

Universidade Federal de Minas Gerais

Instituto de Ciências Exatas

Departamento de Química

Karen Monique Nunes

**Utilização de espectroscopia no infravermelho
médio, fusão de dados e métodos
quimiométricos de classificação na análise de
fraudes em carnes bovinas *in natura***

Belo Horizonte

2015

UFMG / ICEx / DQ. 1.108^a
D. 608^a

Karen Monique Nunes

**Utilização de espectroscopia no infravermelho
médio, fusão de dados e métodos
quimiométricos de classificação na análise de
fraudes em carnes bovinas *in natura***

**Dissertação apresentada
ao Departamento de
Química do Instituto de
Ciências Exatas da
Universidade Federal de
Minas Gerais como
requisito parcial para a
obtenção do grau de
Mestre em Química –
Química Analítica**

Belo Horizonte

2015

N972u Nunes, Karen Monique
2015 Utilização de espectroscopia no infravermelho
D médio, fusão de dados e métodos quimiométricos de
classificação na análise de fraudes em carnes bovinas
in natura [manuscrito] / Karen Monique Nunes. 2015.
[xiii], 99 f. : il.

Orientador: Marcelo Martins de Sena.

Dissertação (mestrado) - Universidade Federal de
Minas Gerais - Departamento de Química.
Inclui bibliografia.

1. Química analítica - Teses 2. Alimentos -
Adulteração e inspeção - Teses 3. Espectroscopia de
infravermelho - Teses 4. Carne bovina - Teses I. Sena,
Marcelo Martins de, Orientador II. Título.

CDU 043



**"Utilização de Espectroscopia no Infravermelho Médio, Fusão de Dados e
Métodos Quimiométricos de Classificação na Análise de Fraudes em Carnes
Bovinas *in natura*"**

Karen Monique Nunes

Dissertação aprovada pela banca examinadora constituída pelos Professores:

Prof. Marcelo Martins de Sena - Orientador
UFMG

Prof. Rodinei Augusti
UFMG

Eng.º Jorge Jardim Zacca
Departamento de Polícia Federal

Belo Horizonte, 13 de agosto de 2015.

*“Uma mente que se abre a uma nova ideia,
jamais voltará ao seu tamanho original”.*

Albert Einstein

Agradecimentos

Agradeço primeiramente a Deus pela oportunidade e por me dar forças durante o percurso.

À minha mãe Tereza por todo esforço e pela convicção de que o melhor que ela poderia fazer era me dar uma boa educação. Tenho tentado seguir seus ensinamentos

À minha família, pelo apoio emocional.

Ao Fernandes, grande amigo, que sempre esteve presente com seus conselhos valiosos.

Ao Wálacson, grande companheiro, por acreditar, me incentivar, me motivar.

Aos meus amigos e amigas pela compreensão e pelo apoio em meio às dificuldades desta etapa, em especial à minha amiga Bruna que acompanhou todo este processo e minhas amigas Gesliene e Dayane.

Às minhas amigas Juliana, Jane, Kele, Edinéia e Márcia pelo incentivo, torcida e apoio.

À minha amiga Louise, pela amizade, apoio, incentivo e conselhos.

Às minhas amigas Carlinha e Angélica, pela amizade, companheirismo, pelo apoio, palavras e motivação!

Às amigas Nancy, Rosy e Ciça com quem tive o prazer compartilhar diversos momentos durante este trabalho.

Ao Prof. Dr. Marcelo por me receber e me orientar, tornando possível o desenvolvimento e a conclusão deste trabalho.

Aos meus amigos do grupo de Química Analítica e Quimiometria, especialmente ao Leandro pelo conhecimento e auxílio compartilhados.

Ao DPF/ SETEC pela parceria e por acreditarem neste trabalho, especialmente ao Marcus Vinícius e ao Pires pelo empenho e disponibilidade como profissionais e, também, como amigos.

Aos meus amigos Maurício, Kraemer, Gyovany, Aline e Cris pelo companheirismo e cumplicidade de sempre.

Ao Lanagro/MG pelo apoio durante o condicionamento e processamento das amostras, em especial à Flávia Coelho, sempre muito receptiva.

À CAPES pelo custeamento pessoal da pesquisa.

À Universidade Federal de Minas Gerais, ao Instituto de Ciências Exatas e ao departamento de Química.

A todos que participaram direta ou indiretamente para a conclusão desta etapa.

Resumo

A preocupação com a autenticidade e segurança alimentar de carnes é crescente nos últimos anos, devido a grandes escândalos de fraudes. A adição de água, sais e outros adulterantes, como carragena (polissacarídeo linear sulfatado obtido dos extratos de algas marinhas vermelhas) pode aumentar a capacidade de retenção de água da carne, propiciando uma fraude econômica por ganho no peso. Esta dissertação teve como principal objetivo o estudo e identificação de adulterações em amostras de carne bovina *in natura* empregando a análise de 5 variáveis físico-químicas (FQ), teores de proteína, cinzas, cloreto, sódio e fosfato, e espectroscopia no infravermelho médio (MIR) com acessório de reflectância total atenuada. Esses dados foram tratados usando métodos quimiométricos de classificação não supervisionada (análise de componentes principais, PCA) e supervisionada (análise discriminante por mínimos quadrados parciais, PLS-DA). Na primeira aplicação, foram analisadas 43 amostras reais de carne bovina de diversos tipos de cortes, obtidas de uma operação da polícia. Os modelos foram construídos a partir destas e de mais 12 amostras controle, garantidamente não adulteradas. Um modelo PCA semi-supervisionado com os dados FQ permitiu detectar 74% das amostras adulteradas. No entanto, melhores modelos foram obtidos usando PLS-DA. O modelo usando somente os dados FQ teve uma capacidade preditiva bem superior ao modelo construído com os espectros MIR. No entanto, o melhor modelo foi obtido pela fusão de dados das variáveis FQ com os espectros MIR, classificando corretamente todas as amostras controle e apresentando apenas 4 falso-negativos para a classificação das amostras adulteradas. Um segundo modelo de fusão de dados foi obtido através da combinação da variável FQ de maior poder preditivo, cloreto, e de 8 regiões espectrais selecionadas a partir de um vetor informativo (*VIPscores*). Este modelo forneceu resultados um pouco inferiores, mas apresentou a vantagem de envolver variáveis medidas de maneira mais simples, rápida e barata. Na segunda aplicação, amostras de carne foram adulteradas de maneira controlada pela injeção de 4 tipos de adulterantes (água, tripolifostafa e misturas contendo sais, colágeno, carragena e maltodextrina (um carboidrato complexo proveniente de amido)). As purgas, líquido exsudado após procedimento de descongelamento da peça, de 51 amostras de carne foram analisadas por espectroscopia MIR e os modelos

PLS-DA elaborados foram capazes de classificar corretamente todas as amostras não adulteradas, permitindo ainda identificar as regiões espectrais mais seletivas para a previsão de cada tipo de adulteração.

Palavras-chave: Adulteração de carnes; PLS-DA; Fusão de dados; Espectroscopia no infravermelho médio; Análise forense

Abstract

The concern with the food safety and authenticity of meats is increasing in the last years, due to the occurrence of great fraud scandals. The addition of some salts and other adulterants, such as carrageen (a linear sulphated polysaccharide extracted from red edible seaweeds), can increase the meat water holding capacity, providing an economic fraud by weight gain. This dissertation aims the detection and characterization of adulterations in bovine meat *in natura* by determining five physico-chemical variables, the contents of protein, ash, chloride, sodium and phosphate, and using attenuated total reflection mid infrared spectroscopy (MIRS). The generated data were treated with non-supervised (principal component analysis, PCA) and supervised (partial least squares discriminant analysis, PLS-DA) classification methods. In the first application, 43 adulterated meat samples of different types of cuts were analysed, which were obtained from a real police operation. The models were built with these samples plus 12 control samples, which were guaranteed to be non-adulterated. A semi-supervised PCA model using the physico-chemical data detected 74% of the adulterated samples. However, better results were obtained with PLS-DA. A model built with only the physico-chemical data had a better predictive ability than a model built with the MIR spectra. But, the best model were obtained with the data fusion of the physico-chemical data and MIR spectra, which correctly classified all the control samples and provided only 4 false-negatives for the prediction of adulterated samples. A second data fusion model were constructed by combining the most predictive physico-chemical variable, chloride, and 8 spectral regions selected from an informative vector (*VIPscores*). This model provided somewhat lower prediction ability, but has the advantage of utilizing variables which were measured simpler, faster and at a low cost. In the second application, meat samples were adulterated in a controlled manner by the injection of 4 types of adulterants (water, tripolyphosphate, and mixtures containing salts, collagene, carrageen and maltodextrin (a complex carbohydrate obtained from starch)). The purges, the liquid exudated after the meat thawing procedure, of 51 meat samples were analysed by MIRS and the constructed PLS-DA models were able to correctly

classify all the non-adulterated samples, allowing in addition to identify the spectral regions more selective for each type of adulteration.

Keywords: Meat adulteration; PLS-DA; Data fusion; Mid infrared spectroscopy; Forensic analysis

Lista de Abreviaturas

a. C.	Antes de Cristo
ATR	Reflectância total atenuada (<i>Attenuated Total Reflectance</i>)
CPs	Componentes Principais
CVCE	Erro de classificação de validação cruzada (<i>cross-validation classification error</i>)
DBO	Demanda bioquímica de oxigênio
DRIFTS	Espectrometria de reflexão (ou reflectância) difusa no infravermelho com transformada de Fourier (<i>diffuse reflectance infrared Fourier transform spectrometry</i>)
EFR	Taxa de eficiência (<i>Efficiency Rate</i>)
FN	Falsos Negativos
FNR	Taxa de Falsos Negativos (<i>False Negative Rate</i>)
FP	Falsos Positivos
FPR	Taxa de Falsos Positivos (<i>False Positive Rate</i>)
FTIR	Infravermelho médio com transformada de Fourier
HCA	Análise de Agrupamentos hierárquica (<i>Hierarchical Cluster Analysis</i>)
ICS	Sociedade Internacional de Quimiometria (<i>International Chemometrics Society</i>)
KNN	Método do K-ésimo vizinho mais próximo (<i>Kth Nearest Neighbour</i>)
KS	Kennard-Stone
LDA	Análise discriminante linear (<i>Linear Discriminant Analysis</i>)
MIR	Infravermelho médio (<i>Mid Infrared</i>)
NIPALS	Algoritmo iterativo não-linear de mínimos quadrados parciais (<i>Non-linear iterative partial least squares</i>)
NIR	Infravermelho próximo (<i>Near Infrared</i>)

PCA	Análise de componentes principais (Principal Component Analysis)
PCR	Reação em cadeia da polimerase (<i>Polymerase Chain Reaction</i>)
pH	Potencial hidrogeniônico
PLS	Mínimos quadrados parciais (<i>Partial Least Squares</i>)
PLS-DA	Análise discriminante por mínimos quadrados parciais (<i>Partial Least Squares Discriminant Analysis</i>)
RIISPOA	Regulamento de Inspeção Industrial e Sanitária de Produtos de Origem Animal
RMN	Ressonância Magnética Nuclear
RMSECV	Raiz quadrada do erro médio quadrático de validação cruzada (<i>root mean square error of cross-validation</i>)
SIMCA	Modelos independentes de similaridade utilizando componentes principais (<i>Soft Independent Modelling of Class Analogies</i>)
SPR	Taxa de especificidade (<i>Specificity Rate</i>)
STR	Taxa de sensibilidade (<i>Selectivity Rate</i>)
SVM	Máquinas de suporte de vetores (<i>Support Vector Machines</i>)
TACO	Tabela Brasileira de Composição de Alimentos
TN	Número total de amostras negativas conhecidas
TP	Número total de amostras positivas conhecidas
UV-Vis	Ultravioleta-visível
VIP scores	Importância das Variáveis na Projeção dos escores (<i>Variable Importance in Projection</i>)
VLs	Variáveis latentes
WHC	Capacidade de retenção de água da carne (<i>Water Holding Capacity</i>)

Lista de Figuras

Figura 1: Componentes básicos de um equipamento que opera na região do infravermelho.	15
Figura 2: Componentes básicos de um espectrômetro FTIR.	16
Figura 3: Esquema típico de uma célula de reflectância total atenuada.	19
Figura 4: Representação da organização da matriz X a partir de dados espectroscópicos.	23
Figura 5: Representação da decomposição da matriz de dados em escores e pesos, de acordo com o modelo PCA	24
Figura 6: Representação do modelo PLS1-DA.	25
Figura 7: Representação do modelo PLS2-DA.	25
Figura 8: representação esquemática dos três níveis de fusão de dados.	31
Figura 9: Peça de Contra-Filé apreendida na Operação Vaca Atolada com formato paralelepipedal e apresentando líquido congelado nas extremidades.	34
Figura 10: Fotografia de amostras apreendidas cortadas após coleta das perdas por descongelamento. As setas indicam pontos de exsudação anormal de líquidos de característica viscosa e presença de espuma.	36
Figura 11: Corte das peças de carne com serra de fita. (a) procedimento de corte. (b) destaque para a amostra cortada em pedaços menores.	36
Figura 12: Processo de trituração e homogeneização da amostra de carne. (a) trituração da amostra. (b) carne triturada sendo homogeneizada com auxílio de espátula.	37
Figura 13: Evidências de perfurações puntiformes na superfície de uma amostra apreendida.	37
Figura 14: Evidência do acúmulo de líquido congelado nas extremidades das peças de amostras congeladas.	39

Figura 15: Volumes de líquido exsudado após procedimento de descongelamento (purga) da amostra controle e das amostras questionadas. As setas indicam o volume esperado proporcional ao peso da amostra controle (corte de 1.692g).	40
Figura 16: (a) Máquinas injetora empregada em adulteração de carne bovina <i>in natura</i> e (b) amaciadora de carne.	41
Figura 17: Amostra teste demonstrando o ganho de peso e volume com a injeção e aumento de 3 % m/m da peça.	45
Figura 18: Marcas de injeção (setas vermelhas). (a) antes e (b) após embalagem à vácuo.	46
Figura 19: (a) Espectrofotômetro de infravermelho médio FTIR Nicolet 380 com acessório de reflectância total atenuada (ATR). (b) análise de amostra de purga.	48
Figura 20: Escores de CP1 x CP2 para dados físico-químicos das amostras reais	50
Figura 21: Escores de CP1 x CP2 para dados físico-químicos de amostras reais - Tipos de corte	51
Figura 22: Escores de CP1 x CP2 para dados físico-químicos de amostras reais - Análise Semi-supervisionada	52
Figura 23: Pesos para os dados físico-químicos das amostras reais.	53
Figura 24: Gráfico de resíduos Q do modelo em função de T^2 de Hotelling (influência) para um modelo com 2 CPs. As linhas pontilhadas indicam limites de confiança de 95% para os parâmetros.	53
Figura 25: Espectros MIR para as 55 amostras. Em vermelho as amostras controle e em verde as amostras suspeitas. Em azul, as regiões espectrais eliminadas são destacadas.	54
Figura 26: Escores de CP1 x CP2 para dados ATR-FTIR de amostras – Tipos de carne.	55
Figura 27: Escores de CP1 x CP2 para dados ATR-FTIR.	56

Figura 28: Previsões para o modelo PLS-DA com os dados físico-químicos. A linha tracejada vermelha indica o limite de decisão (<i>threshold</i>) e a linha tracejada azul indica a separação dos conjuntos treinamento e teste.	58
Figura 29: Coeficiente de Regressão para as variáveis físico-químicas.	59
Figura 30: VIP scores para as variáveis físico-químicas. A linha pontilhada indica o valor de 1,0, sugerido como limite de significância.	60
Figura 31: Previsões para o modelo PLS-DA construído com os espectros MIR.	61
Figura 32: Coeficiente de Regressão para as variáveis de ATR-FTIR.	62
Figura 33: <i>VIP scores</i> para o modelo PLS-DA com os espectros ATR-FTIR.	63
Figura 34: Previsões para o modelo PLS-DA (4 VLs) com fusão dados completa. A linha tracejada em azul separa os conjuntos treinamento e teste.	64
Figura 35: Previsões para o modelo PLS-DA (7 VLs) com fusão dados completa. A linha tracejada em azul separa os conjuntos treinamento e teste.	65
Figura 36: Coeficientes de Regressão das variáveis espectrais do modelo PLS-DA de fusão de dados completo	67
Figura 37: <i>VIP scores</i> das variáveis espectrais do modelo PLS-DA de fusão de dados completo.	67
Figura 38: Previsões para o modelo PLS-DA com fusão dados e seleção de variáveis.	68
Figura 39: Espectro MIR das 51 amostras simuladas	73
Figura 40: Previsões de validação cruzada do modelo PLS-DA construído com os espectros MIR para as amostras simuladas.	74
Figura 41: Coeficientes de Regressão do modelo PLS-DA para as amostras simuladas.	75

Figura 42: <i>VIP</i> scores do modelo PLS-DA para as amostras simuladas.	75
Figura 43: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção de água.	76
Figura 44: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção de água.	77
Figura 45: <i>VIP</i> scores do modelo PLS-DA para as amostras adulteradas por injeção de água.	77
Figura 46: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção de solução de tripolifosfato de sódio.	78
Figura 47: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção de solução de tripolifosfato de sódio.	79
Figura 48: <i>VIP</i> scores do modelo PLS-DA para as amostras adulteradas por injeção de solução de tripolifosfato de sódio.	80
Figura 49: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção da mistura C.	81
Figura 50: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção da mistura C.	82
Figura 51: <i>VIP</i> scores do modelo PLS-DA para as amostras adulteradas por injeção da mistura C.	82
Figura 52: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção da mistura D.	83
Figura 53: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção da mistura D.	84
Figura 54: <i>VIP</i> scores do modelo PLS-DA para as amostras adulteradas por injeção da mistura D.	85

