAFAC com 2 fatores apresentaram comportamento bastante similar àquele observado através do MASLT. Nota-se que as correlações entre as variáveis, determinadas pelo posicionamento no gráfico de *loadings*, foram muito similares àquelas observadas através do MASLT e, deste modo, as conclusões apresentadas na seção 3.4.2 aplicam-se também ao modelo PARAFAC.

A Figura 5.13 mostra também que este pré-processamento, avaliado através do diagnóstico da consistência do núcleo, forneceu um modelo PARAFAC com melhor ajuste que aquele obtido com pré-processamento por escalamento no segundo modo.

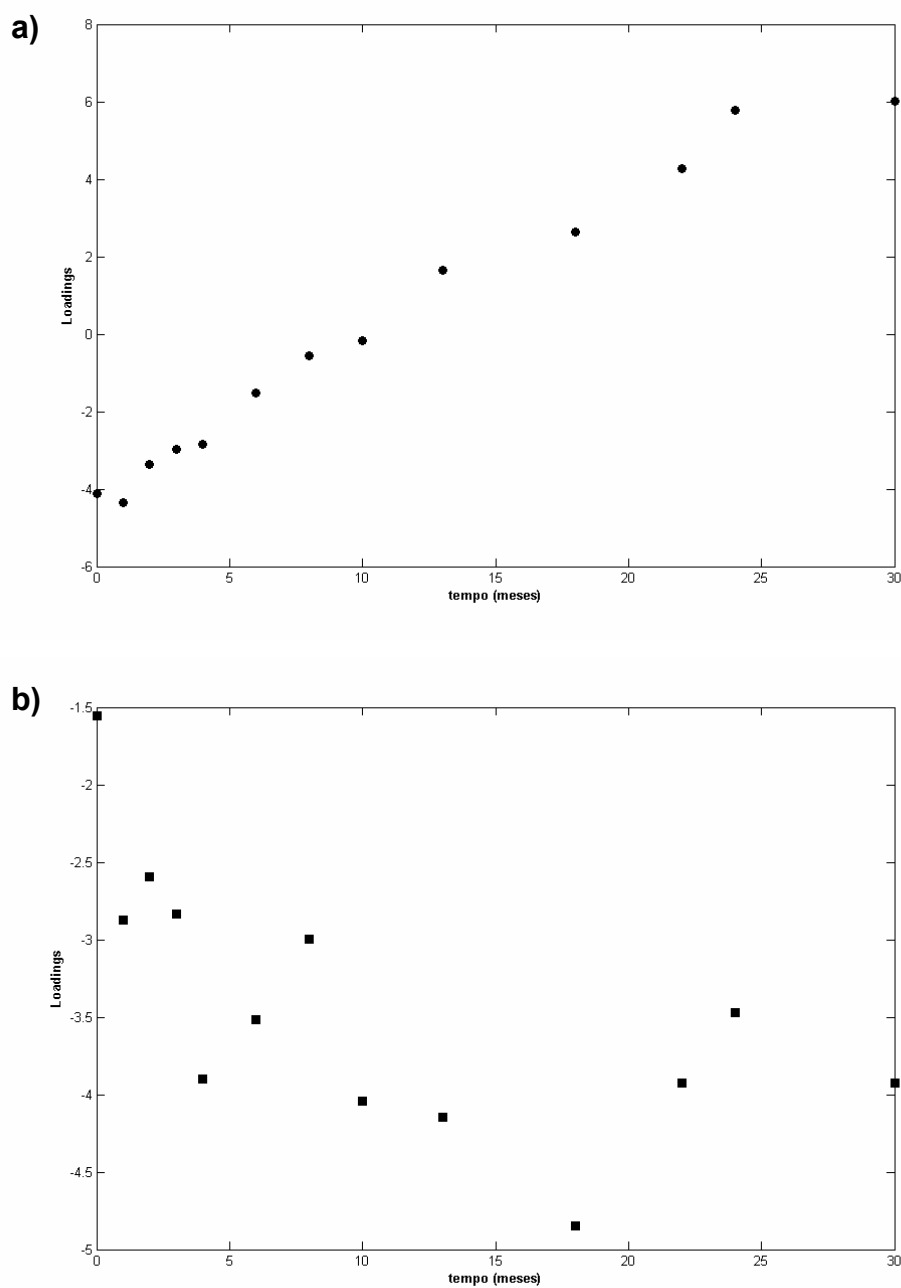


**Figura 5.12** - Gráfico de *loadings* do modo das variáveis (B) para o conjunto de dados físico-químicos e sensoriais para um modelo PARAFAC construído com pré-processamento por auto-escalamento conforme descrito no Capítulo II.



**Figura 5.13** – Gráfico de diagnóstico de consistência do núcleo para 2 fatores para o conjunto de dados físico-químicos e sensoriais com pré-processamento por auto-escalamento. Modelos com ajuste satisfatório devem idealmente apresentar valores da super-diagonal (losangos) iguais à 1 e valores fora da diagonal (pontos) igual a zero.

A Figura 5.14 mostra que os *loadings* temporais também apresentaram comportamento similar àquele obtido através do MASLT quando os dados foram pré-processados adequadamente (vide Figura 3.5, com ênfase para a variação dos *scores* das amostras estocadas a 35°C): o primeiro fator mostrou-se correlacionado com a degradação do produto, enquanto que o segundo modelou basicamente fontes de variação não relacionadas com o tempo.



**Figura 5.14** – Gráfico de *loadings* do modo temporal (A) para o conjunto de dados físico-químicos e sensoriais após pré-processamento conforme descrito no Capítulo II: a) fator 1 e b) fator 2.



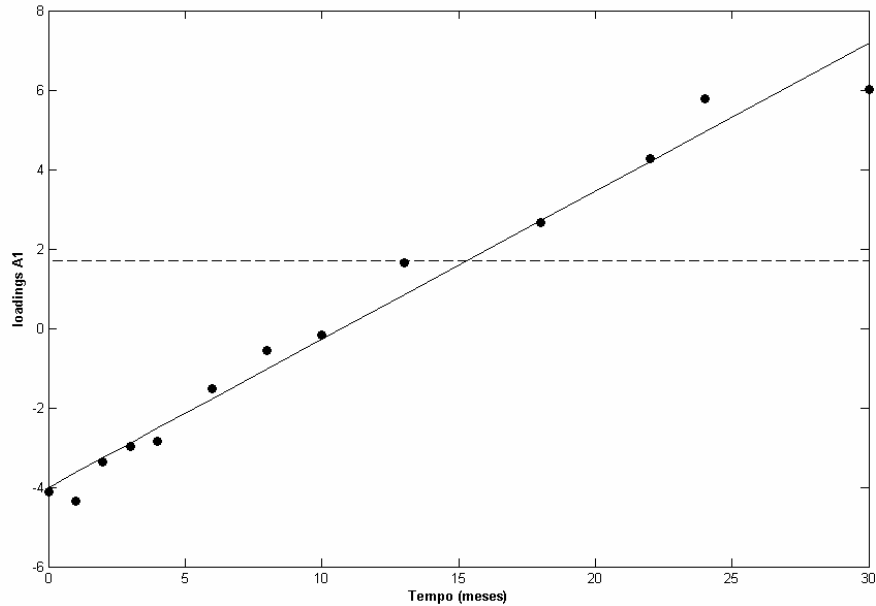
Conforme observado para o conjunto de dados utilizando espectroscopia NIR sem aplicação de pré-processamentos, os *loadings* das condições de estocagem (Tabela 5.3) não mostraram dimensões compatíveis com as constantes de velocidade obtidas através do MASLT (vide Tabela 3.5, Capítulo III). Todavia, ao contrário do observado na seção anterior deste capítulo, o fator de aceleração e a energia de ativação não foram comparáveis entre os dois métodos (PARAFAC e MASLT) para o conjunto de dados físico-químicos e sensoriais.

**Tabela 5.3** – *Loadings* e parâmetros cinéticos do modo das condições de estocagem (C).

Temperatura de Estocagem (°C)	<i>Loadings</i> (Fator 1)	Fatores de Aceleração $\alpha_{T,25}$	Energia de Ativação (kJ.mol <sup>-1</sup> )
8	0,1045		
25	0,4868		55 ± 5
35	0,8672	1,78	

Obs.: para o MASLT,  $\alpha_{25,35}^m = 2,7$  e  $Ea^m = 120$  kJ.mol<sup>-1</sup>.