Lista de Tabelas

Tabela 1: Abate de bovinos e exportação de carne bovina <i>in natura</i> - Brasil – trimestres selecionados de 2013 e 2014	8
Tabela 2: Composição de cortes bovinos por 100 gramas de parte comestível	9
Tabela 3: Regiões espectrais no Infravermelho	14
Tabela 4: Figuras de mérito do modelo PLS-DA para os dados físico-químicos	58
Tabela 5: Figuras de mérito do modelo PLS-DA para os dados ATR-FTIR das amostras reais	62
Tabela 6: Figuras de mérito para o modelo PLS-DA construído com Fusão de Dados Completa	66
Tabela 7: Figuras de mérito para o modelo PLS-DA construído com Fusão de Dados e seleção de variáveis	69
Tabela 8: Comparação dos 4 modelos PLS-DA construídos para as amostras reais	70

Lista de Gráficos

Gráfico 1: Evolução do abate de bovinos por trimestre - Brasil – trimestres 2009-2014	7
Gráfico 2: Evolução do peso acumulado de carcaças de bovinos por trimestre – Brasil -trimestres 2009-2014	7
Gráfico 3: Evolução da participação de machos e fêmeas no abate total de bovinos por trimestre - Brasil - trimestres 2009-2014	8

Sumário

1.	Introdução	1
2.	Revisão Bibliográfica.....	6
2.1	Carne bovina.....	6
2.1.1	Abate de bovinos	6
2.1.2	Composição da carne bovina	9
2.1.3	Fatores de variabilidade da carne bovina ²⁸	9
2.2	Controle de qualidade de alimentos de origem animal no Brasil	11
2.3	Espectroscopia na região do infravermelho	13
2.3.1	Instrumentos para o Infravermelho	15
2.3.1.1	Instrumentos dispersivos.....	15
2.3.1.2	Espectrômetros com Transformada de Fourier.....	16
2.3.2	Aplicações da Espectrometria no Infravermelho.....	17
2.3.2.1	Espectrometria no infravermelho médio.....	17
2.3.2.1.1	Reflexão Difusa.....	18
2.3.2.1.2	Reflexão Total Atenuada ⁴⁰	18
2.4	Quimiometria.....	20
2.4.1	Reconhecimento de Padrões	21
2.4.1.1	Análise de Componentes Principais - PCA.....	22
2.4.1.2	Análise discriminante por mínimos quadrados parciais - PLS- DA 25	
2.4.1.3	Algoritmo de Kennard-Stone	28
2.4.1.4	VIPscores.....	29
2.4.1.5	Fusão de Dados – Data Fusion.....	29
3	Materiais e Métodos	34
3.1	Amostras reais	34
3.1.1	A Operação Vaca Atolada	34
3.1.2	Reagentes e amostras de carne bovina	42
3.1.3	Determinação de parâmetros físico-químicos ²⁴	42
3.1.3.1	Determinação de proteína total	42
3.1.3.2	Determinação do teor de cinzas.....	43
3.1.3.3	Determinação de Na ⁺ , Cl ⁻ e PO ₄ ³⁻ por cromatografia de íons	44
3.2	Amostras simuladas (Plano de estudo).....	44
3.2.1	Preparo das amostras	44
3.3	Dados espectroscópicos	47
3.3.1	Equipamento.....	47

3.3.2	Aquisição dos espectros.....	47
3.4	Processamento dos dados.....	48
4	Resultados e discussão	49
4.1	Amostras Reais.....	49
4.1.1	Modelos PCA.....	49
4.1.1.1	Modelo PCA para os dados Físico-Químicos.....	49
4.1.1.2	Modelo PCA para os dados de ATR-FTIR	54
4.1.2	Modelos PLS-DA	57
4.1.2.1	Modelo PLS-DA para os dados Físico-Químicos	57
4.1.2.2	Modelo PLS-DA para os dados ATR-FTIR.....	60
4.1.3	Fusão de Dados	64
4.1.3.1	Modelo PLS-DA de Fusão de Dados Completa.....	64
4.1.3.2	Modelo PLS-DA de Fusão de Dados com Seleção de Variáveis 68	
4.1.4	Comparação dos modelos PLS-DA para as amostras reais.....	69
4.2	Amostras Simuladas	72
4.2.1	Modelos PLS-DA	72
4.2.1.1	Modelo PLS-DA com os 4 tipos de adulteração.....	73
4.2.1.2	Modelo PLS-DA para amostras adulteradas por injeção de água	75
4.2.1.3	Modelo PLS-DA Modelo PLS-DA para amostras adulteradas por injeção de tripolifostafo.....	78
4.2.1.4	Modelo PLS-DA Com Separação em classes Controle e Grupo C	80
4.2.1.5	Modelo PLS-DA para amostras adulteradas por injeção da mistura D83	
4.2.2	Conclusão parcial	85
5	Conclusão	86
6	Referências Bibliográficas.....	89
	ANEXO A	96
	ANEXO B	98

1. Introdução

A alimentação é o processo pelo qual o organismo dos seres vivos adquire as substâncias necessárias ao seu crescimento, à sua manutenção e à sua reprodução.¹ Os alimentos são diferentes entre si em sua capacidade nutritiva, possuindo teores distintos de diversas substâncias químicas.² O consumo de alimentos de origem animal impacta diretamente na qualidade de vida das pessoas, e dentre tais alimentos destaca-se a carne. Carne é definida como qualquer tecido de origem animal que sirva como alimento.² A carne bovina, especificamente, é um alimento com alta densidade nutricional, pois se apresenta como fonte de proteínas e nutrientes, tais como os minerais zinco, fósforo e ferro, além de ácidos graxos e vitaminas do complexo B.^{3,4}

A adição de substâncias químicas com o intuito comercial de aumentar lucros é um processo conhecido como fraude ou adulteração. Este processo está associado a manobras para mascarar propriedades indesejadas e/ou condições sanitárias inadequadas, incluindo a inserção de substâncias a fim de conferir características que não são usuais, como textura, odor e sabor.⁵ A adulteração inter-espécies em produtos cárneos pode acarretar ofensa aos princípios de grupos religiosos que não consomem determinados tipos de carne e ao direito do consumidor em saber exatamente o que está comprando e consumindo. Outra forma de adulteração de carne pode ocorrer com a injeção de substâncias ou compostos comerciais não descritos nos rótulos, podendo ocasionar problemas de saúde a gestantes e a pessoas alérgicas.⁶

Nos últimos anos, diversos escândalos de adulteração de alimentos tiveram grande repercussão mundial, o que gera preocupação nas autoridades competentes em diversos Países. Como exemplo, pode-se citar a Operação “Ouro Branco”, deflagrada pela Polícia Federal em 2007, no Brasil, cujo objetivo foi investigar a fraude em leites com diversos tipos de substâncias químicas para mascarar propriedades físicas e condições sanitárias inadequadas.⁷ Outra operação relacionada à adulteração de leite foi a Operação “Leite Compensado”, em 2013, em que transportadores do leite cru adicionavam substâncias químicas a fim de aumentar o volume do produto e a sua validade.⁸

Escândalos relacionados a fraudes em carnes ocorreram nos anos de 2012 e 2013. A Operação “Vaca Atolada”, deflagrada pela Polícia Federal em 2012, na qual frigoríficos de Minas Gerais foram investigados por adulterarem carnes bovinas *in natura* a fim de alterar seu peso e valor nutricional, aumentando assim a lucratividade.⁹ Pode-se citar, ainda, o grande escândalo envolvendo a venda de carne de cavalo como carne bovina que ocorreu na Europa em 2013.¹⁰

Comumente, são encontrados relatos de adulteração de carnes bovinas processadas com outros componentes cárneos de menor custo.¹¹ Estes relatos têm fomentado a realização de estudos a fim de propor métodos analíticos para a identificação de fraudes inter-espécies, como: identificação das frações de carne suína e bovina em carne processada utilizando PCR (reação em cadeia da polimerase, do inglês *Polymerase Chain Reaction*) multiplex em tempo real,¹² técnica que permite a quantificação de DNA de espécies animais presentes no produto cárneo em tempo real, entretanto, esta é uma técnica laboriosa por necessitar de padrões de todos os tipos cárneos presentes na amostra em estudo; uso de espectroscopia Raman na determinação do teor de carne de cavalo em mistura com outras carnes,¹³ cuja determinação não foi multivariada/quimiométrica, mas baseada na razão entre alguns números de onda; detecção de adulteração de carne bovina processada com carne de peru por fusão de dados de três diferentes técnicas espectroscópicas (UV-Vis , infravermelho próximo – NIR – e médio - MIR)¹¹; discriminação de carnes bovina, de lhama e de cavalo por modelos de classificação supervisionada (PLS-DA, análise discriminante por mínimos quadrados parciais) a partir de medidas de reflectância para amostras de carne e de transreflectância para amostras de purga (líquido exsudado após o descongelamento da amostra de carne), ambas no NIR;¹⁴ quantificação da adulteração de carne bovina por carne de porco em almôndegas usando calibração multivariada no infravermelho médio com transformada de Fourier – FTIR;¹⁵ entre outros trabalhos na linha de fraude alimentar.

Os principais problemas de autenticação de carnes podem ser classificados em quatro tipos: origem (sexo, corte, raça do animal, etc.), substituição de uma carne por outra, tipo de processamento e adição de ingredientes não cárneos.¹⁶ Conforme relatado no parágrafo anterior, o tipo mais comum de fraude descrita

na literatura é a adulteração de carnes bovinas processadas com outros componentes cárneos de menor custo, ou seja, adulterações inter-espécies. Este também é o tipo de fraude de maior repercussão mundial, como no caso do já mencionado escândalo da carne de cavalo.¹⁰ Entretanto, adulterações com ingredientes não cárneos, com o intuito de aumentar o ganho econômico, são o tipo menos comum, raramente descrito na literatura.¹⁶ Um dos poucos exemplos é o desenvolvimento de um método baseado em cromatografia líquida acoplada com espectrometria de massas para detectar a adulteração de carne suína pela adição de enzimas, as quais se ligam às proteínas do sangue gerando um aumento de massa no produto.¹⁷

A água também pode ser fraudulentamente adicionada à carne, visando ganho de massa, e existe regulamentação para a verificação deste tipo de fraude.¹⁶ O método padrão para determinar água adicionada à carne se baseia no estudo da razão entre os conteúdos de água e proteína.¹⁸ Se água exógena for adicionada à carne, a razão água/proteína se torna muito alta e serve como um claro indicador da fraude. No entanto, proteínas de outras fontes e sais, como fosfatos, também podem ser adicionados a produtos cárneos para aumentar a retenção de água, levando a razão água/proteína a valores próximos à sua proporção natural na carne. Neste caso, é necessário detectar proteínas e/ou sais exógenos para identificar a fraude.¹⁶ Na “Operação Vaca Atolada”,⁹ que é objeto de estudo desta Dissertação, seis frigoríficos foram autuados pela Polícia Federal em função de denúncias, e foram encontrados em loco uma série de produtos possivelmente usados como adulterantes. Os principais produtos encontrados foram sais, como cloretos, fosfatos, tripolifosfato de sódio, sal refinado iodado, pirofosfato ácido de sódio, além de carragena, um polissacarídeo linear sulfatado obtido dos extratos de algas marinhas vermelhas, e maltodextrina, um carboidrato complexo proveniente de amido e colágeno, este último uma proteína. A adição de sais e proteínas é feita a partir de soluções aquosas mantendo a relação água/proteína próxima à proporção natural encontrada na carne.

A adulteração por adição de sais visa o aumento da capacidade de retenção de água da carne (*Water Holding Capacity*, WHC),¹⁹ propiciando uma fraude econômica por ganho no peso, uma vez que o produto absorverá mais água. O aumento da capacidade de retenção de água aumenta a suculência da

carne devido ao relaxamento das fibras musculares, proporcionando a obtenção de carnes mais macias. A adição de sais reduz a relação água/proteína, entretanto, a maior capacidade de retenção de água, eleva esta relação a valores similares aos normais. O efeito da adição de sais, como NaCl, KCl e MgSO₄, no aumento da WHC de carnes tem sido estudado através das espectroscopias MIR e NIR, com o auxílio de métodos quimiométricos, como a PCA (análise de componentes principais), embora não em situações de fraude.^{20,21} O aumento da WHC está ligado a mudanças na conformação das proteínas, provocadas pela adição destes ingredientes. Mais recentemente, a espectroscopia de imagem no infravermelho médio foi aplicada neste tipo de estudo, permitindo um aumento na capacidade de extrair informações, uma vez que esta técnica torna possível estimar a distribuição de componentes químicos na superfície de amostras complexas.²² Assim como alguns sais, a carragena também tem a capacidade de aumentar a WHC de carnes.²³ Entretanto, praticamente não existem estudos na literatura que relacionem a adição de sais ou carragena à fraudes em carnes.

O controle de qualidade da carne bovina consumida no Brasil baseia-se em métodos analíticos clássicos que determinam propriedades físico-químicas e microbiológicas da carne.²⁴ Entretanto, técnicas analíticas espectroscópicas ou cromatográficas podem também ser utilizadas no controle de qualidade da carne. As análises das propriedades físico-químicas da carne bovina *in natura* geram grandes quantidades de dados analíticos. Todavia, a análise individual/univariada dos resultados obtidos não é suficiente para a caracterização da adulteração, pois existem diversos fatores de variabilidade da carne que podem ocasionar alterações nos valores encontrados nas análises, como a idade do animal, sexo, sazonalidade, tipo de alimentação, peso, entre outros.¹⁶

Neste contexto, a análise multivariada surge como alternativa para a análise do conjunto de dados físico-químicos obtido nas análises de controle de qualidade da carne bovina *in natura*, a fim de propor métodos quimiométricos para a identificação de adulterações presentes em amostras suspeitas.

Assim, o presente trabalho tem por objetivo desenvolver métodos para a detecção e caracterização de fraudes econômicas em carne bovina *in natura* a partir da análise das propriedades físico-químicas e espectroscópicas das

amostras analisadas. Nesse desenvolvimento, serão utilizados métodos quimiométricos de análise exploratória/classificação não supervisionada, como a PCA, e de classificação supervisionada, como o PLS-DA. Serão desenvolvidas duas aplicações. Na primeira delas, amostras reais de carne obtidas da Operação “Vaca Atolada” serão analisadas a partir da determinação de cinco diferentes parâmetros físico-químicos e também de seus espectros MIR, obtidos usando um acessório de reflectância total atenuada (ATR, *Attenuated Total Reflectance*). Nesta aplicação, será usada a estratégia multivariada de fusão de dados (*Data Fusion*). Na segunda aplicação, amostras de carne propositadamente adulteradas com as substâncias encontradas na operação da polícia, supostamente utilizadas nessas adulterações, serão analisadas a partir de espectros MIR de suas purgas (líquido extraído da carne).

2. Revisão Bibliográfica

2.1 Carne bovina

A bovinocultura é uma atividade amplamente exercida e importante no agronegócio nacional. Detentor do segundo maior rebanho mundial, o Brasil assumiu a liderança de exportações da carne bovina mundial comercializando para mais de 180 países.²⁵

A carne bovina é o produto final obtido da bovinocultura, tratando-se de um tecido animal fonte de proteínas, minerais, ácidos graxos e vitaminas do complexo B.²⁻⁴ Embora a cadeia produtiva termine com a venda da carne ao consumidor final, a qualidade da carne só poderá ser determinada no momento de seu consumo. Desta forma, a preocupação com a qualidade e segurança quanto à origem, composição e validade da carne bovina tornam-se assuntos de destaque para controle e estudo.²⁶

Atributos sensoriais são relacionados à qualidade da carne, tais como qualidade visual, qualidade gustativa, qualidade nutricional e segurança.²⁶ Técnicas de manipulação da carne com o fim de conservação são utilizadas desde tempos remotos, como em 850 a. C. na época de Homero.²⁷ A fim de garantir segurança sanitária à comercialização de carne, o atual processamento de carcaças bovinas inclui a refrigeração após o abate de forma que a carne atinja temperatura de 7°C antes de seu processamento e comercialização. Entretanto, esta prática diminui a maciez da carne devido à excessiva contração do músculo.²⁸ Os métodos industriais de conservação de carne mais utilizados são os de congelamento, refrigeração, processamento térmico, desidratação e irradiação.²⁹

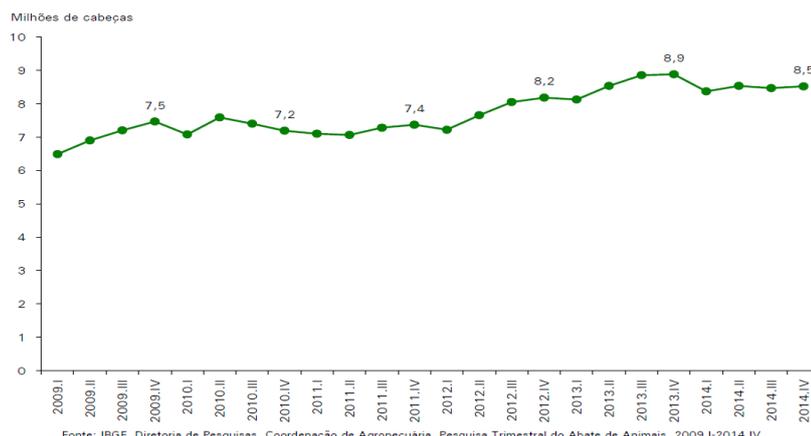
A carne bovina é amplamente consumida no Brasil e em diversas regiões do mundo. No Brasil, as principais formas de comercialização são a venda de carne fresca em frigoríficos e carne embalada comercializada em gôndolas de supermercados.

2.1.1 Abate de bovinos

O Brasil é detentor do segundo maior rebanho bovino no mundo.²⁵ A evolução do abate de bovinos no período de 2009 a 2014 pode ser visualizada na Gráfico 1, no qual se percebe que mais de 8,5 milhões de animais foram

abatidos no 4º trimestre de 2014. Este valor é cerca de 0,7% maior que o encontrado no 3º trimestre de 2014, entretanto, é 4,1% menor que o valor encontrado para o mesmo trimestre do ano anterior. Cabe ressaltar que a oferta restrita de reposição de animais e a seca prolongada a partir do final de 2013 foram fatores que contribuem para esta variação negativa.³⁰

Gráfico 1: Evolução do abate de bovinos por trimestre - Brasil - trimestres 2009-2014



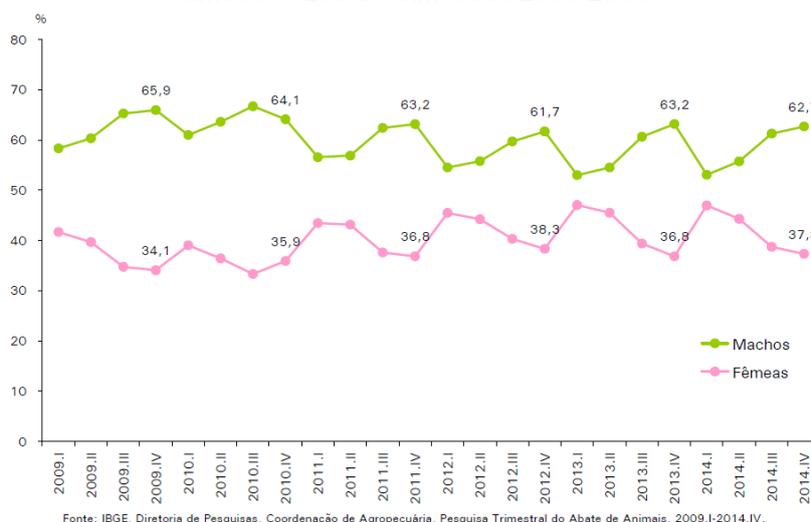
Como há grande diferença de pesos da carcaça, a quantidade acumulada de carne produzida no período de 2009 a 2014 pode ser verificada no Gráfico 2. Percebe-se neste gráfico que a quantidade de carne bovina obtida é 0,9% maior que a obtida no trimestre anterior, entretanto, seguindo o mesmo comportamento da série de abate de bovinos, em relação ao 4º trimestre de 2013 a produção foi menor em 3,7%.³⁰

Gráfico 2: Evolução do peso acumulado de carcaças de bovinos por trimestre - Brasil - trimestres 2009-2014



Dentre os bovinos abatidos neste período, verifica-se que o número de machos é maior em relação ao número de fêmeas, conforme pode ser verificado no Gráfico 3.

Gráfico 3: Evolução da participação de machos e fêmeas no abate total de bovinos por trimestre - Brasil - trimestres 2009-2014



Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Agropecuária, Pesquisa Trimestral do Abate de Animais, 2009.I-2014.IV.

A grande quantidade de carne bovina produzida no Brasil é utilizada para consumo interno e, também, para consumo externo. A carne bovina *in natura* produzida é exportada, principalmente para Hong Kong (23,2%), Egito (19,0%), Rússia (18,6%), Venezuela (16,0%), Chile (3,8%), Itália (3,0%), Argélia (1,7%), Holanda (1,5%), Angola (1,4%) e Irã (1,4%). Estes dez países respondem por 89,6% da carne bovina exportada pelo Brasil no 4º trimestre de 2014. Na Tabela 1 são mostrados valores de abate de bovino e exportação de carne bovina *in natura* entre 2013 e 2014.³⁰

Tabela 1: Abate de bovinos e exportação de carne bovina *in natura* - Brasil – trimestres selecionados de 2013 e 2014

Bovinos abatidos, produção de carcaça e exportação de carne bovina	2013	2014		Variação (%)	
	4º trimestre (1)	3º trimestre (2)	4º trimestre (3)	(3/1)	(3/2)
Bovinos abatidos ¹ (cabeças)	8 888 188	8 469 539	8 525 484	-4,1	0,7
Carcaças produzidas ¹ (t)	2 137 780	2 040 059	2 058 951	-3,7	0,9
Carne <i>in natura</i> exportada ² (t)	333 080	317 461	312 608	-6,1	-1,5
Faturamento da exportação ² (milhões de US\$)	1 538	1 547	1 519	-1,2	-1,9
Preço médio (US\$ FOB/t)	4 616	4 874	4 858	5,2	-0,3

Fonte: ¹Pesquisa Trimestral do Abate de Animais, IBGE e ²Secretaria de Comércio Exterior, Secex/MDIC.

2.1.2 Composição da carne bovina

A composição média dos cortes bovinos analisados neste trabalho é apresentada na Tabela 2. Percebe-se que não há grandes variações na composição química média dos cortes em análise. Tais variações podem estar relacionados a fatores como variabilidade genética, idade, sexo e alimentação a que o animal foi submetido, entre outros.

Tabela 2: Composição de cortes bovinos por 100 gramas de parte comestível³¹

Descrição dos alimentos	Umidade (%)	Proteína (g)	Lipídeo (g)	Cinzas (g)	Cálcio (mg)	Fósforo (mg)	Sódio (mg)
Contra-filé de costela, cru	66,4	19,8	13,1	1,0	3,0	164,0	39,0
Coxão duro, sem gordura, cru	69,8	21,5	6,2	1,1	3,0	189,0	49,0
Coxão mole, sem gordura, cru	68,6	21,2	8,7	1,0	3,0	175,0	61,0
Lagarto, cru	71,0	20,5	5,2	1,1	3,0	185,0	54,0
Músculo, sem gordura, cru	72,4	21,6	5,5	1,0	4,0	162,0	66,0
Patinho, sem gordura, cru	72,9	21,7	4,5	1,0	3,0	170,0	49,0
Miolo de alcatra, sem gordura, cru	69,5	21,6	7,8	1,0	3,0	165,0	43,0

2.1.3 Fatores de variabilidade da carne bovina²⁸

A obtenção de carne com menor quantidade de gordura, ou seja, carne mais dura ocorre a partir da associação de diversos fatores, como o abate de animais de determinadas raças, animais criados em pasto e com idade elevada. O maior teor de gordura na carne e, conseqüentemente, maior maciez, está associado a fatores *ante-mortem* e *post-mortem*. Entre os fatores *ante-mortem* destacam-se os aspectos de raça ou genótipo, alimentação, idade, sexo, aplicação de promotores de crescimento e manejo pré-abate. Entre os fatores *post-mortem* destacam-se efeitos como o resfriamento.

A principal raça bovina utilizada na produção de gado de corte é a raça zebu, puro ou mestiço. A idade de animais prontos para o abate varia em torno dos 30-36 meses, dependendo do tipo de manejo ao qual estão sujeitos. Entretanto, a carne dos zebuínos é considerada uma carne dura, mesmo que eles tenham sido abatidos precocemente e possuam boa cobertura de gordura. Trata-se de uma questão genética e não apenas do manejo de animais desta raça. No início dos anos 90, estudos realizados indicaram que a carne de

animais com mais de 50% de sangue zebu é menos gordurosa e mais dura. Isto se deve a diversos fatores, como acabamento, grau de marmoreio (acumulação de gordura intramuscular), tipo e quantidade de tecido conjuntivo, além da degradação enzimática das proteínas miofibrilares.

O sexo é outro fator que interfere nas propriedades físico-químicas da carne. A textura (ou granulometria) tende a ser mais grosseira em animais machos, relacionada às proteínas presentes nos músculos, tais como as do tecido conjuntivo (colágeno e elastina), miofibrilares (miosina) e as proteínas sarcoplasmáticas. Com o aumento da idade, verifica-se que há a diminuição da solubilidade do colágeno, logo, a carne torna-se mais dura. Três sistemas enzimáticos destacam-se, também, quanto à variabilidade em propriedades da carne bovina. O principal mecanismo responsável é o das calpaínas, que são tiolproteases produzidas pelo músculo e conferem maior maciez à carne. Outro sistema é o das catepsinas, que atuam em pH mais baixos, em torno de 6,0, degradando proteínas miofibrilares e colágeno. O último sistema é o complexo multicatalítico de proteases, que atua sobre os peptídeos conferindo à carne maior maciez.

Relacionada aos fatores *post-mortem*, a refrigeração exerce importante papel sobre a alteração de propriedades da carne bovina. A fim de evitar a contração excessiva da carne animal submetida a resfriamento logo após o abate, são utilizadas as seguintes técnicas: refrigeração retardada, estimulação elétrica das carcaças, adição de enzimas amaciantes, infusão de íons cálcio, pressurização do músculo e maciez por estiramento. Na refrigeração retardada, a carcaça é mantida em temperatura mais alta até que ocorra o *rigor mortis* (endurecimento do músculo após reações químicas que o transforma em carne).³² Em seguida, a carcaça é submetida a resfriamento abaixo de 10°C. Já a estimulação elétrica provoca a fragmentação mecânica de miofibrilas e fibras musculares, aumentando, assim, a solubilidade do colágeno e, como resultado, aumentando também a maciez da carne, pois ligações cruzadas termoestáveis do colágeno dos músculos foram destruídas. A fim de ativar as calpaínas, segue-se o procedimento de infusão de cálcio, ainda que o efeito de contração muscular mediada pelo cálcio não tenha sido sanado.

Diversos fatores ou associação de fatores podem ocasionar diferenças nos valores encontrados para propriedades físico-químicas da carne bovina.

Entretanto, o controle de qualidade deste alimento deve levar em conta valores médios disponíveis em materiais de referência.

2.2 Controle de qualidade de alimentos de origem animal no Brasil

A qualidade de produtos agropecuários é pautada na identificação da origem, composição química, propriedades físicas adequadas, fatores sensoriais satisfatórios, além de garantia de segurança e higiene dos alimentos. A adição fraudulenta de substâncias ou ingredientes mais baratos aos produtos alimentares proporciona perdas econômicas e riscos à saúde. Desta forma, autoridades intensificaram esforços a fim de resguardar a legalidade dos produtos, através da identificação e prevenção de fraudes.³³

A produção e comercialização de alimentos de origem animal são regulamentadas por legislações de inspeção sanitária e controle de qualidade de alimentos. O primeiro regulamento de inspeção sanitária de produtos de origem animal no Brasil foi estabelecido em 1909 com a promulgação do Decreto 7.622/1909, que instituiu a “Diretoria de Indústria Animal”. Modificações na regulamentação oficial foram realizadas até que em 1952 foi normatizada a inspeção de todas as etapas do processo produtivo e de comercialização de alimentos a partir da promulgação do Decreto 30.691/1952, que instituiu o “Regulamento de Inspeção Industrial e Sanitária de Produtos de Origem Animal – RIISPOA”, o qual disserta acerca dos aspectos a serem inspecionados e das responsabilidades pela inspeção.³⁴

A inspeção de todas as etapas do processo de produção e comercialização de produtos alimentícios tem por objetivo garantir a qualidade dos produtos a serem comercializados. Segundo Guenther, o termo qualidade é definido nos meios técnicos como o conjunto de características que atribuem o valor comercial dos produtos, tais como a textura, o peso, forma, tamanho, cor, odor, sanidade entre outros aspectos.¹

A fim de garantir a qualidade dos produtos comercializados, diversos exames devem ser realizados em laboratórios registrados e credenciados junto aos órgãos responsáveis pela inspeção e fiscalização, conforme citado a seguir:

Art. 870 - Os produtos de origem animal prontos para consumo, bem como toda e qualquer substância que entre em sua elaboração, estão sujeitos a exames tecnológicos, químicos e microbiológicos. (Ministério da Agricultura, Regulamento de Inspeção Industrial e Sanitária de Produtos de Origem Animal, 1952, p. 139).³⁵

De acordo com o Código de Defesa do Consumidor, há uma preocupação com a segurança no consumo de serviços e produtos, dentre os quais se encaixam os alimentos de origem animal. Estes alimentos não devem oferecer riscos à saúde humana e não podem conter substâncias não descritas em seus respectivos rótulos:

Da Proteção à Saúde e Segurança

Art. 8º Os produtos e serviços colocados no mercado de consumo não acarretarão riscos à saúde ou segurança dos consumidores, exceto os considerados normais e previsíveis em decorrência de sua natureza e fruição, obrigando-se os fornecedores, em qualquer hipótese, a dar as informações necessárias e adequadas a seu respeito.

Parágrafo único. Em se tratando de produto industrial, ao fabricante cabe prestar as informações a que se refere este artigo, através de impressos apropriados que devam acompanhar o produto.

Art. 9º O fornecedor de produtos e serviços potencialmente nocivos ou perigosos à saúde ou segurança deverá informar, de maneira ostensiva e adequada, a respeito da sua nocividade ou periculosidade, sem prejuízo da adoção de outras medidas cabíveis em cada caso concreto. (Lei nº 8.078, de 11 de setembro de 1990, Código de Defesa do Consumidor).³⁶

Entretanto, a adição de substâncias exógenas é registrada desde épocas remotas em que o homem adicionava sal aos alimentos para melhorar o gosto e conservá-los por mais tempo. Muitas substâncias adicionadas são consideradas normais, entretanto torna-se difícil a conscientização daqueles que utilizam substâncias químicas não informadas no aperfeiçoamento da produção, industrialização e conservação de alimentos quanto aos benefícios ou malefícios dessa prática.¹ Assim, a adição de substâncias químicas com o intuito comercial de aumentar lucros torna-se um processo conhecido como fraude ou adulteração, ou seja, procedimentos que se distanciam das

características normais, como a alteração de peso e preço, utilizando-se de artifícios sem consentimento oficial e que não são aceitos universalmente.¹

O controle de qualidade da carne bovina é baseado, essencialmente, em análises físico-químicas. Entretanto, diversas técnicas instrumentais, como espectroscopia e cromatografia, têm sido utilizadas com êxito na identificação de componentes presentes na carne, quantificação e controle de qualidade. Aliadas ao uso de métodos quimiométricos, tais técnicas, principalmente as baseadas em espectroscopia molecular, são promissoras no desenvolvimento de modelos de identificação de padrões ou classificação exatos e robustos, que podem apresentar inúmeras vantagens, como maior rapidez, menor custo, mínimo tratamento da amostra, e análises mais limpas, sem gasto de reagentes ou geração de resíduos.

2.3 Espectroscopia na região do infravermelho

A espectroscopia na região do infravermelho é uma técnica simples, rápida e não destrutiva (quando usada em conjunto com acessórios de reflectância), que pode ser utilizada na caracterização de compostos orgânicos. Assim como outros tipos de espectroscopia, ela se baseia em interações entre as moléculas ou átomos e a radiação eletromagnética, provocando vibrações de acordo com a amplitude das ligações covalentes existentes na molécula.³⁷

A primeira citação acerca da radiação infravermelha foi realizada por Herschel, em 1800, quando realizou experimentos de refração da luz solar através de um prisma e percebeu o aquecimento de termômetros em presença de luz visível e, também, em presença de luz além dos raios vermelhos, fora do espectro visível.³⁸

A partir do avanço tecnológico dos equipamentos disponíveis para a realização das análises, a espectroscopia na região do infravermelho tornou-se uma técnica analítica amplamente utilizada pelos Químicos na elucidação das estruturas químicas de compostos orgânicos.³⁹ Na década de 1980, a maioria dos equipamentos disponíveis para análises por infravermelho médio eram do tipo dispersivo, fundamentados na utilização de redes de difração e eram, muitas vezes, construídos pelos próprios pesquisadores.^{39,40} Atualmente, a maioria dos espectrofotômetros de infravermelho são do tipo Transformada de

Fourier. Este tipo de espectrômetro ampliou a utilização de equipamentos de análises na região do infravermelho devido à obtenção de melhor relação sinal/ruído e menores limites de detecção.⁴⁰

Na espectroscopia no infravermelho, a frequência da radiação absorvida é resultado da frequência da vibração molecular que ocasiona o processo de absorção.⁴⁰ Tendo como fundamento a teoria corpuscular, a radiação eletromagnética que provoca as vibrações é quantizada, envolvendo *fótons* de energia específica, ou seja, pacotes de energia.^{37,39} A radiação infravermelha não possui energia suficiente para provocar transições eletrônicas,⁴⁰ entretanto, a sua absorção ocorre em frequências específicas para cada grupo funcional, devido ao arranjo característico de átomos e ligações presentes na molécula.³⁷

A radiação eletromagnética na região do infravermelho absorvida por uma molécula na faixa de números de onda entre 10000 e 100 cm^{-1} é transformada em energia vibracional, enquanto na região abaixo de 100 cm^{-1} ela é convertida em transições rotacionais. Entretanto, a região de maior interesse para a caracterização de compostos é a compreendida entre 4000 e 400 cm^{-1} .^{39,40}

A região do infravermelho no espectro eletromagnético corresponde, aproximadamente, a faixa entre 700 a 50000 nm (14290 a 200 cm^{-1}).³⁹ A região compreendida entre 2500 – 50000 nm (4000 – 200 cm^{-1}) corresponde à região do espectro eletromagnético denominada Infravermelho Médio, sendo amplamente utilizada na elucidação de compostos orgânicos. Outra região de interesse na análise de compostos orgânicos é a do Infravermelho Próximo, que está situada entre 780 – 2500 nm (12800 – 4000 cm^{-1}).^{39,40}

Tabela 3: Regiões espectrais no Infravermelho

Região	Intervalo de Comprimento de Ondas (λ), μm	Intervalo de Número de Ondas ($\bar{\nu}$), cm^{-1}	Região de Frequência (ν)
Próximo	0,78 a 2,5	12.800 a 4.000	$3,8 \times 10^{14}$ a $1,2 \times 10^{14}$
Médio	2,5 a 50	4.000 a 200	$1,2 \times 10^{14}$ a $6,0 \times 10^{12}$
Distante	50 a 1.000	200 a 10	$6,0 \times 10^{12}$ a $3,0 \times 10^{11}$

Fonte: Skoog, D. A.; Holler, F. J.; Nieman, T. A. *Princípios de análise instrumental*⁴⁰

2.3.1 Instrumentos para o Infravermelho

Três tipos de equipamentos para a análise no infravermelho estão disponíveis no mercado: Instrumentos dispersivos; baseados na Transformada de Fourier; e fotômetros não-dispersivos (usados para amostras gasosas, e por isso não discutidos a seguir). Os espectrofotômetros mais utilizados até meados de 1980 eram os dispersivos. Entretanto, com o avanço tecnológico de computadores e velocidade de processamento de dados, os equipamentos amplamente utilizados atualmente são os espectrômetros com Transformada de Fourier. A Fig. 1 representa um esquema de componentes básicos de um espectrofotômetro de infravermelho:

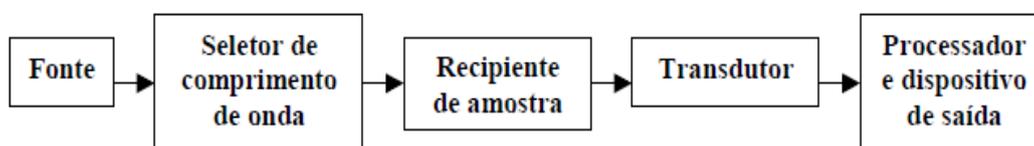


Figura 1: Componentes básicos de um equipamento que opera na região do infravermelho.