O prazo de validade foi determinado através do modelo PARAFAC conforme procedimento descrito na seção 4. A Figura 5.15 mostra o gráfico cinético dos *loadings* temporais para o primeiro fator, onde a linha tracejada representa o *loading* crítico, que corresponde a um prazo de validade acelerado médio de 15 meses. Aplicando-se o fator de aceleração de 1,78, obtém-se que o prazo de validade final do produto, à condição de mercado, seria de 27 meses. Nota-se então que a divergência na estimativa final do prazo de validade do produto não foi tão díspar quanto os parâmetros cinéticos apresentados na Tabela 5.3 entre os métodos PARAFAC e MASLT.



**Figura 5.15** – Gráfico cinético dos *loadings* temporais (A) para o primeiro fator do modelo PARAFAC aplicado aos dados físico-químicos e sensoriais. A linha tracejada representa o *loading* temporal crítico.

### 5.3 – Considerações Gerais Sobre a Utilização do Método PARAFAC para Determinação do Prazo de Validade

Apesar dos *loadings* temporais e relativos às variáveis fornecerem informações similares para os métodos MASLT e PARAFAC, verificou-se que a quantidade de informação trazida pelos modelos, medida através da somatória quadrática residual (Eq. 5.6) foi consideravelmente diferente entre os métodos (Tabela 5.4). Entretanto, segundo Smilde, Bro, e Geladi [1] o principal objetivo do método PARAFAC é fornecer modelos de melhor interpretação, e não necessariamente com o melhor ajuste. Esta diferença também pode ser explicada pela característica dos modelos u-PCA de utilizar um número maior de graus de liberdade que a compressão de dados pelo método PARAFAC, conforme explanado anteriormente [1-3].

**Tabela 5.4** – Somatória quadrática dos resíduos ( $SS_E$ ) para os modelos MASLT e PARAFAC.

Conjunto de Dados	MASLT	PARAFAC
Físico-Químicos e sensoriais	98	110
Espectroscopia NIR	$1,9 \times 10^{-5}$	$4,3 \times 10^{-3}$

Observou-se que o método PARAFAC resultou consistentemente em previsões de prazos de validade inferiores às aquelas obtidas com o MASLT. Isto ocorre devido ao fato dos *loadings* temporais do PARAFAC, contidos na matriz **A**, refletirem a média da cinética de degradação do produto em todas as temperaturas de estocagem. Assim, a projeção do fator de aceleração nestes *loadings* para determinação do prazo de validade na condição de mercado sofre influência das condições mais brandas de estocagem, resultando em previsões de validade sistematicamente menores que aquelas obtidas através das técnicas convencionais e do MASLT. As previsões para os conjuntos de dados aqui demonstrados não foram mais díspares em relação ao método MASLT porque a variância da degradação nestas aplicações foi majoritariamente influenciada pelas condições de estocagem a temperaturas elevadas (vide Figuras 3.6 e 4.5 nos Capítulos III e IV, respectivamente). Deste modo, o uso da equação 1.62 para determinação do prazo de validade não é conceitualmente correto e novas estratégias devem ser desenvolvidas para que o método PARAFAC seja utilizado para determinação do prazo de validade de bens de consumo.

## 6 – Conclusões

Apesar de ambos os métodos poderem ser utilizados para estudar os processos degradativos de bens de consumo, o MASLT mostrou-se mais completo que o método PARAFAC no sentido em que permite estudar os *scores* nas diferentes condições de estocagem individualmente. Apesar de ser conceitualmente mais complexo - uma vez que une duas teorias distintas: compressão de dados e cinética química -, os *scores* e *loadings* do MASLT são de interpretação intuitiva em estudos de validade.

Ademais, para que o método PARAFAC seja aplicado de maneira preditiva em estudos acelerados de validade de produtos, faz-se necessária a pesquisa e desenvolvimento de novas premissas teóricas e/ou algoritmos matemáticos para acomodar de maneira adequada os efeitos das condições de estocagem.

## 7 – Bibliografia

1. Smilde A, Bro R, Geladi P, Multi-Way Analysis, John Wiley & Sons, Chichester, 2004.
2. Bro R, Multi-way analysis in the food industry, Ph.D. thesis, University of Amsterdam, 1998.
3. Bro R, PARAFAC: Tutorial & Applications, *Chemometrics Intel. Lab. Sys.*, **1997**, 38, 149-171.
4. Labuza TP, Shelf-Life Dating of Foods, Food & Nutrition Press, Westport, 1982.
5. Steele R (ed.), Understanding and Measuring the Shelf-Life of Food, CRC Press, Boca Raton, 2004, pp. 3-67, 128-140, 317-354
6. Kiers HAL, Hierarchical relations among three-way methods, *Psychometrika*, **1991**, 56, 449 - 470.
7. Malinowski ER, Factor Analysis in Chemistry, John Wiley & Sons, New York, 1991.
8. Jolliffe IT, Principal Component Analysis, 2<sup>nd</sup> Ed., Springer, New York, 2002.
9. Beebe KR, Pell RJ, Seasholtz MB, Chemometrics: a Practical Guide, John Wiley & Sons, New York, 1998.
10. Jolliffe IT, Principal Component Analysis, 2<sup>nd</sup> Ed., Springer, New York, 2002.
11. Niazi A, Ghasemi J, Yazdanipour A, PARAFAC decomposition of three-way kinetic-spectrophotometric spectral matrices based on phosphomolybdenum blue complex chemistry for nitrite determination in water and meat samples, *Anal. Letters*, **2005**, 38, 2377-2392.
12. Silva JCGE, Oliveira CJS, Parafac decomposition of three-way kinetic-spectrophotometric spectral matrices corresponding to mixtures of heavy metal ions, *Talanta*, **1999**, 49, 889-897.
13. Nahorniak ML, Cooper GA, Kim YC, Booksh KS, Three- and four-way parallel factor (PARAFAC) analysis of photochemically induced excitation-emission kinetic fluorescence spectra, *Analyst*, **2005**, 130, 85-93.
14. Reis MM, Biloti DN, Ferreira MMC, Pessine FBT, Teixeira GM, PARAFAC for spectral curve resolution: A case study using total luminescence in human dental tartar, *Appl. Spect.*, **2001**, 55, 847-851.
15. Bro R, Jakobsen M, Exploring complex interactions in designed data using GEMANOVA. Color changes in fresh beef during storage, *J. Chemom.*, **2002**, 16, 294-304.
16. Bosco M, Larrechi MS, Rapid and quantitative evaluation of the effect of process variables on the kinetics of photocatalytic degradation of phenol using experimental design techniques and parallel factor (PARAFAC) analysis, *Anal. Bioanal. Chem.*, **2008**, 390, 1203-1207.
17. Tan YX, Jiang JH, Wu HL, Cui H, Yu RQ, Resolution of kinetic system of simultaneous degradations of chlorophyll a and b by PARAFAC, *Analytica Chimica Acta*, **2000**, 412, 195–202.

18. Let MB, Jacobsen C, Meyer AS, Effects of fish oil type, lipid antioxidants and presence of rapeseed oil on oxidative flavour stability of fish oil enriched milk, *Eur. J. Lipid Sci. Tech.*, **2004**, 106, 170-182.
19. Christensen J, Povlsen VT, Sorensen J, Application of fluorescence spectroscopy and chemometrics in the evaluation of processed cheese during storage, *J. Dairy Sci.*, **2003**, 86, 1101-1107.
20. Harshman RA, Berenbaum SA, Basic concepts underlying the PARAFAC-CANDECOMP three-way factor analysis model and its application to longitudinal data, in "Present and past in middle life". (Eds. D.H. Eichorn, J.A. Clausen, N. Haan, M.P. Honzik and P.H. Mussen), Academic Press, NY, 1981, 435-459.
21. Burdick DS, An introduction to tensor products with applications to multiway data analysis. *Chemom. Intell. Lab. Syst.*, **1995**, 28, 229 - 237.
22. Wold S, Esbensen K, Geladi P, Principal component analysis: a tutorial, *Chemom. Intel. Lab. Sys.*, **1987**, 2, 37-52.
23. Kowalki BR, Chemometrics: Mathematics and Statistics in Chemistry, D. Reidel, Dordrecht, 1984
24. Harshman RA, Foundations of the PARAFAC procedure: Model and conditions for an 'explanatory' multi-mode factor analysis, *UCLA Working Papers in phonetics*, **1970**, 16, 1 - 84.
25. Carroll JD, Chang J, Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition, *Psychometrika*, **1970**, 35, 283 – 319.
26. Scarminio I, Kubista M, Analysis of correlated spectral data, *Anal. Chem.*, **1993**, 65, 409 - 416.
27. Sarabia L, Ortiz MC, Leardi R, Drava G, A program for non-orthogonal factor analysis, *Trends Anal. Chem.*, **1993**, 12, 226 - 230.
28. Faber NM, Buydens LMC, Kateman G, Generalized rank annihilation method. I: derivation of eigenvalue problems. *J. Chemom.*, **1994**, 8, 147 - 154.
29. Faber NM, Bro R, Hopke PK, Recent developments in CANDECOMP/PARAFAC algorithms: a critical review, *Chemometrics Intel. Lab. Sys.*, **2003**, 65, 119-137.
30. Tomasi G, Bro R, A comparison of algorithms for fitting the PARAFAC model, *Computational Stat Data Anal.*, **2006**, 50, 1700-1734.
31. Kiers H, Krijnen WP, An efficient algorithm for PARAFAC of 3-way data with large numbers of observations, *Psychometrika*, **1991**, 56, 147-152.
32. Andersson CA, Bro R, The N-way toolbox for Matlab, *Chemometrics Intel. Lab. Sys.*, **2000**, 52, 1-4.
33. Strang G, Linear Algebra and its Applications, 3<sup>rd</sup> Ed., Harcourt Brace Jovanovich, Pub, San Diego, 1988.
34. Noble B, Daniel JW, Applied Linear Algebra, Prentice Hall, Englewood Cliffs, 1988.
35. Golub G, van Loan CF, Matrix Computations, 2<sup>nd</sup> Ed., Johns Hopkins Press, London 1989.

36. J.M.F. ten Berge, Convergence of PARAFAC preprocessing procedures and the Deming-Stephan method of iterative proportional fitting, In "Multiway data analyses", (Eds. R. Coppi and S. Bolasco), Elsevier Science Pub., Amsterdam, 1989.
37. Bro R, Andresson C, The N-way toolbox for Matlab, disponível em <http://www.models.kvl.dk/source/nwaytoolbox/index.asp>, 15/02/2009.

## **Capítulo VI**

### **Desenvolvimento de Métodos Rápidos e Não-Destrutivos de Análise de Produtos de Tomate para Determinação do Prazo de Validade**

## 1 – Introdução [1]

O tomate ocupa uma posição importante no mercado agrícola global. É o segundo vegetal mais consumido mundialmente, o primeiro lugar sendo ocupado pela batata [2]. Em 2006 mais de 30 milhões de toneladas de tomate *in natura* foram produzidas [3], das quais aproximadamente 85% foram utilizadas para fabricação de produtos industrializados como extrato e molho de tomate [4].

O tomate é composto principalmente por água, sólidos (solúveis e insolúveis), ácidos orgânicos (predominantemente cítrico) e micronutrientes como carotenóides e vitaminas A e C. Sólidos solúveis são constituídos principalmente por açúcares e sais e são geralmente expressos como °Brix. Sólidos insolúveis são majoritariamente constituídos por fibras como celulose e pectina. Os sólidos totais correspondem à soma dos sólidos solúveis e insolúveis e estão presentes nos frutos na proporção de 4,5% a 8,5% (excluindo-se sementes e pele), dependendo da variedade, solo e condições climáticas durante o cultivo [2, 5, 6].

Os sólidos do tomate constituem um parâmetro de extrema relevância na indústria porque determinam o rendimento fabril: quanto mais alto seu valor, menor a quantidade de frutos necessários para manufatura dos produtos industrializados. Por outro lado, a proporção entre açúcares e ácidos que compõe os sólidos solúveis determina o sabor dos produtos finais, definindo assim a preferência dos consumidores [2,5].

Apesar da relevância destes atributos para a indústria de alimentos, sua quantificação ainda é demorada, além de demandar mão-de-obra especializada. Por exemplo, os sólidos totais são quantificados por secagem em estufa sob vácuo, um procedimento que demanda aproximadamente 4 horas. A acidez total dos produtos atomatados é usualmente quantificada por titulação ácido-base usando fenolftaleína como indicador, mas a determinação acurada do ponto de equivalência é difícil devido à cor avermelhada do produto. Açúcares são individualmente determinados por cromatografia líquida (HPLC), uma técnica que demanda uma considerável quantidade de solventes orgânicos. Além disso, a análise é demorada e requer mão-de-obra especializada não somente para operar a instrumentação analítica mas, principalmente, porque um procedimento robusto de preparo de amostra faz-se necessário [2, 5-8].

Para resolver problemas analíticos como os acima expostos, os químicos aliaram técnicas rápidas e não-destrutivas, com ênfase à espectroscopia no infravermelho próximo (NIR)



aos métodos quimiométricos de calibração. A química de alimentos foi, de fato, um dos campos mais beneficiados pelo uso da espectroscopia NIR para determinação de uma miríade de propriedades em matrizes complexas [9-14]. A primeira publicação para parâmetros de qualidade de tomates deve-se a Hong e Tsou [15]. Após, Goula e Adamopoulos determinaram umidade, açúcares, acidez total, sal e proteína em produtos de tomate [16], enquanto Jha e Matsuoka desenvolveram uma calibração para a razão ácido/Brix para vários sucos de tomate [17]. Mais recentemente, Pedro e Ferreira [18] desenvolveram um modelo de calibração PLS para determinar sólidos (totais e solúveis), licopeno e  $\beta$ -caroteno, micronutrientes importantes em produtos de tomate [19-21], utilizando a espectroscopia NIR.

O método de calibração mais popular em quimiometria e amplamente aliado à espectroscopia NIR é a regressão por quadrados mínimos parciais, ou *Partial Least Squares* (PLS) [22, 23]:

$$\mathbf{X} = \mathbf{T}\mathbf{L}^T + \mathbf{E} \quad (6.1)$$

$$\mathbf{Y} = \mathbf{T}\mathbf{Q}^T + \mathbf{F} \quad (6.2)$$

onde  $\mathbf{T}$  é a matriz de *scores*;  $\mathbf{L}$  e  $\mathbf{Q}$  são, respectivamente, as matrizes de *loadings* de  $\mathbf{X}$  e  $\mathbf{Y}$ ; e  $\mathbf{E}$  e  $\mathbf{F}$  são as matrizes de resíduos correspondentes. O número de fatores ( $R$ ) na regressão PLS é geralmente muito menor que o número de amostras  $N$ , assim como usualmente  $N \ll K$  [22, 23].

A principal vantagem do método PLS é sua capacidade em lidar com dados que apresentam correlação significativa (o que geralmente ocorre em espectroscopia), porque a calibração é realizada utilizando-se variáveis latentes ou fatores (contidos em  $\mathbf{T}$ ,  $\mathbf{L}$  e  $\mathbf{Q}$ ), que são definidos como combinações lineares das variáveis originais. Esses fatores têm a importante característica de serem geometricamente ortogonais entre si e, além disso, são extraídos de modo a apresentar a máxima covariância com a propriedade sendo calibrada [22, 23].

Duas rotas distintas de regressão PLS podem ser utilizadas quando existem muitas variáveis dependentes disponíveis para calibração: cada propriedade pode ser calibrada individualmente utilizando-se o algoritmo conhecido como PLS1 – e, assim, a matriz  $\mathbf{Y}$  é

substituída pelo vetor  $\mathbf{y}$  em 6.2 - ou, alternativamente, todas as propriedades podem ser calibradas em um único procedimento conhecido como PLS2. Na regressão PLS1 um conjunto diferente de fatores é extraído para cada modelo de calibração enquanto que na regressão PLS2 um único conjunto de fatores é extraído para todas as propriedades [22, 23].

Enquanto os fatores na PLS2 são obtidos em uma única execução da rotina de cálculo, a PLS1 geralmente fornece modelos mais precisos [22, 23]. Todavia, a regressão PLS2 é especialmente vantajosa quando propriedades cujos dados de calibração são imprecisos apresentam correlação significativa com outras que podem ser determinadas com alta precisão. Para estas propriedades, cujos métodos de referência fornecem resultados imprecisos, a PLS2 tende a produzir modelos de calibração que possuem, em geral, melhores capacidades de previsão que a PLS1. Isso ocorre porque os *loadings* de  $\mathbf{X}$  são estabilizados na regressão PLS2, reduzindo os efeitos dos desvios aleatórios das variáveis imprecisas [22-24].

## 2 - Objetivo

O objetivo deste capítulo é apresentar o desenvolvimento de um procedimento analítico para determinação simultânea e não-destrutiva de importantes parâmetros de produtos industrializados de tomate, visando acelerar a aquisição de dados para determinação de seu prazo de validade.

## 3 – Parte Experimental

### 3.1 – Métodos Físico-Químicos e Espectroscópicos

Quarenta e duas amostras de concentrado de tomate com teores de sólidos totais variando entre 6,9 e 35,9% foram coletadas na América Latina (Brasil e Argentina), Estados Unidos e Europa (Holanda, Espanha, Itália e Grécia).