2.3.1.1 Instrumentos dispersivos

A maioria dos equipamentos dispersivos disponíveis são equipamentos de feixe duplo que utilizam redes de difração. Há preferência pela utilização deste tipo de equipamento em relação ao feixe simples, pois água e gás carbônico atmosféricos são absorvidos nessa região espectral, podendo gerar interferências, e a utilização de um feixe de referência é capaz de compensar tal absorção. A utilização de feixe duplo é importante, também, devido à baixa intensidade das fontes de infravermelho, à baixa sensibilidade dos transdutores e à necessidade de amplificação do sinal.^{40,43} Os equipamentos dispersivos apresentam tempo de resposta lento dos transdutores, logo, necessitam de moduladores com velocidade lenta e de baixa frequência, que permitam a distinção de bandas espúrias. Após a dispersão por um prisma, a radiação incide no transdutor que converte em sinal elétrico os feixes incididos.⁴⁰ O grande problema dos equipamentos dispersivos reside no monocromador, que possui fendas de entrada e saída muito pequenas, as quais limitam a radiação que chega ao detector, prejudicando a leitura de amostras que exijam maior sensibilidade analítica.⁴³ Este tipo de equipamento foi amplamente utilizado até

a década de 1960, quando os equipamentos com transformada de Fourier tornaram-se mais baratos e confiáveis.⁴³

2.3.1.2 Espectrômetros com Transformada de Fourier

Espectrofotômetros com Transformada de Fourier (FTIR) baseiam-se na interferência da radiação que passa por dois feixes para a criação de um interferograma, sendo que a medição se deve à mudança do comprimento da trajetória entre os feixes. A interconversão entre trajetória e frequência é realizada pela ferramenta matemática Transformada de Fourier.⁴³ O espectrômetro FTIR apresenta um sistema ótico que altera a tonalidade do sinal policromático permitindo a análise pelo conteúdo intensidade/frequência sem dispersão. A Transformada de Fourier fundamenta-se na transformação em função cosseno da relação entre a distribuição da radiação incidente no interferômetro e o sinal emitido pelo detector ao receber a radiação. Para isto, utiliza-se o algoritmo de Cooley-Tukey.⁴⁴

Estes espectrômetros apresentam maior sensibilidade na conversão das frequências de infravermelho.⁴⁴ O tipo mais comum de interferômetro é o de Michelson, que consiste em dois espelhos planos perpendiculares, sendo que um deles pode mover-se perpendicularmente ao plano. Os materiais utilizados como separadores dos feixes são escolhidos de acordo com a região espectral a ser analisada. O brometo de potássio é usado no infravermelho próximo e médio, enquanto películas orgânicas são utilizadas no infravermelho distante.

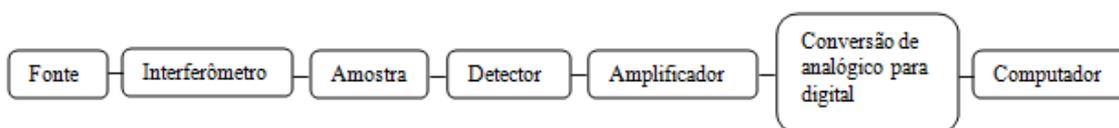


Figura 2: Componentes básicos de um espectrômetro FTIR.⁴³

Diversas vantagens são apresentadas pelo espectrômetro FTIR, dentre elas pode-se citar, primeiramente, a vantagem de Jaquinot ou vantagem de transporte, que é relacionada à utilização de poucos elementos óticos e nenhuma fenda, permitindo uma melhora na relação sinal-ruído devido à maior potência de radiação que incide no detector. Outras vantagens são a alta resolução e reprodutibilidade do comprimento de onda, a obtenção de dados para um espectro inteiro em segundos, devido ao fato de todos os elementos

atingirem simultaneamente o detector, além de o interferômetro estar livre de radiações espúrias.⁴⁰ Existe, ainda, a vantagem de Fellgett (multiplex) que reflete a melhoria na resolução quando da análise de uma grande número de elementos simultaneamente.⁴³

2.3.2 Aplicações da Espectrometria no Infravermelho

A espectrometria na região do Infravermelho é uma técnica versátil e rápida utilizada na identificação, caracterização e quantificação de diversos compostos moleculares. Sua aplicação está definida de acordo com a região espectral a ser analisada. Historicamente, a região mais amplamente utilizada é a região do infravermelho médio. Até o começo da década de 1990, essas aplicações eram predominantemente qualitativas e univariadas, limitadas à identificação de compostos orgânicos e a atribuições espectrais. No entanto, o uso quantitativo desta técnica teve um grande impulso com o desenvolvimento da análise multivariada e da quimiometria, aliado ao avanço da instrumentação analítica, principalmente com a difusão do uso de acessórios de reflectância, que permitiram uma maior versatilidade na análise de amostras sólidas e líquidas. A região do infravermelho próximo (NIR) também se beneficiou destes mesmos avanços. Como os picos do NIR são menos seletivos e mais sobrepostos, praticamente não é possível desenvolver métodos analíticos qualitativos ou quantitativos univariados baseados nesta região espectral. Por isso, a difusão da espectrometria NIR nos últimos anos, reforçada pelo desenvolvimento da quimiometria, foi marcante.⁴²

2.3.2.1 Espectrometria no infravermelho médio

A espectrometria de absorção no infravermelho médio é uma técnica fundamental para a determinação estrutural de espécies orgânicas e bioquímicas. Com o aparecimento de equipamentos com registradores mais baratos e de fácil manuseio a partir fim da década de 1950, estima-se que o tempo necessário para a caracterização estrutural de um composto foi reduzido cerca de dez, cem ou, até mesmo, mil vezes. A análise de dados dos espectros de infravermelho médio baseia-se na análise das bandas atribuídas a grupos funcionais mais prováveis. Em seguida, compara-se o espectro da amostra desconhecida com espectros de compostos puros que possuam todos os

grupos identificados na etapa anterior. Uma região de extrema importância é a região de impressão digital (*fingerprint*), entre 1200 e 600 cm^{-1} .⁴⁰

Encontrando grandes aplicações na análise de substâncias sólidas, a espectrometria de reflexão no infravermelho médio fornece espectros com as mesmas informações que os espectros de absorção, podendo ser utilizada em análises qualitativas e quantitativas.⁴⁰ Existem quatro tipos de reflexão da radiação no infravermelho: reflexão especular; reflexão interna; reflexão difusa e reflexão total. Destes, os dois últimos são mais amplamente utilizados e serão destacados a seguir.

2.3.2.1.1 Reflexão Difusa

A espectrometria de reflexão (ou reflectância) difusa no infravermelho com transformada de Fourier (em inglês, DRIFTS – *diffuse reflectance infrared Fourier transform spectrometry*) é uma técnica utilizada para a obtenção de espectros diretamente de amostras pulverizadas, sem pré-tratamento. O seu uso intensificou-se após a década de 1970, com a disponibilização de equipamentos com transformada de Fourier, pois a radiação reflexiva possui baixa intensidade para ser medida em equipamentos dispersivos. A Reflexão Difusa acontece quando a radiação incide sobre a superfície finamente dividida de um pó e ocorre a reflexão especular em cada superfície plana. Entretanto, as partículas estão organizadas de forma aleatórias e a radiação é refletida em todas as direções.³³ Acessórios baseados neste tipo de reflexão são largamente usados no infravermelho próximo. Eles também são usados no infravermelho médio, embora, neste caso, acessórios baseados em outro tipo de reflexão, descrita na seção seguinte, sejam preferidos.

2.3.2.1.2 Reflexão Total Atenuada⁴⁰

A reflectância total atenuada (ATR, *Attenuated Total Reflection*) é a técnica de reflexão mais usada no infravermelho médio, apresentando grande versatilidade na análise de amostras sólidas e líquidas, podendo ser empregada em amostras de difícil manipulação, como as de solubilidade limitada e pós. Este tipo de reflexão baseia-se na passagem de um feixe de radiação de um meio mais denso para outro meio menos denso. Ocorre, então, a reflexão de uma fração da radiação incidente que aumenta com o ângulo de

incidência; acima de um ângulo crítico,⁴⁰ função dos índices de refração das duas superfícies,⁴³ ocorre a reflexão total. A Fig. 3 apresenta um esquema típico de uma célula de ATR.

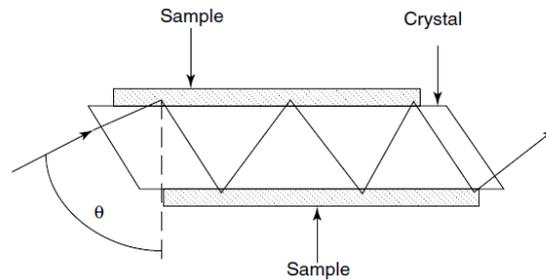


Figura 3: Esquema típico de uma célula de reflectância total atenuada.⁴³

O meio mais denso é o cristal de ATR, que faz parte do acessório, e o meio menos denso é a amostra analisada. Os materiais mais usados como cristal de ATR são seleneto de zinco e diamante. Experimentalmente, percebeu-se que a radiação penetra um pouco no meio menos denso antes da reflexão, sendo variável a profundidade de penetração.⁴⁰ A profundidade da penetração é uma função do comprimento de onda, λ , do índice de refração do cristal, n_2 , e do ângulo de incidência da radiação, θ . A profundidade de penetração, d_p , é dada pela equação 1, em que n_1 é o índice de refração da amostra:

$$d_p = \frac{\left(\frac{\lambda}{n_1}\right)}{\left\{2\pi\left[\text{sen}\theta - \left(\frac{n_1}{n_2}\right)^2\right]^{1/2}\right\}} \quad (1)$$

A radiação penetrante recebe a denominação de onda evanescente. A reflectância total atenuada ocorre quando a onda evanescente é absorvida pelo meio menos denso, sendo atenuada nos comprimentos de onda de absorção, em função da composição química da amostra. A vantagem da utilização de ATR é que os espectros são rapidamente obtidos em grande variedade de amostras sem necessidade de pré-tratamentos, além de permitir a análise de soluções aquosas, desde que o cristal não seja solúvel em água.

2.4 Quimiometria

O surgimento das atividades de pesquisa em Quimiometria teve início no começo da década de 1970 com a chegada dos microprocessadores e computadores aos laboratórios químicos, permitindo e ampliando a utilização da estatística multivariada em análises químicas.⁴⁵⁻⁴⁷ Na década de 1980 surgiram equipamentos analíticos mais acessíveis, com microcomputadores acoplados, que possibilitaram a obtenção dos resultados analíticos de maneira mais ágil e rápida. Equipamentos menores e com interface computacional surgiram, então, na década de 1990, sendo amplamente difundidos em laboratórios de análises químicas.⁴⁷ Com isto, os químicos analíticos passaram a gerar cada vez maiores quantidades de dados de maneira mais rápida. A necessidade de tratar esses dados está intimamente ligada ao desenvolvimento da Quimiometria.

Paralelamente ao desenvolvimento dos equipamentos analíticos, na década de 1980, quando a Quimiometria foi introduzida no Brasil, surgiu a demanda por cursos específicos para empresas e universidades, além da necessidade de publicação de bibliografia especializada em português. Assim, inúmeras pesquisas aplicando métodos quimiométricos começaram a ser desenvolvidas, com destaque a partir da década de 1990, quando os computadores tornaram-se mais acessíveis e houve um aumento significativo no número de publicações devido à grande vantagem da análise multivariada poder tratar simultaneamente diversas variáveis medidas para uma determinada amostra.⁴⁵

O campo de aplicação da Quimiometria é enorme abrangendo, por exemplo, estudos de cinética e equilíbrio químico em físico-química, estudos de reações e estruturas de compostos em química orgânica, estudos de relação estrutura-atividade em química teórica, e diversas técnicas analíticas, com maior destaque para as técnicas de espectroscopia molecular, tais como espectroscopia de absorção molecular no UV/vis, no infravermelho próximo – NIR, e no médio – MIR, espectroscopia Raman e espectrofluorimetria. Estas técnicas naturalmente produzem sinais analíticos sobrepostos para matrizes complexas e vem sendo aplicadas junto com métodos quimiométricos a

diversas áreas, tais como arqueologia, bioanalítica, ciência forense, processos industriais, análise farmacêutica, ciência de alimentos, dentre outras.

Segundo a Sociedade Internacional de Quimiometria (ICS, *International Chemometrics Society*), esta é definida como a “disciplina química que utiliza métodos matemáticos e estatísticos, entre outros, empregando uma lógica formal para: 1) definir ou selecionar condições ótimas para procedimentos de análise e experimentos; e 2) obter máximo de informação química relevante a partir da análise de dados químicos”.⁴⁸ Genericamente, pode-se dividir a Quimiometria em três áreas principais, planejamento e otimização de experimentos, reconhecimento de padrões e calibração multivariada. Os métodos de reconhecimento de padrões podem ser utilizados para análise exploratória de dados, classificação de amostras e resolução de curvas.⁴⁵

2.4.1 Reconhecimento de Padrões

O reconhecimento de padrões é baseado na classificação de objetos a partir de medidas indiretas relacionadas às propriedades da amostra. Caso a informação prévia sobre a classificação da amostra não esteja disponível, procede-se ao reconhecimento de padrões de forma não supervisionada, buscando-se identificar padrões de agrupamentos.⁴⁵ Dentre as aplicações de reconhecimento de padrões, a análise exploratória (ou classificação não supervisionada) é utilizada para identificar a associação de padrões dentro de um conjunto de dados, estabelecendo relações entre os objetos e verificando a existência de amostras anômalas na matriz de dados analíticos. Os dois métodos mais utilizados para isso são a Análise de Agrupamentos hierárquica (HCA – *Hierarchical Cluster Analysis*) e Análise de Componentes Principais (PCA – *Principal Component Analysis*).⁴⁵

Métodos de reconhecimento de padrões também podem ser usados de forma supervisionada, quando forem conhecidas as atribuições de classes das amostras. Havendo um conjunto de amostras para as quais se sabe *a priori* a que classes pertencem, chamado conjunto de treinamento, é possível a utilização de um método supervisionado, no qual regras de classificação são fornecidas baseadas em medidas multivariadas.⁴⁶ Dentre os métodos mais comuns de classificação supervisionada, podemos citar a Análise discriminante linear - LDA (*Linear Discriminant Analysis*), Modelos independentes de

similaridade utilizando componentes principais - SIMCA (*Soft Independent Modelling of Class Analogies*), o método do K-ésimo vizinho mais próximo – KNN (*Kth Nearest Neighbour*) e a Análise discriminante por mínimos quadrados parciais - PLS-DA (*Partial Least Squares Discriminant Analysis*). Dentre esses métodos, o que vem tendo o maior aumento no número de aplicações encontradas na literatura quimiométrica nos últimos anos é o PLS-DA, devido à sua robustez e versatilidade.

2.4.1.1 Análise de Componentes Principais - PCA

Em 1901, Pearson introduziu a ideia de componentes principais, a qual, entretanto, só veio a ser mais amplamente utilizada décadas depois, com o desenvolvimento tecnológico dos computadores.⁴⁷ Atualmente, pode-se afirmar com segurança que a PCA é o método quimiométrico mais amplamente usado. A PCA permite a redução da dimensionalidade da matriz de dados preservando o máximo de informação. Ela permite, também, a interpretação dos resultados devido à determinação das variáveis mais significantes a partir de combinações lineares das variáveis originais, as chamadas Componentes Principais (CPs). Estas componentes descrevem as tendências relativas dos dados analíticos e são calculadas em ordem decrescente de importância. A primeira CP é descrita no eixo de maior variância da informação e a segunda CP é descrita no eixo ortogonal à primeira CP que apresente a maior variância residual. A terceira CP explica a maior variância residual e é ortogonal às duas primeiras, e assim por diante.⁴⁹ Desta forma, as CPs são independentes e descrevem variações no sistema de modo não-correlacionado,⁵⁰ permitindo a identificação das variáveis realmente importantes no conjunto de dados analisado e dos agrupamentos de amostras com características semelhantes.

A PCA, por ser uma técnica de reconhecimento de padrões, pode ser utilizada, também, para a análise de sinais analíticos sobrepostos.⁴⁶ O objetivo é obtenção de informação relevante a partir de dados analíticos brutos. Duas necessidades principais da Química se destacam: a interpretação das CPs quantitativamente e obtenção de padrões.⁴⁹

Para um conjunto de dados formado por espectros de infravermelho, por exemplo, a matriz de dados **X** será organizada de modo que cada linha

corresponda a uma amostra e cada coluna corresponda a uma variável, como mostrado na Fig. 4.