Sólidos totais (%) foram determinados em triplicata por secagem em estufa (70°C) utilizando uma estufa Fanen EV8 (Fanen Co., São Paulo) sob vácuo (~150 mmHg de pressão absoluta) produzido por uma bomba de vácuo Edwards E2M8. Aproximadamente 3g de amostra foram pesadas em cápsulas de alumínio contendo 1g de terra diatomácea e mantidas na estufa até peso constante (~4h).

Sólidos solúveis foram determinados em duplicata utilizando-se refratômetro de bancada Abbe (Anerucab Optical), com precisão de 0,1°Brix. Acidez total (% ácido cítrico) foi quantificada através de um titulador automático Metrohm 702 equipado com um eletrodo combinado de vidro e utilizando-se solução titulante de NaOH 0,1 mol.L<sup>-1</sup>.

Açúcares (sacarose, frutose e glicose) foram determinados em duplicata injetando-se 5µL de uma solução a 5% da amostra previamente clarificada com 10mL de cada reagente de Karenz (K<sub>4</sub>Fe(CN)<sub>6</sub> 0,25mol.L<sup>-1</sup> e acetado de zinco 1mol.L<sup>-1</sup>) em um HPLC Shimadzu equipado com detector por índice de refração modelo RID-6A. A separação dos açúcares foi realizada numa coluna Shodex NH<sub>2</sub>P-50 4E acoplada a uma pré-coluna NH<sub>2</sub>P-50G. Utilizou-se como fase móvel acetonitrila:água (75:25) em regime isocrático a 1mL.min<sup>-1</sup>. Curvas analíticas de quantificação foram construídas utilizando padrões Sigma-Aldrich (S8501 para sacarose, D9434 para glicose e F9048 para frutose).

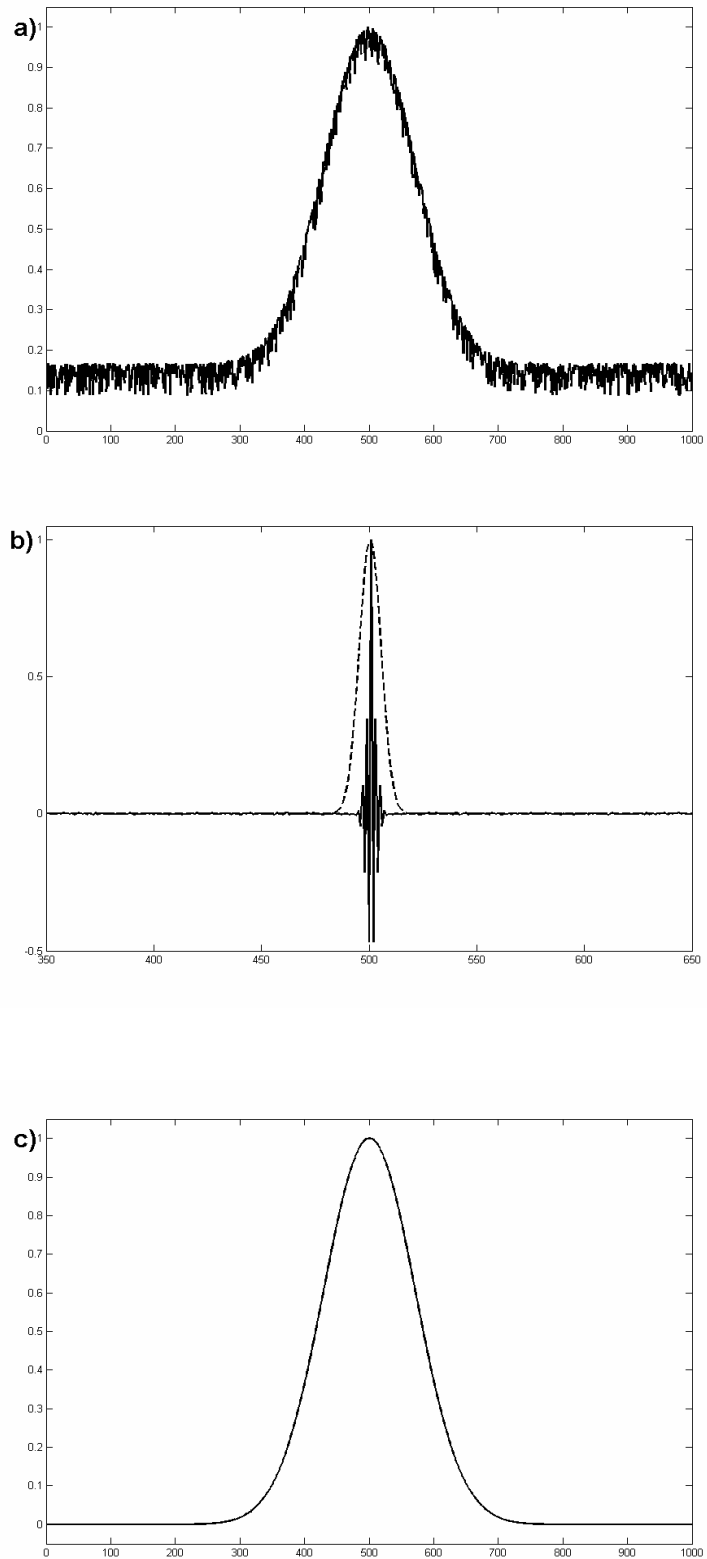
Os espectros NIR foram adquiridos utilizando-se um espectrômetro Büchi NIRLab N-200 equipado com uma célula de reflectância difusa MSC-100 com sistema de rotação da amostra. Três espectros foram coletados para cada amostra, utilizando-se 100 varreduras na região entre 4000 e 10 000 cm<sup>-1</sup> (4 cm<sup>-1</sup> de resolução), à temperatura ambiente (20 – 25°C).

### 3.2 – Métodos Quimiométricos

Para reduzir o ruído e os desvios na linha de base dos espectros NIR aplicou-se um alisamento pela média com janelas de diferentes amplitudes (números de pontos) seguido pela correção multiplicativa de sinais (MSC – *Multiplicative Signal Correction*) [22, 23]. Alternativamente, filtros de Fourier com funções Gaussianas de convolução (também conhecidas como funções de apodização) de várias larguras ( $w$  na Eq. 6.3) foram aplicadas (Figura 6.1), seguidos por MSC ou por primeiras- ou segundas-derivadas de acordo com o algoritmo descrito por Savitzky e Golay [22, 24, 25].

$$f(x) = e^{-(x-\bar{x})^2/w^2} \quad (6.3)$$

Os dois procedimentos de regressão multivariada PLS1 e PLS2 foram utilizados para construir modelos de calibração, sendo que a matriz  $\mathbf{Y}$  de variáveis dependentes foi auto-escalada nas calibrações onde a regressão PLS2 foi utilizada. Realizou-se a validação cruzada retirando-se uma amostra a cada iteração (*leave-one-out*) para definição do número ótimo de fatores a se manter no modelo através do  $RMSECV_p$  (vide equação 6.8 e considerações correspondentes).



**Figura 6.1** – Alisamento por Fourier de um sinal genérico com unidades arbitrárias: (a) pico apresentando ruído; (b) uma função Gaussiana de convolução (linha descontínua) é aplicada ao interferograma (linha sólida) para remover os componentes de alta frequência que caracterizam o ruído; (c) sinal após alisamento.

A detecção de amostras anômalas foi realizada através de gráficos de *leverage* por resíduos de Student (vide Equações 6.5 e 6.6, bem como a Figura 6.2). Amostras que apresentassem valores de *leverage* acima do valor crítico (Equação 6.7) e resíduos de Student maiores que 2 foram consideradas anômalas e removidas do conjunto de calibração antes da construção de novos modelos [22-24, 26].

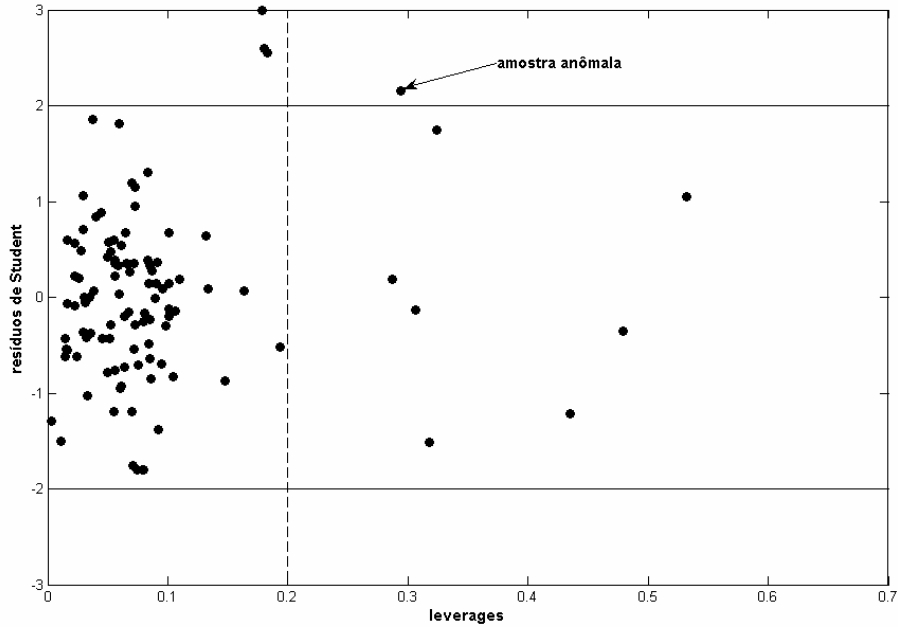
$$\mathbf{h} = \text{diag}(\mathbf{X}(\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T) \quad (6.4)$$

$$Lresc_i = \sqrt{\frac{(y_i - \hat{y}_i)^2}{(Nc - 1)(1 - h_i)}} \quad (6.5)$$

$$SR_i = \frac{(y_i - \hat{y}_i)}{Lresc_i \cdot \sqrt{1 - h_i}} \quad (6.6)$$

$$h_{\text{crit}} = \frac{3R}{Nc} \quad (6.7)$$

onde  $\mathbf{h}$  corresponde ao vetor com os valores de *leverage* de cada amostra;  $\mathbf{T}$  à matriz de *scores* para o conjunto de calibração;  $Nc$  ao número de amostras no conjunto de calibração;  $Lresc_i$  ao resíduo da amostra  $i$ , padronizado por seu valor de *leverage*;  $SR_i$  corresponde ao resíduo de Student;  $y_i$  e  $\hat{y}_i$  são, respectivamente, o valor medido e o previsto da propriedade  $y$  da  $i$ -ésima amostra;  $h_{\text{crit}}$  é o *leverage* crítico e  $R$  é o número de fatores utilizados no modelo de calibração.



**Figura 6.2** – Gráfico de *leverage* vs. resíduos de Student. As linhas horizontais representam os resíduos de Student com 95% de significância enquanto que a linha vertical tracejada representa o valor crítico de *leverage*.

Sete amostras foram separadas para validação externa enquanto que as outras 35 constituíram o conjunto de calibração. O modelo PLS2 que apresentou a melhor capacidade de previsão foi escolhido de acordo com o  $RMSEP_p$  (Equação 6.8) e com  $r_{val,p}$  (Equação 6.9) [22-24].

$$RMSEP_p = \sqrt{\frac{\sum_{j=1}^K \sum_{i=1}^{N_v} (y_{ij} - \hat{y}_{ij})^2}{K \times N_v}} \quad (6.8)$$

$$r_{val,p} = \frac{\sum_{j=1}^K \sum_{i=1}^{N_v} (y_{ij} - \hat{y}_{ij})^2}{\sqrt{s^2(\mathbf{Y}) \times s^2(\hat{\mathbf{Y}})}} \quad (6.9)$$

onde  $RMSEP_p$  é a média quadrática ponderada do erro de previsão (*pooled Root Mean Squared Error of Prediction*),  $r_{val,p}$  é o coeficiente de correlação ponderado entre os valores

previstos e os estimados pelo modelo (o sub-índice  $p$  indica média ponderada entre todas as propriedades);  $s^2(\mathbf{Y})$  e  $s^2(\hat{\mathbf{Y}})$  são, respectivamente, as variâncias dos valores medidos e estimados para todas as propriedades em estudo,  $N_v$  é o número de amostras no conjunto de validação e  $K$  o número de propriedades sendo calibradas pela regressão PLS2. O erro médio ponderado de validação cruzada, ou  $RMSECV_p$  (*Root Mean Squared Error of Cross Validation*), foi calculado de maneira análoga ao  $RMSEP_p$  substituindo-se  $N_v$  por  $N_c$  na Equação 6.8, em que  $N_c$  é o número de amostras no conjunto de calibração. Para cada propriedade individualmente, a comparação entre os desempenhos dos modelos obtidos pelas regressões PLS1 e o melhor modelo PLS2 foi realizada através do  $RMSEP$  e  $r_{val}$  convencionais, obtidos desconsiderando-se a somatória em  $j$  nas equações 6.8 e 6.9 [22-24].

## 4– Resultados e Discussão

A Figura 6.3a mostra os espectros originais coletados para amostras contendo níveis baixos (linhas contínuas) e altos (linhas tracejadas) de sólidos de tomate. Sinais ruidosos foram obtidos nas regiões compreendidas entre 4000 e 5500  $\text{cm}^{-1}$ , e entre 6300 e 7100  $\text{cm}^{-1}$ , devido à baixa intensidade de radiação infravermelha atingindo o detector (note os elevados valores de  $\log(1/R)$ ). Este efeito foi especialmente significativo nos espectros das amostras contendo baixos teores de sólidos de tomate, pois quanto menor a quantidade de material particulado no produto, maior o caminho óptico efetivo. Para remover este ruído, alisamentos pela média ou por filtros de Fourier foram aplicados.

Determinou-se que a melhor janela para o alisamento pela média possuía 50 pontos (Figura 6.3b), uma vez que forneceu modelos PLS2 com melhor habilidade de previsão que aqueles obtidos com outras janelas. O melhor modelo de calibração PLS2 utilizando alisamento por filtro de Fourier foi obtido com  $w = 17$  pontos.

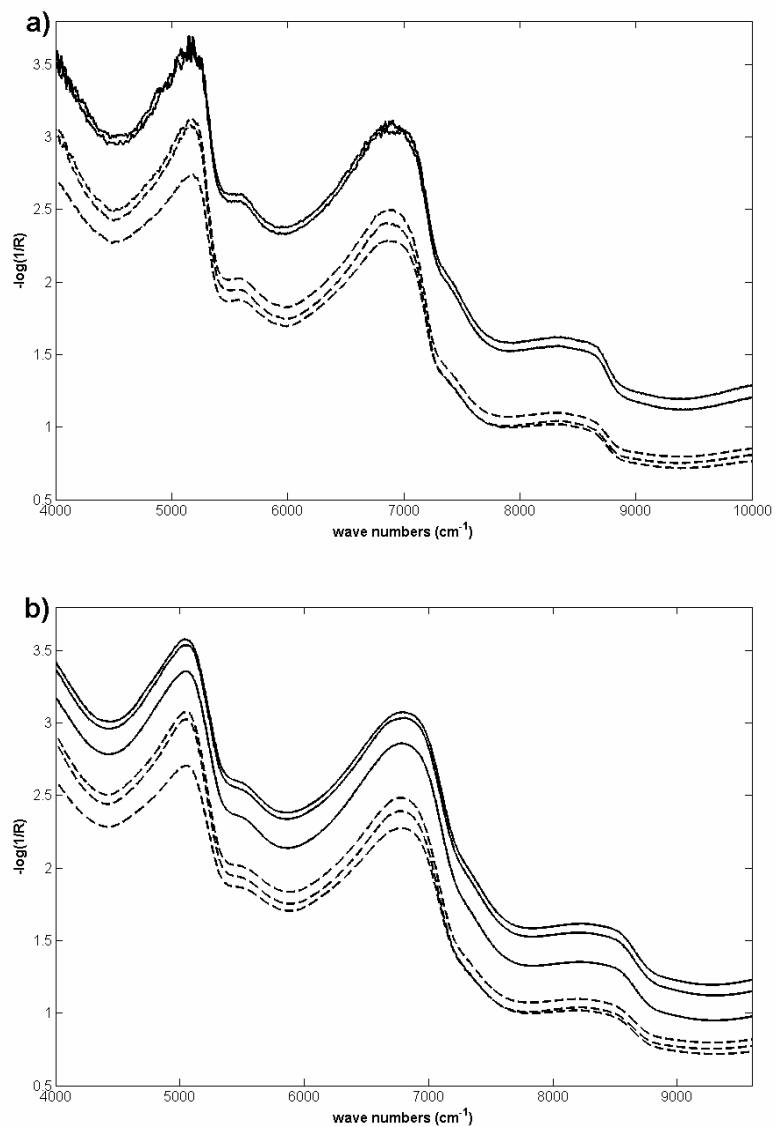
Também se observa, na Figura 6.3, a presença de desvios e inclinação nas linhas de base, características bastante comuns em espectros NIR adquiridos por reflectância difusa devido a efeitos ópticos. Em espectroscopia por transmitância, a absorção da radiação é diretamente proporcional ao caminho óptico e à concentração das espécies que absorvem em um determinado comprimento de onda. Todavia, a medida do caminho óptico em espectroscopia por reflectância difusa é geralmente imprecisa e irreprodutível devido ao