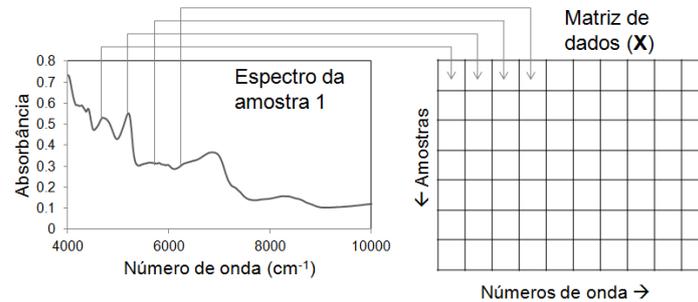


Figura 4: Representação da organização da matriz X a partir de dados espectroscópicos.⁵¹

A PCA realiza uma transformação matemática de mudança de sistema de coordenadas descrita pela seguinte equação, que é sua base estrutural:

$$X = T \cdot P^T + E \quad (2)$$

Onde T é a matriz de escores, que possui o mesmo número de linhas que a matriz original e o número de colunas igual ao número de CPs; P é a matriz de pesos (*loadings*) e tem o número de linhas igual ao de colunas na matriz original; E contém os resíduos. A decomposição dos dados descrita na Eq. 2 é feita através de uma operação de autovetores e o número de CPs é igual ao posto da matriz original, ou seja, ao número de vetores linearmente independentes da matriz. O posto da matriz é no máximo igual à sua menor dimensão (número de linhas ou colunas). Na prática, é importante definir o pseudo-posto da matriz original, que corresponde ao número de CPs significativas, que contêm variância sistemática relevante.

A matriz de escores consiste em uma série de vetores colunas, e a matriz de pesos é formada por uma série de vetores linhas. O primeiro vetor de escores e o primeiro vetor de pesos correspondem à primeira CP, conforme ilustrado na Fig. 5. A variância explicada por cada CP é proporcional ao quadrado do autovalor do autovetor equivalente, de acordo com a decomposição dos dados.

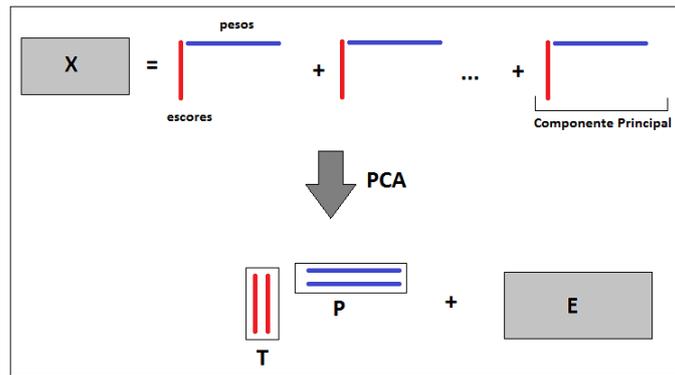


Figura 5: Representação da decomposição da matriz de dados em escores e pesos, de acordo com o modelo PCA

Escores e pesos têm importantes propriedades, entre elas a ortogonalidade, que significa independência, pois o produto escalar entre quaisquer dois vetores pesos ou escores é igual a 0.⁴⁹ Multiplicando o vetor \mathbf{t}_1 pelo vetor \mathbf{p}_1 é possível a conversão da matriz original na matriz \mathbf{X}_1 , que representa os dados reconstruído somente com a primeira CP:

$$\mathbf{X}_1 = \mathbf{t}_1 \mathbf{p}_1^T \quad (3)$$

A segunda CP pode ser calculada extraíndo o produto $\mathbf{t}_2 \mathbf{p}_2^T$ de \mathbf{X}_1 , e assim por diante, até o número de CPs significativo. Portanto, o modelo PCA pode ser representado pela seguinte soma de produtos vetoriais:

$$\mathbf{X}_1 = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \dots + \mathbf{t}_A \mathbf{p}_A^T + \mathbf{E} \quad (4)$$

A informação relevante está contida nas primeiras CPs e as restantes descrevem apenas variações residuais aleatórias ou informação sistemática irrelevante para a interpretação dos dados. A escolha do número de CPs é uma decisão crucial para a interpretação do modelo e deve levar em conta a quantidade de variância modelada e o conhecimento químico do sistema para não haver sobre-ajustes.⁵²

Finalmente, um aspecto fundamental na interpretação de modelos PCA é a observação conjunta de escores e pesos. Os primeiros indicam a composição de cada CP em termos das amostras, enquanto os segundos indicam essa mesma composição em termos de variáveis. Portanto, essa observação conjunta permite relacionar os agrupamentos de amostras com as variáveis mais discriminantes.

2.4.1.2 Análise discriminante por mínimos quadrados parciais - PLS-DA

O método dos Mínimos Quadrados Parciais (PLS – *Partial Least Squares*) foi inicialmente desenvolvido como uma ferramenta de calibração multivariada. Embora ele não tenha sido projetado para resolução de problemas de classificação e discriminação, seu algoritmo iterativo não-linear de mínimos quadrados parciais (NIPALS, *non-linear iterative partial least squares*) foi modificado para utilização com propósitos de classificação.^{52, 53}

A teoria do PLS-DA descreve dois tipos de abordagem. Na primeira, cada classe (variável dependente) é discriminada separadamente em relação às demais, de modo que um vetor binário \mathbf{y} designe valor 1 para amostras modeladas pertencentes à classe específica e valor 0 para amostras não pertencentes a esta classe. Este método é o PLS1-DA,⁵² conforme descrito na Fig. 6. O número 1 indica que a matriz de dados independentes (\mathbf{X}) é correlacionada com um arranjo unidimensional de cada vez (vetor \mathbf{y}).

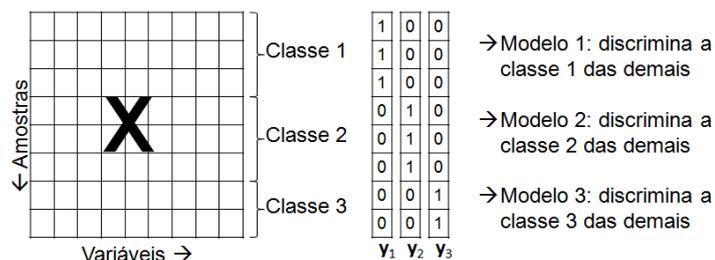


Figura 6: Representação do modelo PLS1-DA.⁵¹

A segunda abordagem refere-se ao método PLS2-DA, em que todas as classes são modeladas conjuntamente,⁵² conforme a Fig. 7. O número 2 indica que a matriz \mathbf{X} é correlacionada com um arranjo bidimensional (matriz \mathbf{Y}).

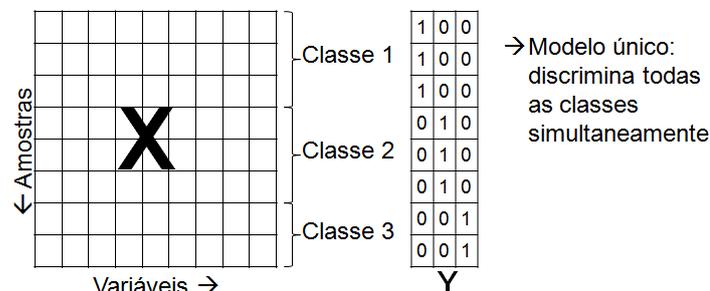


Figura 7: Representação do modelo PLS2-DA.⁵¹

A principal diferença do PLS-DA em relação ao PLS é o fato das variáveis dependentes representarem valores qualitativos e não quantitativos.

Desta forma, o bloco Y contém as denominadas variáveis categóricas (*dummy variables*), que definem se uma amostra pertence ou não a uma determinada classe (1 ou 0, respectivamente). O modelo, então, estimará um valor para Y , entretanto, este valor poderá não ser exatamente igual a 0 ou 1. Um limite de decisão (*threshold*) precisa ser definido, podendo arbitrariamente ser fixado em 0,5 ou, mais frequentemente, estimado pela teoria bayesiana. Valores de Y acima deste limite indicam que a amostra pertence à respectiva classe, enquanto valores abaixo dele indicam que a amostra não pertence à classe. Para cada classe, é definido um vetor y .⁵²

A validação cruzada é uma etapa essencial para a escolha do número de variáveis latentes (VLs) a ser usado na construção de qualquer modelo PLS ou PLS-DA. Nela, separa-se uma parte(ou apenas uma) das amostras de calibração e constrói-se o modelo com as restantes. Em seguida, estimam-se os erros de previsão para as amostras que foram separadas, utilizando diferentes números de VLs. Esse processo é repetido para outras amostras, até que todas tenham ficado de fora. Existem vários tipos de validação cruzada, dependendo de como a amostra, ou o subconjunto de amostras, é retirada dos dados. Os mais comuns são *leave-one-out*, blocos contínuos, subconjuntos aleatórios e venezianas (*venetian blinds*). Em toda esta dissertação será usada validação cruzada do tipo *leave-one-out*, que é recomendada para pequenos conjuntos de calibração, contendo normalmente até 20 amostras. Embora na primeira aplicação desenvolvida (amostras reais), o número de amostras de calibração seja um pouco maior (38), considera-se que o uso desta opção é ainda razoável. Em aplicações de calibração multivariada usando PLS e nas primeiras aplicações de classificação supervisionada usando PLS-DA, desenvolvidas no início da década passada, o critério para a escolha do número de VLs era o menor valor da raiz quadrada do erro médio quadrático de validação cruzada (RMSECV, *root mean square error of cross-validation*). No entanto, como nos modelos PLS-DA é mais importante o número de amostras corretamente classificadas do que a magnitude do erro na diferença do valor de previsão da variável categórica, o critério atualmente adotado é o menor erro de classificação de validação cruzada (CVCE, *cross-validation classification error*).⁵⁴

Segundo a Teoria Bayesiana, a probabilidade de uma amostra pertencer a uma classe A, dado um valor particular de y , pode ser calculado da seguinte forma:

$$P(A|y) = \frac{P(y|A)*P(A)}{[P(y|A)*P(A) + P(y|B)*P(B)]} \quad (5)$$

Onde $P(A)$ e $P(B)$ são as probabilidades de observação das classes A e B no futuro. Caso seja assumido que a possibilidade de observar as classes A e B é similar à quantidade de amostras no conjunto treinamento original, ou seja, as probabilidades bayesianas de um resultado ser previsto *a priori* como adulterado ($P(A)$) ou não ($P(B)$) são iguais a 0,5. Portanto, $P(A)$ se iguala a $P(B)$, obtendo-se as seguintes equações:

$$P(A|y) = \frac{P(y|A)}{P(y|A) + P(y|B)} \quad (6)$$

$$P(B|y) = \frac{P(y|B)}{P(y|A) + P(y|B)} \quad (7)$$

Assumindo que uma amostra em teste pertença a uma de duas classes possíveis, A ou B, pode-se afirmar que:

$$P(A|y) + P(B|y) = 1 \quad (8)$$

As duas distribuições - $P(A|y)$ e $P(B|y)$ – geralmente se cruzam em um único ponto, que é definido como o limite de decisão.⁵⁵

Após a definição do *threshold*, procede-se à detecção do número de amostras que não pertencem a uma classe e foram classificadas como pertencentes, ou seja, o número de Falsos Positivos (FP), e à detecção do número de amostras que pertencem a uma classe e foram classificadas como não pertencentes, ou seja, o número de Falsos Negativos (FN). A Taxa de Falsos Positivos (FPR – *False Positive Rate*) é definida como a razão entre FP e a soma de FP e o número total de amostras negativas conhecidas (TN):⁵⁵

$$FPR = \frac{FP}{FP + TN} \quad (9)$$

De forma análoga, a Taxa de Falsos Negativos (FNR – *False Negative Rate*) é definida como a razão entre FN e a soma de FN e o número total de amostras positivas conhecidas (TP):

$$FNR = \frac{FN}{FN + TP} \quad (10)$$

A validação analítica qualitativa do método pode ser obtida a partir do cálculo de figuras de mérito específicas, como sensibilidade, especificidade ou seletividade e taxa de eficiência.⁵⁵

A sensibilidade consiste na habilidade de detecção de amostras verdadeiramente positivas.⁵⁶ A taxa de sensibilidade (STR – *Selectivity Rate*) é calculada como a porcentagem referente à razão entre TP e a soma de TP e FN, conforme equação abaixo.⁵⁵

$$STR = \frac{TP}{TP + FN} * 100 \quad (11)$$

A especificidade ou seletividade consiste na habilidade de detecção de amostras verdadeiramente negativas.⁵⁶ A taxa de especificidade (SPR – *Specificity Rate*) é calculada como a porcentagem referente à razão entre TN e a soma de TN e FP, conforme demonstrado abaixo.⁵⁵

$$SPR = \quad (12)$$

A taxa de eficiência (EFR – *Efficiency Rate*) é calculada como a diferença entre o percentual total de resultados e a soma das taxas de falso positivo e falso negativo.

$$EFR = 100 - (FPR + FNR) \quad (13)$$

As figuras de mérito FPR, FNR e EFR expressam a veracidade da análise qualitativa, enquanto as figuras STR e SPR estão relacionadas à seletividade do método analítico.^{55,56} É importante ressaltar que o significado dos termos sensibilidade e especificidade, quando usados em análise qualitativa, é diferente de seu uso em análise quantitativa.

2.4.1.3 Algoritmo de Kennard-Stone

Para a separação das amostras em conjuntos de treinamento e teste foi usado o algoritmo de seleção Kennard-Stone. Este algoritmo, com base na distância euclidiana, define as duas amostras mais distantes entre si. Em seguida, ele seleciona a amostra mais distante das duas amostras inicialmente selecionadas. Este processo é repetido até que a quantidade de amostras a ser selecionada, previamente definida pelo analista, seja alcançada. Desta forma, o algoritmo Kennard-Stone garante sistematicamente a presença de amostras homogeneamente distribuídas no espaço amostral e representativas do modelo no conjunto de treinamento.⁵⁷ Em modelos de classificação supervisionada, como o PLS-DA, este algoritmo é aplicado em cada classe separadamente.

2.4.1.4 VIP scores

A interpretação espectral de modelos PLS e PLS-DA é usualmente realizada baseando-se na análise dos coeficientes de regressão do modelo. Entretanto, a interpretação não deve se basear somente nesses coeficientes, pois eles são dependentes da composição das amostras no conjunto de calibração/treinamento, da covariância implícita entre os componentes dessas amostras e da relação sinal/ruído dos dados analíticos.⁵⁸

Uma ferramenta mais robusta e confiável para a interpretação espectral são os gráficos de Importância das Variáveis na Projeção dos escores (*VIP scores – Variable Importance in Projection*). Esta ferramenta estima a importância de cada variável na projeção utilizada pelo modelo PLS ou PLS-DA através dos coeficientes em cada componente, juntamente com a significância de cada componente na regressão em módulo.^{58,59} Desta forma, os *VIP scores* são mais robustos que os coeficientes de regressão para identificar quais variáveis são mais significativas para o modelo preditivo. Os *VIP scores* podem ser utilizados, também, para a seleção de variáveis. O critério mais usado para a seleção da variável j é a média dos quadrados dos *VIP scores* serem maiores ou iguais a 1. A importância da variável de previsão j com base no modelo com h VLs pode ser calculada por:

$$VIP_j = \sqrt{\frac{J}{\sum_h SS(b_h t_h)} \sum_h w_{hj}^2 SS(b_h t_h)} \quad (14)$$

Onde J é o número de variáveis previstas, w_{hj} é a importância do peso da j ésima variável prevista no h ésimo fator PLS e $SS(b_h t_h)$ é a porcentagem de Y explicada pela h ésima VL.⁵⁹

2.4.1.5 Fusão de Dados – Data Fusion

Técnicas clássicas de determinação de propriedades de alimentos geralmente são destrutivas, exigem pessoal treinado e demandam muito tempo em sua realização devido ao tratamento das amostras. A fim de tornar o procedimento de análise mais eficaz, técnicas analíticas rápidas e minimamente destrutivas têm sido empregadas na análise de autenticidade de alimentos, como as técnicas espectroscópicas UV-Visível e Infravermelho (médio ou próximo). Estas técnicas permitem, ainda, identificar regiões de impressão digital (*fingerprint*) não-seletivas, que podem ser analisadas por

métodos multivariados fornecendo bons modelos quantitativos ou de classificação.⁶⁰

A Fusão de dados é uma estratégia de modelagem que combina dados provenientes de diferentes instrumentos analíticos, sensores ou variáveis físico-químicas discretas medidas isoladamente. Como um grande número de variáveis costuma ser gerado, o uso de métodos quimiométricos/multivariados é mandatório para potencializar a interpretação e modelagem do conjunto de dados sob análise.⁶¹ A união de dados provenientes de técnicas não específicas complementares pode fornecer modelos com maior capacidade preditiva ou classificatória, permitindo interpretações mais abrangentes e facilitando a descrição completa do produto ou matriz analisado.³³ Naturalmente, se uma técnica analítica isolada propiciar um modelo melhor, não há necessidade de fusão de dados. Mas, em vários casos, a existência de sinergia entre as informações oriundas de diversos tipos de dados leva ao melhor desempenho de modelos baseados em fusão de dados.

Desde o final dos anos 1980, a fusão de dados tem sido aplicada em áreas como engenharia e robótica,⁶² mas seu uso em química analítica é relativamente recente, concentrando-se na análise de alimentos, um tipo de matriz quase sempre bastante complexa.³³ As principais técnicas analíticas utilizadas em aplicações de fusão de dados de amostras de alimentos são baseadas em espectroscopia molecular, como infravermelho (NIR e MIR), Raman, UV-Visível, Fluorescência, RMN, e espectrometria de massas. Outras aplicações envolvem sensores (narizes ou línguas eletrônicas), análise de imagens digitais, propriedades físicas, químicas, de composição ou pureza, etc.

Essas aplicações envolvem modelos de classificação, principalmente supervisionada, e calibração multivariada. Os principais métodos quimiométricos usados são PCA, LDA, PLS-DA, máquinas de suporte de vetores (SVM - *Support Vector Machines*), redes neurais artificiais e PLS.³³ A fusão de dados pode ocorrer em três níveis: baixo, médio ou alto. No nível baixo de fusão de dados todas as variáveis são concatenadas diretamente, após as etapas de pré-processamento, em uma única matriz na qual o número de linhas é o mesmo número de amostras analisadas e o número de colunas é igual ao número de variáveis medidas em diferentes instrumentos. Após a

união destes dados, emprega-se o método quimiométrico adequado à classificação ou predição.^{33,63}

No nível intermediário (ou médio) de fusão de dados, primeiramente se extraem as informações relevantes de cada instrumento individualmente e, então, se concatena os dados extraídos (na maioria das vezes, os escores de modelos PCA, PLS-DA ou PLS) em uma única matriz que será submetida aos métodos de classificação/regressão. No alto nível de fusão de dados, também chamado de nível de decisão, os modelos de classificação/regressão de cada equipamento são construídos separadamente. Em seguida, todos os resultados dos modelos individuais são combinados para a obtenção do modelo final. Na maioria dos casos em que há comparação, a fusão de alto nível fornece resultados inferiores aos outros níveis.³³ A Fig. 8 apresenta a representação esquemática dos níveis de fusão de dados. Neste trabalho, será adotada a fusão de dados de baixo nível, mais simples e mais comum na literatura.

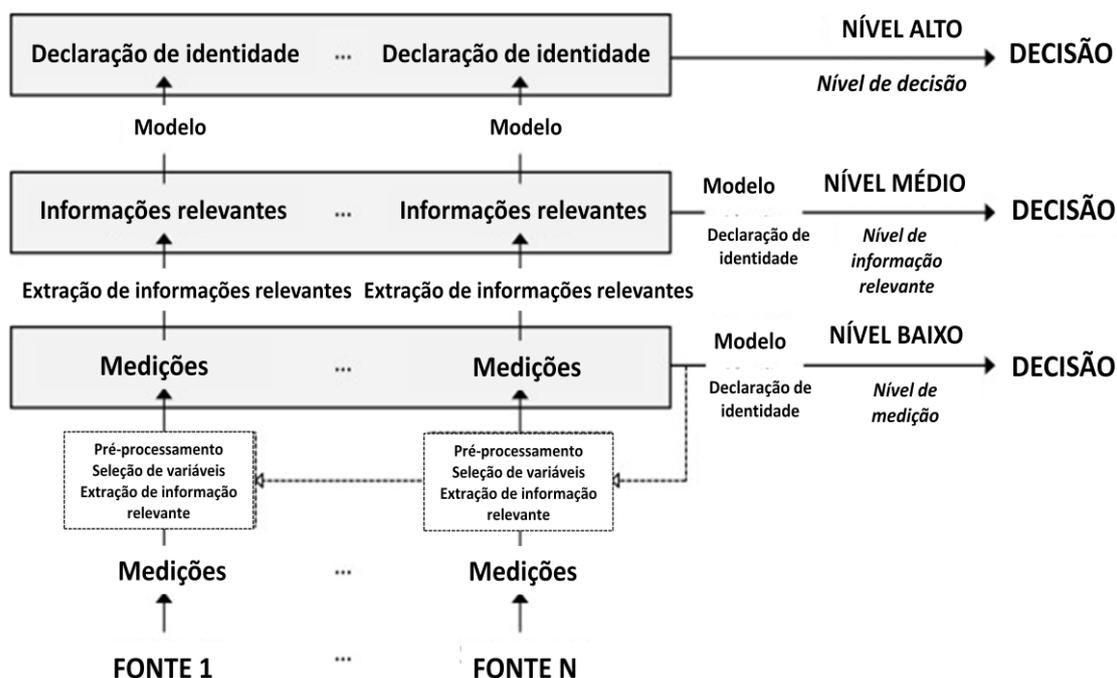


Figura 8: representação esquemática dos três níveis de fusão de dados.³³

Uma das matrizes mais estudadas por fusão de dados é o azeite de oliva. Casale e colaboradores usaram as técnicas espectroscópicas UV-Vis, NIR e MIR na construção de modelos individuais e de fusão de dados para determinar os teores de ácido oléico e linoléico, além de checar a autenticidade de amostras de uma região demarcada. Foram utilizados dois níveis diferentes

de fusão de dados: seleção de regiões seletivas dos espectros de cada técnica separadamente e, em seguida, combinação em uma única matriz (nível médio); e unificação das matrizes de dados espectrais seguida por seleção de variáveis relevantes (baixo nível). Os resultados encontrados demonstraram que houve uma sinergia no modelo de classificação ao utilizar a fusão de dados, melhorando a eficiência do modelo de classificação; entretanto, o mesmo não foi verificado para os modelos de regressão.⁶⁰ Em outro artigo, os mesmos autores desenvolveram modelos de classificação para a previsão da origem de azeites de oliva extra-virgem, usando espectros de massas, NIR e UV-Vis. O melhor modelo foi obtido com a fusão dos três tipos de espectro.⁶⁴ Também buscando classificar azeites de oliva, Pizarro e colaboradores utilizaram a fusão de dados de espectros na região do visível com descritores químicos, tais como acidez livre, índice de peróxidos e constantes de absorção no UV. Os resultados mostraram que a fusão dos dados foi adequada à discriminação dos azeites devido aos efeitos de sinergia entre as diferentes variáveis.⁶⁵

Biancolillo e colaboradores aplicaram fusão de dados oriundos de cinco técnicas instrumentais para caracterização da autenticidade de uma cerveja artesanal italiana: termogravimetria e espectroscopias nas regiões NIR, MIR, UV e visível. Foram construídos modelos de fusão de dados de baixo e médio nível e os melhores resultados de classificação foram obtidos com o nível médio.⁶⁶ Vera e colaboradores também construíram modelos de fusão para a classificação e descrição sensorial de cervejas a partir de sensores baseados em técnicas espectroscópicas: um nariz eletrônico baseado em espectros de massas, uma língua eletrônica baseada em espectros de MIR e um olho eletrônico baseado em espectros de UV-Vis.⁶⁷

Pode-se encontrar ainda na literatura artigos que desenvolveram modelos de fusão de dados para a identificação de adulterações inter-espécies em carne processada, o tipo de adulteração de maior interesse em carnes,¹¹ além de aplicações em outros tipos de matriz, como amostras biológicas em estudos de metaboloma,⁶¹ e na quantificação de propriedades de óleos isolantes.⁶³

Finalmente, destaca-se que neste trabalho será usada a fusão de dados espectroscópicos (MIR) e variáveis químicas medidas de maneira isolada. Embora a fusão entre dados espectroscópicos e conjuntos de variáveis

discretas seja menos comum, ela pode ser encontrada em vários trabalhos, como no já citado artigo de Pizarro classificando amostras de azeite,⁶⁰ na autenticação de amostras de trutas cruas e cozidas a partir de espectros NIR e parâmetros colorimétricos e mecânicos,⁶⁸ e na discriminação de amostras de queijo a partir da fusão de espectros NIR com variáveis químicas.⁶⁹

3 Materiais e Métodos

3.1 Amostras reais

Foram utilizadas 55 amostras reais apreendidas pela Polícia Federal na Operação Vaca Atolada, sendo 43 amostras suspeitas de adulteração e 12 amostras controle sabidamente não adulteradas (retiradas da própria carcaça animal no momento da desossa). Foram utilizadas peças de carnes de seis tipos de cortes, coxão mole, coxão duro, patinho, músculo, contra filé e lagarto, oriundas de cinco diferentes frigoríficos suspeitos de praticarem adulteração nas peças *in natura*. As amostras controle e suspeitas foram submetidas ao mesmo processamento.

3.1.1 A Operação Vaca Atolada

Em agosto de 2012, foi deflagrada em Belo Horizonte e região metropolitana a Operação Vaca Atolada da Polícia Federal. Amostras de carne bovina suspeitas de adulteração foram apreendidas, lacradas e enviadas para análise a fim de identificar e caracterizar seus principais componentes. A Fig. 9 apresenta a imagem de uma amostra de carne do tipo Contra Filé apreendida na Operação Vaca Atolada. O formato anômalo da peça de carne pode ser verificado pela determinação da linha tracejada (1ª foto) e pela presença de líquido congelado nas extremidades (indicado pelas setas na 2ª foto).



Figura 9: Peça de Contra-Filé apreendida na Operação Vaca Atolada com formato paralelepipedal e apresentando líquido congelado nas extremidades.

As amostras apreendidas foram processadas e subamostradas utilizando duas formas de pré-tratamento, o descongelamento controlado e a

trituração seguida de homogeneização. Tais tratamentos permitiram a obtenção de três tipos de subamostras, sendo: líquido resultante do descongelamento (purga), amostra de carne congelada totalmente triturada e homogeneizada, e amostra de carne descongelada sem purga totalmente triturada e homogeneizada.

As amostras controle foram coletadas nas instalações dos próprios frigoríficos sob investigação durante o momento em que a carne foi retirada da carcaça (desossa), podendo-se afirmar, portanto, que elas não sofreram injeção de qualquer tipo de substância química. Os mesmos procedimentos adotados para as amostras congeladas de carnes suspeitas de adulteração foram utilizados para amostras controle, tais como a retirada da gordura superficial das peças que passariam pelo pré-tratamento de trituração e homogeneização. A escolha de quais amostras seriam submetidas a cada tipo de pré-tratamento foi determinada de forma aleatória a fim de garantir uma amostragem mais diversificada.

Após pré-processamento e subamostragem, foram realizadas diversas análises para a identificação e quantificação de componentes na carne. Subamostras submetidas a um processo de avaliação controlado do descongelamento da peça cárnea foram colocadas em estufa incubadora, normalmente usada para determinação de DBO (demanda bioquímica de oxigênio), por um período de 48 horas, a fim de descongelar a amostra de carne sob temperatura controlada entre 4 °C e 7 °C. Após o descongelamento, foram avaliadas as perdas de líquido, tomando-se o peso inicial (massa bruta) e peso dos cortes secos (enxutos com papel toalha) após o descongelamento e pesagem da embalagem. O resultado obtido nesta análise foi a Perda por Descongelamento. A Fig. 10 apresenta duas fotografias de partes de amostras apreendidas ao serem cortadas, após coleta das perdas ocorridas posteriormente ao descongelamento controlado. As setas indicam pontos de exsudação anormal de líquidos de característica viscosa e com presença de espuma.

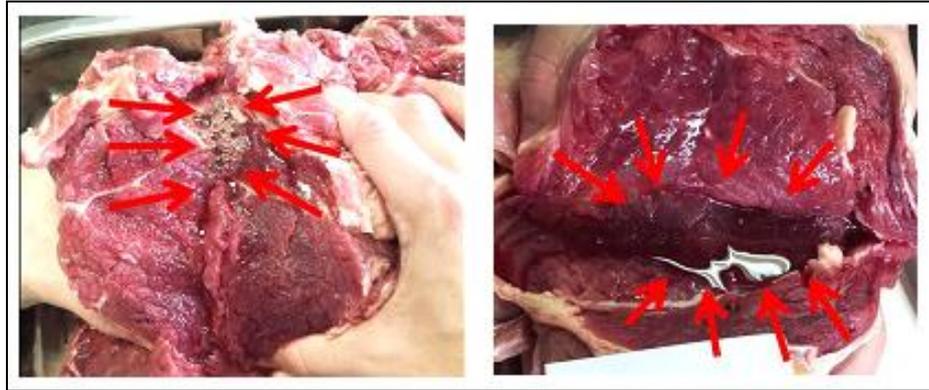


Figura 10: Fotografia de amostras apreendidas cortadas após coleta das perdas por descongelamento. As setas indicam pontos de exsudação anormal de líquidos de característica viscosa e presença de espuma.

As subamostras de carne congeladas foram cortadas com serra de fita para ossos (C.A.F. Máquinas, Rio Claro, SP) a fim de obter menores pedaços para a trituração das amostras. A Fig. 11 apresenta uma imagem mostrando o corte das amostras.

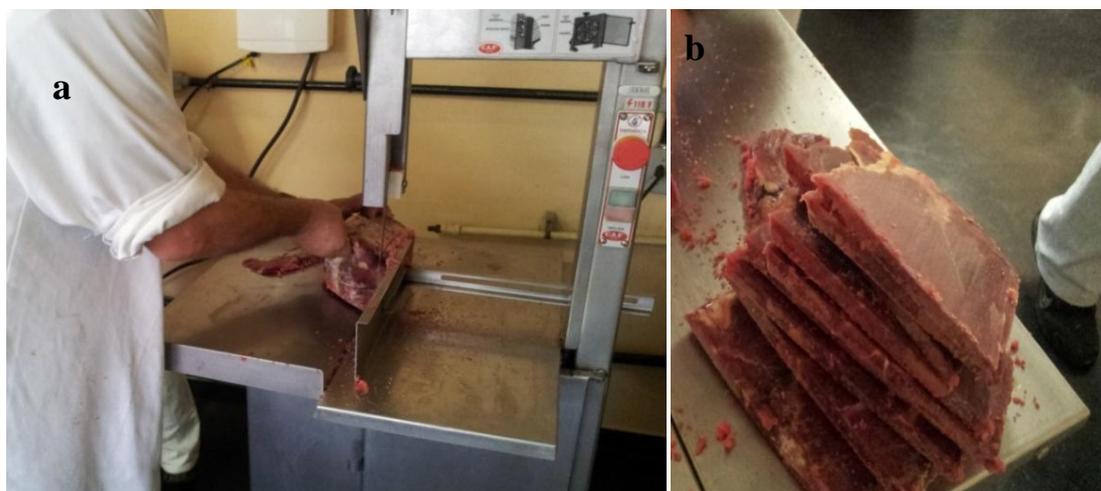


Figura 11: Corte das peças de carne com serra de fita. (a) procedimento de corte. (b) destaque para a amostra cortada em pedaços menores.

Em seguida, as peças congeladas cortadas foram trituradas com triturador de carne (Marconi, Piracicaba, SP). A Fig. 12 apresenta uma fotografia da trituração e homogeneização da peça de carne bovina congelada.



Figura 12: Processo de trituração e homogeneização da amostra de carne. (a) trituração da amostra. (b) carne triturada sendo homogeneizada com auxílio de espátula.

Foram realizadas, também, análises organolépticas de cor, odor e textura dos cortes em comparação com as amostras controle para verificação de alterações físicas na carne, identificação visual de materiais não-cárneos nas peças, marcas de perfuração e marcas de injeção de soluções ou temperos. A Fig. 13 mostra fotografias de amostras que apresentam evidências de perfurações puntiformes simétricas em suas superfícies. A presença de apenas um ponto de perfuração não necessariamente está associada à fraude, podendo ser característica dos ganchos utilizados para suspensão da carcaça. Mas, a presença de perfurações em malhas simétricas sugere a ocorrência de adulteração das carnes por injeção de substâncias exógenas.

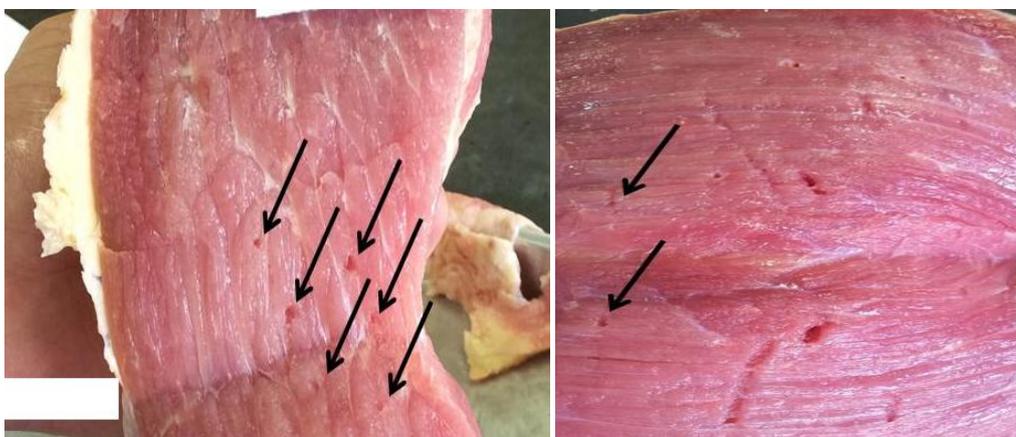


Figura 13: Evidências de perfurações puntiformes na superfície de uma amostra apreendida.

Análise centesimal também foi empregada para a determinação de proteínas, umidade e cinzas nas amostras, através de determinações físico-químicas. O teor de gordura não foi dosado, pois há variação natural entre os diferentes tipos de cortes de carne, além de variação intramuscular em uma mesma peça inviabilizando, assim, a comparação entre os valores determinados e o uso deste parâmetro para a detecção de fraudes. Os resultados obtidos foram comparados aos valores de referência da Tabela Brasileira de Composição de Alimentos (TACO)³¹ Por fim, foram realizadas análises para a quantificação de íons sódio, cloreto e fosfato usando a técnica cromatografia de íons (CI/TI) com detector de condutividade elétrica, cuja finalidade foi a determinação de quantidade extra de sais acima dos limites normais, considerando as condições de pré e pós-abate. A descrição dessas análises é apresentada nas próximas seções.

Todas as amostras analisadas foram peças de carne bovina *in natura*, as quais não possuem determinação legal que permita a injeção de líquidos ou substâncias químicas independente de quaisquer objetivos. A Fig. 14 apresenta fotografias de uma das amostras congeladas, evidenciando o acúmulo de líquido congelado nas extremidades das peças. Tais partes são facilmente quebradiças, característica anômala à carne considerada controle. Embora a inspeção visual indique claramente uma quantidade excessiva de líquidos exsudados pela peça de carne, esta análise não é suficiente para comprovação da sua adulteração. Para a comprovação da ocorrência de fraude, é necessária a aplicação de métodos analíticos para detecção de substâncias exógenas adicionadas à carne com o objetivo de proporcionar maior confiabilidade ao resultado encontrado, corroborando para a convicção pericial quanto à adulteração.