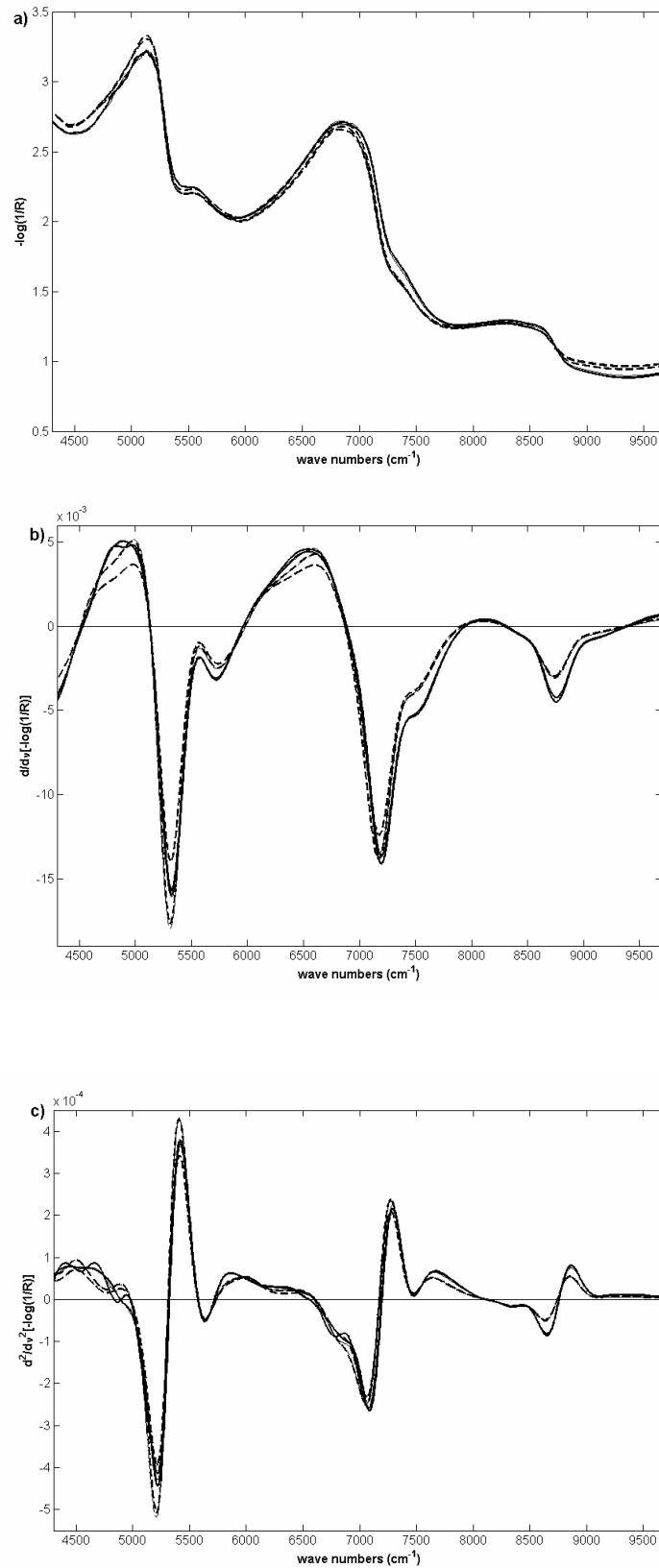
espalhamento da luz que ocorre na amostra. Este espalhamento da luz é por sua vez função do comprimento de onda, do tamanho e do formato das partículas, bem como da densidade de material particulado na amostra: quanto maior o número de partículas, e quanto menor seu tamanho, maior o espalhamento e, conseqüentemente, menor o caminho óptico [27, 28]. E é por esta razão que amostras de concentrados de tomate contendo maior concentração de sólidos de tomate e, portanto, de material particulado, apresentam maiores desvios das linhas de base.

Para produtos contendo altos teores de tomate, observa-se também uma desproporção entre as bandas de vibração O-H em 4900 e em 6800  $\text{cm}^{-1}$ . Entretanto, como esta desproporção praticamente desaparece após aplicação de correção por MSC (Figura 6.4) conclui-se que a maior parte desta desproporção é devida a efeitos de linha de base. Além disso, em 4900  $\text{cm}^{-1}$  há a sobreposição da vibração das ligações O-H de ácidos carboxílicos [27] e, portanto, parte desta desproporção deve-se também à variação de acidez entre produtos de diferentes países e regiões.

A Figura 6.4a mostra que o pré-processamento por MSC reduziu de forma efetiva os desvios de linha de base originalmente presentes nos espectros, mas manteve inalteradas as inclinações das linhas de base. Derivadas também foram aplicadas como pré-processamento dos espectros, após alisamento por filtros de Fourier; enquanto a primeira-derivada removeu completamente os desvios nas linhas de base - as inclinações sendo convertidas em um desvio constante -, a segunda-derivada foi capaz de remover tanto desvios quanto inclinações das linhas de base. Entretanto, como as derivadas são numericamente calculadas como diferenças entre as intensidades em intervalos constantes de números de onda, elas têm o efeito negativo de reduzir a razão sinal-ruído [22-24].



**Figura 6.3** – (a) Espectros originais e (b) espectros alisados. As linhas contínuas e tracejadas representam, respectivamente, amostras contendo teores baixos e elevados de sólidos de tomate.



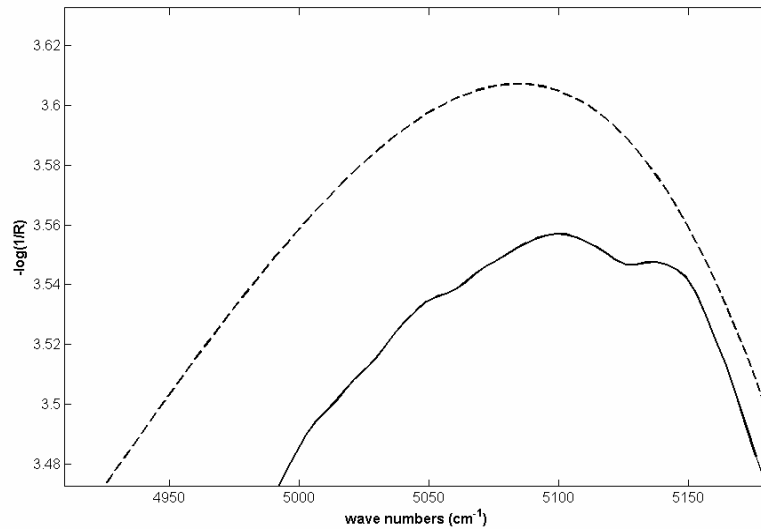
**Figura 6.4** – Pré-processamento de espectros por a) MSC; b) primeira- e c) segunda-derivada após alisamento por filtro de Fourier com função de convolução de 17 pontos.

A Tabela 6.1 mostra os modelos de calibração PLS2 que apresentaram as melhores capacidades de previsão - em termos de  $RMSEP_p$  e  $r_{val,p}$  - dentro de cada técnica de pré-processamento de espectros (nenhuma amostra anômala foi encontrada em qualquer dos modelos). Os modelos que apresentaram os menores  $RMSEP_p$  e maiores  $r_{val,p}$  foram obtidos por pré-processamento de espectros por MSC após alisamento, sendo que ambos os processos – alisamento pela média ou por filtro de Fourier – produziram resultados similares. Todavia, apesar de ambos os procedimentos de alisamento aumentarem significativamente a razão sinal-ruído sem alterar a quantidade de informação nos espectros, a Figura 6.5 mostra que o filtro de Fourier foi mais efetivo do que o alisamento pela média. Esta diferença foi particularmente importante para calibrações onde derivadas foram utilizadas como pré-processamento e é devida ao fato dos filtros de Fourier reduzirem os sinais de alta frequência diretamente nos interferogramas dos espectros.