Figura 14: Evidência do acúmulo de líquido congelado nas extremidades das peças de amostras congeladas.

Os resultados dos exames de perdas por descongelamento, características organolépticas, teor de proteína, umidade, relação umidade/proteína, cinzas, teores dos íons sódio, cloreto e fosfato nas amostras de carne congelada e na purga obtida após descongelamento controlado convergiram no sentido de configurar a adição de soluções salinas nas amostras analisadas. Grandes quantidades de purga foram obtidas para as amostras suspeitas, incompatíveis até mesmo com princípios fisiológicos relacionados à osmolaridade (promoção do equilíbrio entre os meios extracelular e intracelular pela passagem de líquido por uma membrana permeável) de tecidos e fluidos orgânicos. A Fig. 15 apresenta a quantidade de líquido exsudado por amostras questionadas (suspeitas) e por uma amostra controle. Percebe-se a grande quantidade de líquido nas amostras questionadas comparativamente à amostra controle. Embora a quantidade de purga das amostras suspeitas seja na maioria das vezes muito maior que a quantidade de purga das amostras controle, este parâmetro não permite correlacionar a quantidade de líquido exsudado ao tipo de fraude, pois não identifica qual o tipo de substância exógena adicionada à carne. Este parâmetro, portanto, não pode ser utilizado para a análise de possível fraude nas peças *in natura*, pois a carne injetada com altos valores em percentagem em massa de solução adulterante, em alguns casos, não é capaz de suportar a grande quantidade de líquido injetado, perdendo ao longo do processo frações desta solução.

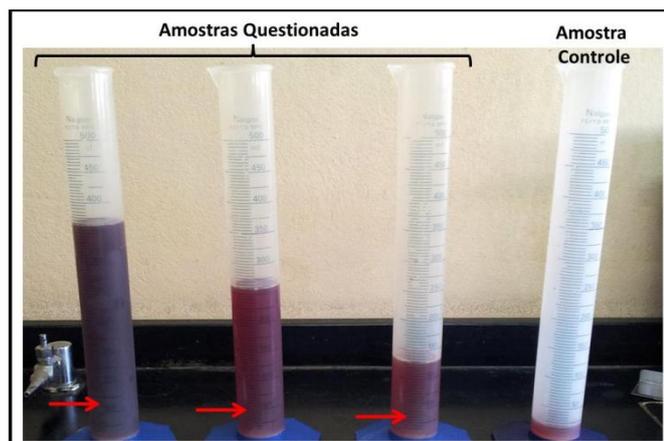


Figura 15: Volumes de líquido exsudado após procedimento de descongelamento (purga) da amostra controle e das amostras questionadas. As setas indicam o volume esperado proporcional ao peso da amostra controle (corte de 1.692g).

Nos frigoríficos sob investigação foram encontrados embalagens de produtos industriais utilizados na injeção de carne. Nestas misturas foram encontrados compostos, como colágeno hidrolisado, tripolifosfato de sódio, sal refinado (NaCl) iodado, pirofosfato ácido de sódio, carragena, maltodextrina proteína de soja, e derivados lácteos como soro de leite, cuja função é aumentar a retenção do líquido pela carne, reduzindo seu gotejamento até que a peça seja consumida, e, conseqüentemente, aumentando o peso final.

À carne *in natura*, por não sofrer nenhum processamento como a trituração, não é permitida a adição de nenhuma substância exógena (conservantes, corantes, acidulantes, flavorizantes, diluidores, temperos, etc.). Desta forma, não é exigido pelo Ministério da Agricultura a especificação da composição nutricional da carne *in natura* no rótulo do produto,²⁴ pois esta equivale à composição natural da carne bovina. Entretanto, as peças de carne submetidas à fraude, analisadas comparativamente antes e depois de uma injeção de líquido, apresentam alterações nos teores de seus componentes (proteínas, aminoácidos, sais, macromelementos, etc.), resultando, necessariamente, em redução do valor nutritivo. Outro problema associado à fraude com injeção de substâncias é o risco potencial à segurança alimentar, pois não existe qualquer controle de qualidade em relação às substâncias injetadas em relação à origem, às condições mínimas de segurança quanto à presença de componentes tóxicos, à potabilidade da água, ao teor de sais e aditivos desconhecidos pelo consumidor, entre outros.

Fraudes por injeção de substâncias exógenas são realizadas em grandes indústrias de processamento de carne, envolvem aquisição de maquinário, que é integrado à linha de produção dos cortes, e a manutenção de estoques de produtos usados na adulteração. Máquinas injetoras e máquinas para amaciamento de carne foram encontradas nos frigoríficos autuados. A Fig. 16 mostra duas máquinas, uma injetora de soluções e outra de amaciamento de carne, encontradas em um dos frigoríficos investigados.

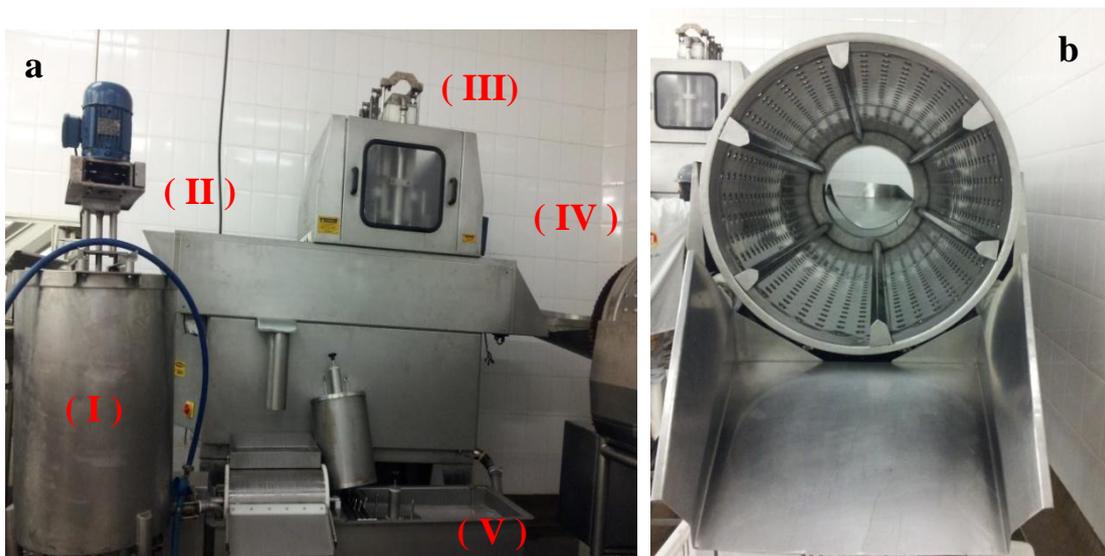


Figura 16: (a) Máquinas injetora empregada em adulteração de carne bovina *in natura* e (b) amaciadora de carne.

A máquina injetora possui um tanque de homogeneização (I) da solução preparada para injeção. Este tanque possui um motor elétrico responsável pela movimentação das pás de homogeneização. As peças de carne são inseridas pela abertura (II) manualmente. O sistema de injeção (III) tem as agulhas posicionadas de modo que a aparência na superfície da carne mostrará pontos de injeção nos vértices de um quadrado. Após a injeção as amostras são encaminhadas para uma rampa de saída (IV) acoplada à máquina amaciadora. O tanque (V) é utilizado para coletar e recircular a solução na máquina injetora.

O amaciamento da carne é realizado quando a máquina amaciadora gira em sentido anti-horário e uma de suas pás transporta a amostra de carne até um ponto em que a mesma caia por gravidade dentro da máquina. Este processo é repetido várias vezes até que a carne percorra toda a extensão da máquina. Este procedimento é utilizado no momento de fraude para fechar as fibras da carne e impedir a percepção dos pontos de injeção.

3.1.2 Reagentes e amostras de carne bovina

Todos os reagentes foram de grau analítico e utilizados sem prévia purificação. Água deionizada (Milli-Q) foi empregada em todos os procedimentos. Foram coletadas amostras em 5 frigoríficos distintos suspeitos de realizar adulterações. As 43 amostras suspeitas coletadas correspondem a peças embaladas a vácuo de coxão mole, lagarto, patinho, chã de fora, músculo e contra-filé. Outras 12 amostras foram coletadas diretamente de carcaças animais, no momento da desossa, sendo consideradas amostras controle comprovadamente não adulteradas por mecanismos suspeitos.

3.1.3 Determinação de parâmetros físico-químicos²⁴

Para avaliação da possível ocorrência de fraudes, cinco parâmetros físico-químicos foram determinados para todas as amostras: teores de proteína total, cinzas, sódio, cloreto e fosfato.

3.1.3.1 Determinação de proteína total

A determinação do teor de proteína total foi realizada pelo Método de Kjeldahl. Este método baseia-se na transformação do nitrogênio da amostra em sulfato de amônio através da digestão com ácido sulfúrico. Após a conversão do nitrogênio presente na amostra, ocorre a destilação liberando amônia, que é fixada em solução ácida e, então, titulada com HCl 0,1 mol/L. A multiplicação de fatores constantes específicos pela porcentagem de nitrogênio encontrado na análise representa a quantidade total de proteína presente na amostra.

A mistura catalítica foi preparada com K_2SO_4 e $CuSO_4 \cdot 5H_2O$ na proporção de 10:1, triturados em gral de porcelana até obtenção de um pó fino. Em seguida, foram adicionadas soluções de NaOH 50 % (p/v) e de H_3BO_3 4 % (p/v). A mistura foi então agitada e 4 g de H_3BO_3 foram adicionados. A mistura foi transferida para um béquer, ao qual foram adicionados 80 mL de água deionizada, antes do aquecimento sob agitação branda até dissolução. Em seguida, a solução foi resfriada à temperatura ambiente e transferida para um balão volumétrico de 100 mL, cujo volume foi completado com água. Para a preparação do indicador misto, foram pesados 0,132 g de vermelho de metila e 0,060 g de verde de bromocresol. Ambos foram dissolvidos em 200 mL de álcool etílico a 70% (v/v) e a solução obtida foi armazenada em frasco âmbar.

Para cada amostra, 0,50 g foram pesadas e transferidas para o tubo de Kjeldahl. Na sequência, foram adicionados 2,5 mL da mistura catalítica e 18 mL de H₂SO₄. A mistura foi aquecida em bloco digestor (FOSS, modelo Tecator Digestor Auto, Dinamarca) a 50°C por uma hora. Em seguida, a temperatura foi gradativamente elevada até atingir 400 °C. Quando o líquido tornou-se límpido e transparente, de tonalidade azul-esverdeada, a solução foi retirada do aquecimento e resfriada à temperatura ambiente. Em seguida, foram adicionados 10mL de água deionizada. O enlenmeyer contendo 20 mL de HBO₃ 4% e 4 gotas do indicador misto foi acoplado ao destilador. Em seguida, o tubo de Kjeldahl foi adaptado ao destilador e foi adicionado NaOH 50% até que a solução se tornasse negra (cerca de 20mL). A destilação foi realizada monitorando-se o pH até que não ocorresse mais reação alcalina. A solução receptora foi mantida fria durante a destilação. A solução obtida foi titulada com solução de HCl 0,1 mol/L.

O cálculo do teor de proteína total foi realizado a partir da Equação (15), em que *V* é o volume em mL de solução de HCl gasto na titulação após a correção do branco, *M* é a concentração exata em mol/L da solução de HCl, *f* é o fator de correção da solução de HCl 0,1mol/L e *p* é a massa da amostra em gramas:

$$\% \text{ Proteína} = \left(\frac{V * M * fc * 0,014 * 100}{p} \right) * 6,25 \quad (15)$$

3.1.3.2 Determinação do teor de cinzas

O método de determinação de cinzas fundamenta-se na eliminação da matéria orgânica e inorgânica volátil à temperatura de 550 °C. O produto obtido é denominado de resíduo mineral fixo ou cinzas.²⁴

Um cadinho de porcelana foi deixado em uma mufla (Sybron Termolyne) a 550 °C durante 30 min, resfriado em dessecador até temperatura ambiente e pesado em balança analítica. A amostra foi seca em estufa a 80°C por 2 h. Em seguida, foram pesadas nos cadinhos massas entre 2 a 5 g de amostra seca. Após esta etapa, o conjunto foi aquecido até a carbonização completa da amostra. Logo após, o cadinho foi levado à mufla, não ultrapassando 550°C, para evitar perda de cloretos. O conjunto foi incinerado até obtenção de cinzas

claras, resfriado no dessecador até atingir temperatura ambiente, e então pesado em balança analítica.

O cálculo do teor de cinzas foi realizado a partir da Equação (16), em que m é a diferença em gramas entre a massa do cadinho com amostra antes e após a calcinação e p é massa da amostra em gramas:

$$\% \text{ Cinzas} = \frac{m * 100}{p} \quad (16)$$

3.1.3.3 Determinação de Na^+ , Cl^- e PO_4^{3-} por cromatografia de íons

O resíduo obtido na determinação do teor de cinzas foi dissolvido em 5 mL de água deionizada e aquecido em banho-maria por 15 min. Em seguida, a solução obtida foi transferida para um béquer de 250 mL, diluída em 60 mL de água deionizada e mantida sob aquecimento por 30 min. Após o resfriamento da solução em temperatura ambiente, ela foi transferida para um balão volumétrico de 100 mL e o volume completado com água. Preparam-se curvas de calibração utilizando padrões certificados e verificando suas linearidades de maneira adequada. Os íons Na^+ , Cl^- e PO_4^{3-} presentes na solução resultante foram, então, quantificados utilizando-se a técnica de cromatografia de troca iônica (CI/TI). Para tal, foi utilizado um cromatógrafo Metrohm 761, Compact IC (Metrohm AG, Herisau, Suíça), com coluna para análise de ânions, com supressão química, METROSEP A SUPP 5-100, e com coluna para análise de cátions METROSEP CATION 1-2, e detector de condutividade elétrica.

3.2 Amostras simuladas (Plano de estudo)

3.2.1 Preparo das amostras

As amostras do denominado Plano de Estudo foram preparadas em frigorífico industrial seguindo a dinâmica de adulteração desenvolvida pelos frigoríficos investigados na Operação Vaca Atolada. Todas as etapas foram acompanhadas por um Veterinário a fim de garantir a integridade e autenticidade das amostras.

Para minimizar as fontes de variabilidade foram selecionados animais criados em um mesmo pasto, submetidos ao mesmo padrão de alimentação com ração e pastagem. Foram escolhidos animais machos, com peso variando

entre 243 e 351 kg e idade máxima de dois anos (idade verificada pela técnica de fórmula dentária).

Os animais foram abatidos em frigorífico industrial na cidade de Caeté. As peças de lagarto foram retiradas e armazenadas lacradas em câmara fria do mesmo frigorífico. As amostras escolhidas como peças controle não foram submetidas à injeção de soluções de forma que as propriedades físico-químicas das peças *in natura* fossem preservadas.

Antes de iniciar a injeção das misturas adulterantes nas amostras do Plano de Estudo (conforme Anexo B), foram realizados testes em algumas peças de carne para ajuste da máquina injetora quanto à quantidade de líquido a ser injetado e, também, para verificar o efeito causado pela injeção no aumento de volume da carne. A Fig. 17 apresenta fotografias de injeção em uma amostra teste antes da simulação com as amostras do plano de estudo. Percebeu-se um aumento significativo no volume da amostra teste quando o ganho de peso variou em torno de 3%, sendo possível verificar diferença de volume entre as peças injetadas e as peças iniciais: após a injeção, as fibras da carne ficaram mais inchadas e com uma textura mais macia, justificando a simulação das amostras por injeção. Este valor de concentração de solução injetada é inferior à concentração usada no plano de estudo, conforme descrito a seguir.



Figura 17: Amostra teste demonstrando o ganho de peso e volume com a injeção e aumento de 3 % m/m da peça.

A simulação de adulteração das carnes foi realizada em dois níveis: 10% m/m e 30% m/m de misturas preparadas. Entretanto, a máquina injetora possui sistema de ajuste de pressão manual, não sendo possível garantir uma injeção homogênea nas peças de carne, inclusive pelos diferentes pesos de cada uma. Foram utilizadas quatro misturas, sendo: (A) água potável; (B) mistura comercial de tripolifosfato de sódio; (C) mistura contendo tripolifosfato de sódio, maltodextrina, sal refinado iodado, pirofosfato ácido de sódio e carragena; (D) mistura contendo tripolifosfato de sódio, maltodextrina, sal refinado iodado, pirofosfato ácido de sódio, carragena e colágeno. Após a simulação de adulteração, as amostras foram embaladas à vácuo e mantidas sob refrigeração industrial. O processo de embalagem à vácuo auxilia no processo de fechamento das fibras para disfarçar as marcas de injeção. Na Fig. 18 percebe-se as marcas deixadas pela injeção, antes e após a embalagem à vácuo.

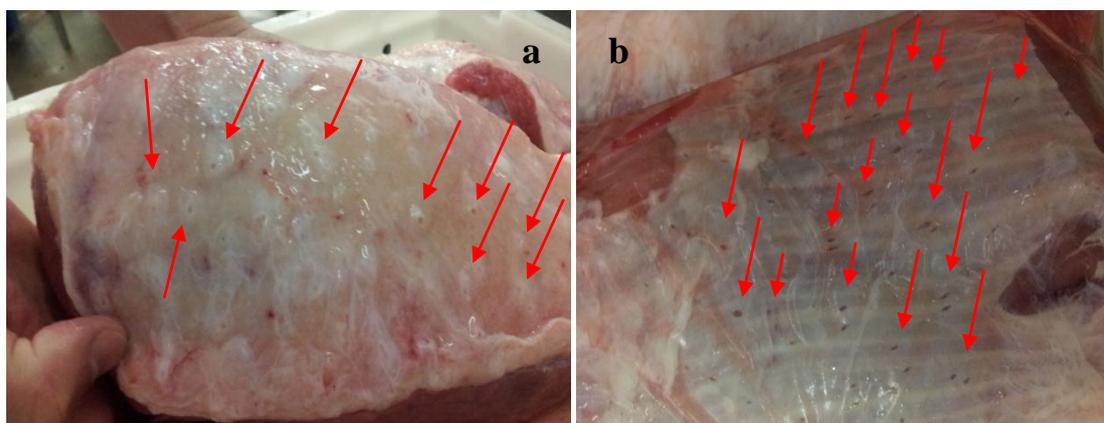


Figura 18: Marcas de injeção (setas vermelhas). (a) antes e (b) após embalagem à vácuo.