**Tabela 6.1** – Número de fatores e erros de previsão para os modelos obtidos com regressão PLS2.

Pré-Processamento		Fatores	RMSECV <sub>p</sub>	RMSEP <sub>p</sub>	r <sub>val,p</sub>
Alisamento	Linha de Base				
Nenhum	Nenhum	8	1,11	1,09	0,970
	MSC	5	0,89	0,94	0,979
	1 <sup>a</sup> -derivada	9	1,32	1,35	0,957
	2 <sup>a</sup> -derivada	9	2,02	2,42	0,841
Alisamento pela Média	Nenhum	8	0,88	0,89	0,981
	MSC	4	0,83	0,83	0,982
	1 <sup>a</sup> -derivada	9	1,06	1,19	0,971
	2 <sup>a</sup> -derivada	9	2,12	2,15	0,856
Filtro de Fourier	Nenhum	8	0,84	0,83	0,982
	MSC	4	0,82	0,83	0,982
	1 <sup>a</sup> -derivada	7	0,99	0,93	0,978
	2 <sup>a</sup> -derivada	7	1,08	1,00	0,970

**Obs.** Os erros de previsão para os modelos PLS2 são globais, ou seja, referem-se às médias quadráticas ponderadas de todas as propriedades em estudo.



**Figura 6.5** - Diferença na eficiência de alisamento pela média (linha sólida) e por filtro de Fourier (linha tracejada). Os espectros foram deslocados para maior clareza.

Os modelos obtidos com as derivadas mostraram-se mais complexos – requereram um número maior de fatores – e apresentaram menores capacidades preditivas que aqueles obtidos com a aplicação de pré-processamento por MSC. Apesar do alisamento por filtro de Fourier ser mais efetivo em reduzir o ruído dos espectros originais, mais fatores foram necessários para acomodar o ruído re-inserido pelas derivadas.

Modelos de calibração PLS1 também foram individualmente construídos para cada propriedade; a Tabela 6.2 mostra a comparação entre o melhor modelo PLS2 e os modelos PLS1 que apresentaram as melhores capacidades preditivas. Pode-se observar que o modelo obtido com a regressão PLS2 mostrou capacidades de previsão similares às aquelas obtidas com a regressão PLS2 (testes-F, Tabela 6.2). Todavia, a regressão PLS2 demandou um número consideravelmente menor de fatores que a PLS1 para quase todas as propriedades, a exceção sendo o teor de frutose, para o qual o número de fatores foi igual para ambos os métodos de calibração.

**Tabela 6.2** – Comparação entre os melhores modelos de calibração PLS2 e PLS1.

Propriedade	PLS2 (4 fatores) <sup>a</sup>		PLS1			Teste-F Valor-p
	RMSEP	$r_{val}$	Fatores	RMSEP	$r_{val}$	
Sólidos Totais (%)	0,63	0,999	8	0,81	0,998	0,53
Sólidos Solúveis (%)	0,68	0,999	9	0,35	0,999	0,30
Acidez Total (% ac cítrico)	0,22	0,970	10	0,21	0,974	0,62
Açúcares Totais (%)	1,96	0,962	6	2,19	0,943	0,60
Glicose (%)	0,54	0,991	6	0,59	0,990	0,61
Frutose (%)	0,88	0,973	4	0,93	0,967	0,62

**Obs.:** a – pré-processamentos para PLS2: alisamento pela média com janela de 50 pontos seguida por MSC.

A regressão PLS2 forneceu modelos com capacidade preditiva satisfatória e com menor número de fatores que a maioria dos modelos PLS1 devido a uma combinação de características específicas deste conjunto de dados: as propriedades em estudo apresentam correlação (Tabela 6.3), especialmente com o teor de sólidos totais, uma propriedade que pode ser determinada com bastante precisão pelo método de referência.

**Tabela 6.3** – Coeficientes de correlação entre as propriedades em estudo.

	Sólidos Totais	Sólidos Solúveis	Acidez Total	Açúcares Totais	Glicose	Frutose
Sólidos Totais	1,000	0,995	0,796	0,746	0,916	0,897
Sólidos Solúveis		1,000	0,791	0,750	0,902	0,897
Acidez Total			1,000	0,550	0,646	0,624
Açúcares Totais				1,000	0,750	0,859
Glicose					1,000	0,935
Frutose						1,000

A Tabela 6.4 mostra os valores medidos e previstos, bem como os erros-padrão de previsão para as sete amostras utilizadas na validação externa. Os modelos PLS2 apresentaram erros comparáveis àqueles obtidos com os métodos de referência para sólidos totais e solúveis. Entretanto, os erros foram ligeiramente maiores para açúcares e consideravelmente maiores para acidez total, provavelmente devido ao fato desta propriedade estar sendo calibrada de forma indireta pela técnica NIR.

**Tabela 6.4** - Valores medidos, valores previstos e desvios de previsão para as amostras do conjunto de validação externa.

Amostra	Sólidos Totais Faixa de Calibração: 6,94-35,96%			Sólidos Solúveis Faixa de Calibração: 6,8 – 31,1%		
	Medido	Previsto	Desvio (%)	Medido	Previsto	Desvio (%)
1	8,23	7,92	3,72	7,8	7,8	0,39
2	19,32	19,72	2,06	18,3	17,9	1,83
3	16,92	16,72	1,19	16,8	16,7	0,21
4	19,97	19,82	0,77	19,1	19,5	2,34
5	25,74	27,13	5,39	25,4	25,5	0,34
6	9,69	9,86	1,69	9,3	9,4	0,84
7	14,76	14,19	3,87	14,0	13,7	2,09
Amostra	Acidez Total Faixa de Calibração: 0,28 – 2,46%			Açúcares Totais Faixa de Calibração: 0,87–17,73%		
	Medido	Previsto	Desvio (%)	Medido	Previsto	Desvio (%)
1	0,36	0,32	11,11	4,45	5,41	21,60
2	0,78	0,80	2,56	4,89	6,00	22,70
3	0,51	0,57	11,64	4,12	3,08	25,24
4	1,20	1,30	9,16	4,75	4,09	13,80
5	1,26	1,10	2,39	13,99	16,40	17,23
6	0,66	0,59	12,66	4,18	3,20	23,40
7	0,88	0,79	9,71	3,24	3,46	6,86
Amostra	Frutose Faixa de Calibração: 0,92 – 9,55%			Glicose Faixa de Calibração: 1,30 – 8,18%		
	Medido	Previsto	Desvio (%)	Medido	Previsto	Desvio (%)
1	1,71	1,52	10,98	1,34	1,46	9,06
2	3,96	3,06	22,74	3,39	2,77	18,31
3	3,15	3,22	2,26	2,67	2,22	16,87
4	5,22	4,43	15,20	3,07	3,48	13,44
5	6,73	6,09	9,53	7,26	8,51	17,21
6	1,78	1,95	9,78	1,61	1,49	7,44
7	3,33	3,70	11,13	2,63	2,32	11,97

**Obs.:** Erros típicos dos métodos de referência: sólidos totais: 5%; sólidos solúveis: 3%; acidez total: 10%; açúcares (totais, frutose e sacarose): 15%.

## **5 – Conclusões**

Modelos de calibração satisfatórios para determinação do prazo de validade de produtos atomatados foram obtidos utilizando-se a regressão PLS2 para determinação de sólidos, acidez e açúcares (totais, glicose e frutose) em produtos de tomate por espectroscopia NIR.

Apesar dos erros-padrão de previsão para acidez e açúcares apresentarem-se ligeiramente mais elevados que aqueles obtidos pelos métodos de referência, os modelos de calibração para estas propriedades podem ser melhorados (a) pela inclusão de um número maior de amostras no modelo de calibração e (b) pela construção de modelos específicos para cada região de origem dos produtos atomatados.



## 6 – Referências

1. Pedro AMK, Ferreira MMC, Simultaneously Calibrating Solids, Sugars and Acidity of Tomato Products Using PLS2 and NIR Spectroscopy, *Analytica Chimica Acta*, **2007**, 595. 221-227.
2. Gould WA, Tomato Production, Processing & Technology, 3<sup>rd</sup> ed., CTI Pub, Baltimore, 1992, pp 15-120, 236-349.
3. World Processing Tomato Council (WPTC), release # 12, <http://www.wptc.to/releases/releases12.pdf>, oct.05.06.
4. USDA Foreign Agricultural Service (FAS), *World Tomato and Tomato Products Situation and Outlook*, [http://www.fas.usda.gov/http/Hort\\_Circular/2005/08-05/08-01-05%20Tomato%20article.pdf](http://www.fas.usda.gov/http/Hort_Circular/2005/08-05/08-01-05%20Tomato%20article.pdf), oct.05.2006.
5. Salunkhe DK, Kadam SS, Handbook of Vegetable Science and Technology, Marcel Dekker, New York, 1998, pp. 171-2002.
6. Nielsen SS, Food Analysis, 3<sup>rd</sup> ed., Gaithersburg Pub, Aspen, 2003, pp. 112-32.
7. Cháves-Servín JL, Castellote AI, López-Sabater MC, Analysis of mono- and disaccharides in milk-based formulae by high performance liquid chromatography with refractive index detector, *J. Chromatogr. A*, 1043 (2004) 211-215.
8. Moskowitz HR, Product Testing and Sensory Evaluation of Foods, Food & Nutrition Press, Westport, 1983.
9. Chen JY, Iyo C, Kawano S, Development of calibration with sample cell compensation for determining the fat content of unhomogenised raw milk by a simple near infrared transmittance method, *J. Near Infrared Spectrosc.*, **1999**, 7, 265-273.
10. Lammertyn J, Non-destructive measurement of acidity, soluble solids, and firmness of Jonagold apples using NIR-spectroscopy, *Trans. ASAE*, **1998**, 41, 1089-1094.
11. Armenta S, Garrigues S, de la Guardia M, Determination of edible oil parameters by near infrared spectrometry, *Analytica Chimica Acta*, **2003**, 596, 330-337.
12. Kawano S, Fujiwara T, Iwamoto M, Nondestructive determination of sugar content in satsuma mandarine using near-infrared (NIR) transmittance, *J. Jap. Soc. Hortic. Sci.*, **1993**, 62, 465-470.
13. Osborne SD, Kunnemyer R, Jordan RB, A low-cost system for the grading of kiwifruit, *J. Near Infrared Spectrosc.*, **1999**, 7, 9-15.
14. Kawano S, Watanabe H, Iwamoto M, Determination of sugar content in intact peaches by near-infrared spectroscopy with fiber optics in interactance mode, *J. Jap. Soc. Hortic. Sci.*, **1992**, 61, 445-451.
15. Hong TL, Tsou SCS, Determination of tomato quality by near infrared spectroscopy, *J. Near Infrared Spectrosc.*, **1998**, 6, A321-A324.
16. Goula AM, Adamopoulos KG, Estimating the composition of tomato juice products by near infrared spectroscopy *J. Near Infrared Spectrosc.*, **2003**, 11, 123-136.
17. Jha SN, Matsuoka T, Non-destructive determination of acid-brix ratio of tomato juice using near infrared spectroscopy, *Intern. J. Food Sci. Technol.*, **2004**, 39, 425-430.

18. Pedro AMK, Ferreira MMC, Nondestructive determination of solids and carotenoids in tomato products by near-infrared spectroscopy and multivariate calibration, *Anal. Chem.*, **2005**, 77, 2505-2511.
19. Bramley P, Is lycopene beneficial to human health?, *Phytochem.*, **2000**, 54, 233-236.
20. Tapiero H, Townsend DM, Tew KD, The role of carotenoids in the prevention of human pathologies, *Biomed. Pharmacother.*, **2004**, 58, 100-110.
21. Willcox J, Catignani GL, Lazarus S., *Crit. Rev. Food Sci. Nutr*, **2003**, 43, 1-18.
22. Martens H, Naes T, Multivariate Calibration, John Wiley & Sons, Chichester, 1993, pp. 116-165, 237-290, 314-354.
23. Brereton RG, Chemometrics: Data Analysis for the Laboratory and Analytical Plant, Wiley & Sons, Chichester, 2003, pp. 271-315.
24. Vandeginste DGM, Massart DL, Buydens LMC, de Jung S, Lewi PJ, Smeyers-Verbeke J, Handbook of Chemometrics Part B, Elsevier, Amsterdam, 1998, pp. 349-370.
25. Savitzky A, Golay MJE, Smoothing plus differentiation of data by simplified least squares procedures, *Anal. Chem.*, **1964**, 36, 1627.
26. Ferreira MMC, Antunes AM, Melgo MA, Volpe PLO, Quimiometria I: calibração multivariada, um tutorial, *Quím. Nova*, **1999**, 22, 724-731.
27. Olinger JM, Griffiths PR, Effects of sample dilution and particle size morphology on diffuse reflecton spectra of carbohydrate systems in the near-infrared and midinfrared 2. durum-wheat, *Appl. Spectrosc.*, **1993**, 47, 695-701.
28. Williams P, Norris K (eds.), NIR Technology in the Agricultural and Food Industries, Am. Ass. Cereal Chemists, St. Paul, 1990, pp.1-31.

**CONSIDERAÇÕES FINAIS E  
PERSPECTIVAS FUTURAS**

O Método Multivariado Acelerado para Determinação do Prazo de Validade constitui um avanço considerável sobre os protocolos tradicionais para determinação da validade e estudo dos mecanismos de degradação de produtos até então disponíveis. Baseando-se na análise de perfis multivariados de degradação, e não apenas na determinação dos parâmetros cinéticos para cada propriedade individualmente, o método possibilita que relações não-óbvias entre as diferentes propriedades dos produtos sejam determinadas. Além disso, o MASLT pode ser considerado um método auto-consistente no sentido em que nenhum conhecimento *a priori* é necessário na definição dos perfis de degradação e para que possíveis correlações entre propriedades sejam observadas.

O MASLT também possibilita que dados originários de instrumentação analítica multivariada, como espectrômetros, cromatógrafos, calorímetros e reômetros, entre outros, sejam utilizados diretamente no estudo da degradação e determinação do prazo de validade de produtos. Atualmente, estes prazos são determinados, por exemplo, tomando como base a variação de picos de compostos específicos em cromatogramas, ou o máximo de absorção em espectros, ou ainda a variação de parâmetros termodinâmicos calculados a partir de termogramas. Entretanto, não raras vezes o desempenho dos produtos não está relacionado com os valores absolutos de uma ou algumas propriedades, mas com suas proporções. Nestes casos, a utilização direta dos sinais fornecidos pelos instrumentos analíticos torna possível que estas proporções e correlações sejam determinadas através da análise dos perfis de fornecidos pelo MASLT. Esta vantagem foi, de fato, bem exemplificada no caso da determinação da validade de um cosmético utilizando a espectroscopia NIR, apresentado no Capítulo IV, onde benefícios secundários, ligados ao desempenho da fragrância do produto, mostraram-se inesperadamente relacionados com a variação do comportamento da linha de base dos espectros.

Neste trabalho apenas uma aplicação do MASLT utilizando dados de origem instrumental foi apresentada, mas é bastante provável que várias outras técnicas, além das acima citadas, possam ser exploradas para que estudos de validade mais robustos, rápidos e eficientes sejam conduzidos. O autor desta tese consegue visualizar duas aplicações imediatas para alimentos e cosméticos: a determinação da degradação de fragrâncias e/ou aromas utilizando cromatografia (gasosa e/ou líquida) e a decomposição de corantes utilizando espectroscopia UV-Vis. Atualmente a degradação de fragrâncias e aromas é estudada comparando-se a variação da concentração de alguns poucos compostos,

conhecidos como marcadores, com avaliações sensoriais. Entretanto, não raras vezes estas correlações são pobres, evidenciando que a técnica instrumental, aliada aos protocolos clássicos para determinação do prazo de validade, não reflete os fenômenos percebidos pelos consumidores no momento do uso dos produtos. A análise dos perfis cromatográficos através do MASLT, por sua vez, tem o potencial de revelar as correlações entre diferentes compostos, sem que nenhum conhecimento prévio seja necessário.

Uma grande vantagem do MASLT sobre os protocolos clássicos para determinação do prazo de validade é a capacidade de reunir dados de naturezas distintas, como físico-químicos, instrumentais e sensoriais, em um único conjunto de dados para definição dos perfis de degradação dos produtos. Além de evidenciar possíveis correlações entre variáveis de diferentes origens, o MASLT fornece uma visão geral integrada dos mecanismos de degradação dos produtos.

Adaptações podem ser incorporadas ao MASLT de modo a acomodar especificidades de alguns estudos. Existem casos onde uma ou algumas variáveis são mais relevantes que outras para determinação do prazo de validade do produto: variáveis funcionais podem ser mais relevantes que aquelas que trazem benefícios secundários, por exemplo. Nesses casos, a u-PCA pode ser substituída pela wu-PCA (*weighted unfolding Principal Component Analysis*) onde pesos diferentes são atribuídos às propriedades e, deste modo, aquelas mais relevantes contribuirão majoritariamente para a determinação do prazo de validade mesmo que não constituam as principais fontes de variação no conjunto de dados.

Além dessa modificação, outras propriedades inerentes aos métodos multivariados podem ser incorporadas ao MASLT, como a capacidade de lidar com dados faltantes através da utilização de uma grande gama de estratégias.