Em seguida, 51 amostras foram submetidas a descongelamento controlado em estufa incubadora de determinação de DBO por um período de 48 h, em temperatura de 4 °C a 7 °C a fim de descongelar a peça de carne sob temperatura controlada. Após o descongelamento, os líquidos exsudados (purga) foram coletados para serem utilizados nas análises espectroscópicas (Fig. 15).

Dezoito amostras de carne congeladas foram cortadas com serra de fita para ossos a fim de obter menores pedaços para a trituração (Fig. 11).

Posteriormente, as peças cortadas foram colocadas no triturador de carne (Fig. 12). As subamostras de carne e de purga foram armazenadas em freezer a temperatura de - 22°C. Todas as amostras de purga foram analisadas à temperatura ambiente.

3.3 Dados espectroscópicos

3.3.1 Equipamento

Foi utilizado um espectrofotômetro de infravermelho médio FTIR Nicolet 380 (Thermo Fisher Scientific Inc., Madison, EUA), usando um acessório de reflectância total atenuada (ATR) equipado com cristal de diamante de uma reflexão (Smart Orbit Diamond ATR, Thermo Fisher, EUA), presente no Laboratório de Setor Técnico-Científico da Superintendência Regional do Departamento de Polícia Federal, em Belo Horizonte.

3.3.2 Aquisição dos espectros

Pequenas porções das amostras de carne (quantidade suficiente para cobrir a abertura do cristal) foram colocadas no acessório de ATR do espectrômetro com auxílio de espátula. Na análise das purgas pequenos volumes foram adicionados, usando pipetas de Pasteur, até a cobertura total da abertura do cristal, conforme mostrado na Fig. 19. Os espectros foram obtidos em triplicata de 4000 a 525 cm^{-1} com resolução de 2 cm^{-1} e 32 varreduras. Para a construção dos modelos foi utilizada a média dos valores dos espectros obtidos para cada amostra.

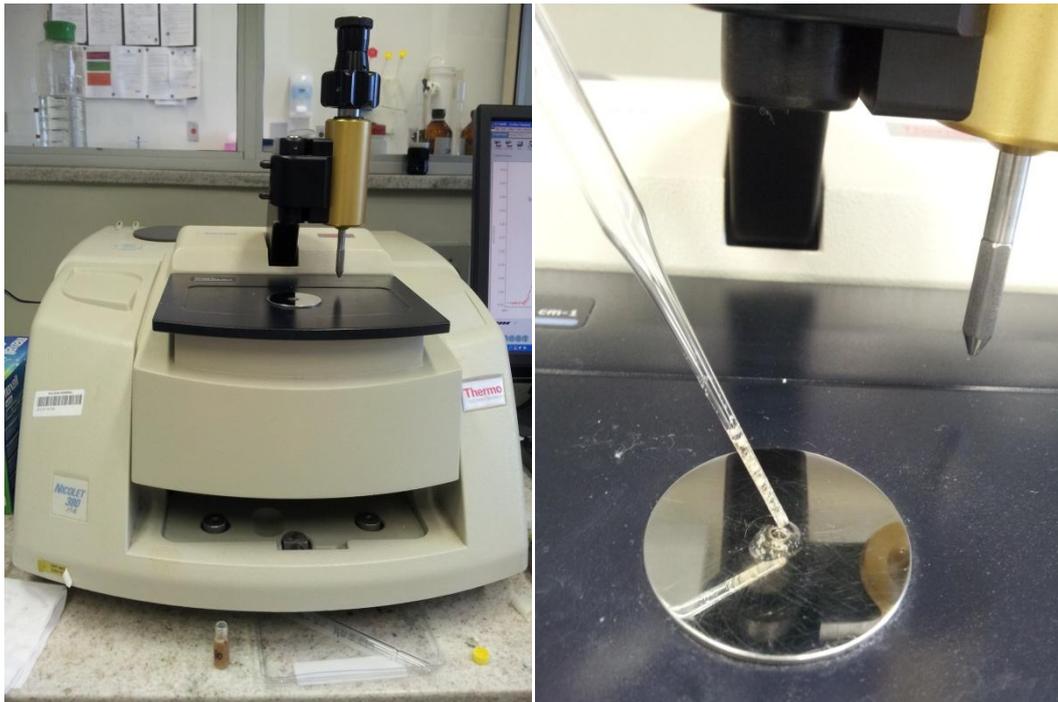


Figura 19: (a) Espectrofotômetro de infravermelho médio FTIR Nicolet 380 com acessório de reflectância total atenuada (ATR). (b) análise de amostra de purga.

3.4 Processamento dos dados

Os dados obtidos foram processados com o programa MATLAB versão 7.13 (The MathWorks, Natick, EUA), utilizando o pacote PLS Toolbox versão 6.5 (Eigenvector Technologies, Manson, EUA).

4 Resultados e discussão

Os resultados obtidos nesta dissertação serão apresentados em duas seções: amostras reais e amostras do plano de estudo. A seção apresentando os resultados das amostras reais, amostras de carne apreendidas na operação “Vaca Atolada”, está dividida em três subseções as quais apresentam os modelos PCA, PLS-DA individuais e PLS-DA com fusão de dados. Nas duas primeiras subseções são apresentados os modelos construídos a partir dos dados físico-químicos e de MIR para as amostras reais. Os dados físico-químicos utilizados são mostrados no Anexo A.

A segunda seção, que mostra os resultados para a análise de purgas das amostras propositadamente adulteradas, apresenta os modelos PLS-DA construídos com espectros MIR. A planilha do planejamento para o preparo das amostras simuladas é mostrada no Anexo B.

4.1 Amostras Reais

4.1.1 Modelos PCA

4.1.1.1 Modelo PCA para os dados Físico-Químicos

Os resultados encontrados para os parâmetros Proteína Total, Cinzas, íons Sódio, Cloreto e Fosfato foram agrupados em uma matriz de dados para análise por PCA. Por apresentarem variáveis de diferentes naturezas e distribuições, os dados foram autoescalados a fim de conferir atribuição de pesos iguais aos resultados analíticos, independente da escala natural de cada variável.

No modelo construído, as duas primeiras CPs explicaram 90,81% da variância total, enquanto a 3ª CP explicou apenas 5,18%. Portanto, considerou-se que a observação das duas primeiras CPs é suficiente para a interpretação dos dados. Conforme observado na Fig. 20, que mostra os escores do modelo, sem atribuição de origem das amostras, é possível verificar que algumas delas encontram-se mais próximas de outras. Entretanto, não se percebe um agrupamento claro de amostras, sem a distinção classes.

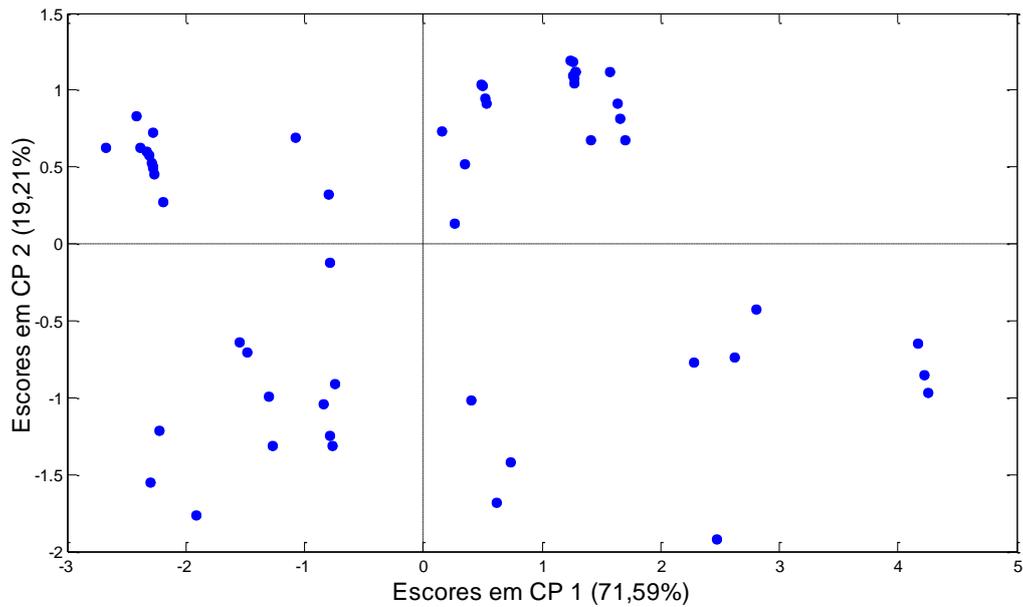


Figura 20: Escores de CP1 x CP2 para dados físico-químicos das amostras reais.

Entretanto, embora a PCA seja um método não-supervisionado, é possível realizar análises semi-supervisionadas atribuindo informações acerca das classes às quais as amostras pertencem. Na Fig. 21 é apresentado o gráfico de escores de CP1 *versus* CP2 para o mesmo conjunto de dados, com a identificação dos tipos de corte de carne bovina das amostras analisadas. Percebe-se neste gráfico que há uma tendência das amostras de Contra-filé e de Músculo se agruparem em valores positivos de CP1 e CP2 (elipses vermelha e verde, respectivamente). Já as amostras de Chã de fora são caracterizadas por valores negativos de CP1 (elipse rosa). As amostras de Patinho apresentam uma tendência negativa em CP1, embora duas delas tenham mostrado valores positivos em CP1, próximos ao valor zero. Já as amostras de Coxão mole caracterizam-se por valores positivos em CP1, com exceção de duas que apresentaram valor negativo nesta mesma componente.

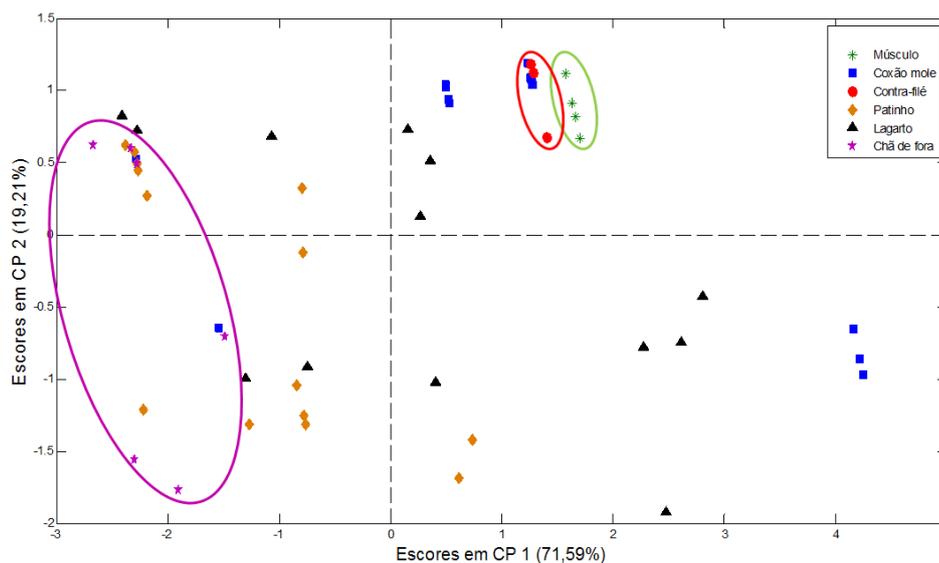


Figura 21: Escores de CP1 x CP2 para dados físico-químicos de amostras reais - Tipos de corte

De um modo geral, o modelo PCA apresentou uma razoável capacidade de agrupamento para as amostras em função do tipo de corte de carne. No entanto, para os objetivos deste projeto, o mais importante é a diferenciação das classes de amostras suspeitas e controle através do modelo. Na Fig. 22 é apresentado o gráfico de escores de CP1 versus CP2 para o mesmo modelo, com a identificação das amostras controle como triângulos vermelhos e das amostras questionadas como círculos pretos. Foi calculada uma elipse de confiança em torno dos escores das amostras controle, caracterizadas por valores negativos de CP1, com 90% de confiança. Percebe-se que 32 amostras questionadas diferiram significativamente do padrão de carnes controle não adulteradas. O restante das amostras questionadas, 11, não diferiu significativamente das amostras controle. Destas 11 amostras, 7 não apresentavam evidências visuais de adulteração, pois não foram observados pontos de injeção e as análises físico-químicas e centesimais não detectaram alterações em valores ou a presença de substâncias exógenas adicionadas às peças de carne. Entretanto, podem ter sido adicionadas pequenas concentrações de soluções adulterantes não permitindo, assim, a identificação e classificação da possível adulteração.

Com o objetivo de tentar construir melhores modelos classificatórios serão aplicados outros tipos de metodologias, como métodos quimiométricos supervisionados, técnicas espectroscópicas e estratégia de fusão de dados.

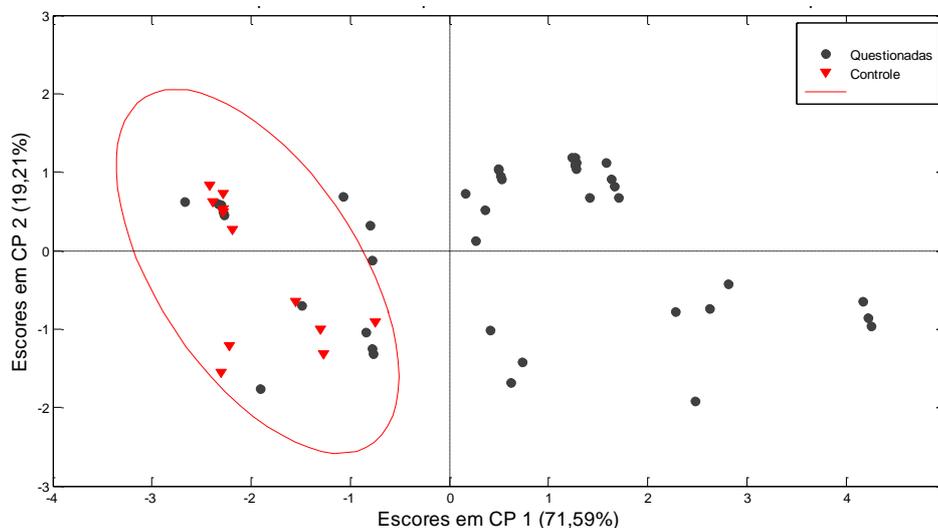


Figura 22: Escores de CP1 x CP2 para dados físico-químicos de amostras reais - Análise Semi-supervisionada

O modelo de PCA construído com os dados físico-químicos das amostras reais permitiu detectar que 74,4% (32 amostras em 43) das amostras suspeitas diferem significativamente das amostras não adulteradas.

Analisando os pesos para os dados físico-químicos das amostras reais (Fig. 23), verificam-se dois padrões: em CP1 o teor de proteínas está em oposição ao teor de sais de modo geral, enquanto na CP2 o teor de proteínas está em oposição ao teor de fosfato, especificamente. Tal resultado deve-se à adição de substâncias comerciais contendo sais fosfatados. Verificou-se, ainda que os teores de cinzas e de íons sódio estão muito correlacionados. Avaliando os gráficos de escores e de pesos simultaneamente, percebe-se que as amostras controle são caracterizadas por valores negativos de CP1, ou seja, elas possuem maior teor de proteínas e menores teores de íons cloreto, fosfato, sódio e cinzas do que a maioria das amostras adulteradas.

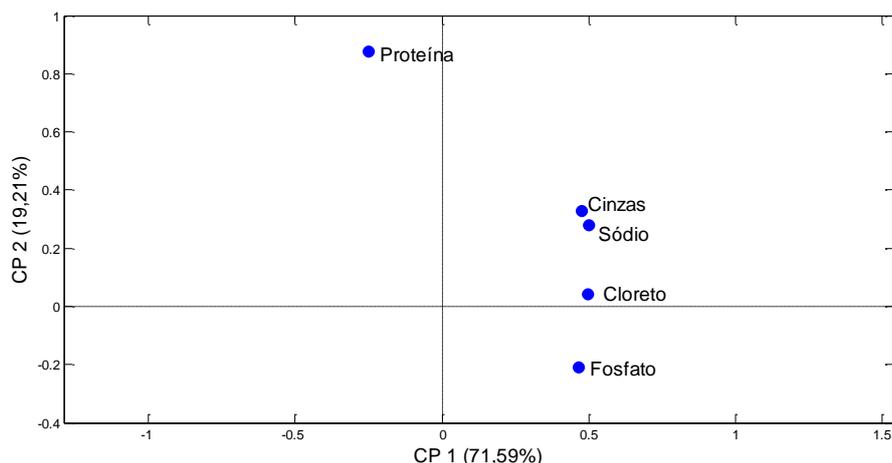


Figura 23: Pesos para os dados físico-químicos das amostras reais.

Comparando-se as figuras 22 e 23, percebe-se que em CP1 as amostras controle (valores negativos) estão diretamente correlacionadas a um maior teor de Proteína Total e um menor teor de sais, de um modo geral, enquanto as amostras suspeitas de adulteração possuem maiores concentrações de sais e menores teores de proteína. Como explicado anteriormente, este resultado era esperado, pois a adição de sais aumenta a capacidade de retenção de água pela peça de carne, reduzindo a razão água/proteína a valores próximos dos normais. Em relação à CP2, verifica-se que as amostras controle (valores positivos) apresentam maior teor de Proteína e menor teor de íons Fosfato especificamente.

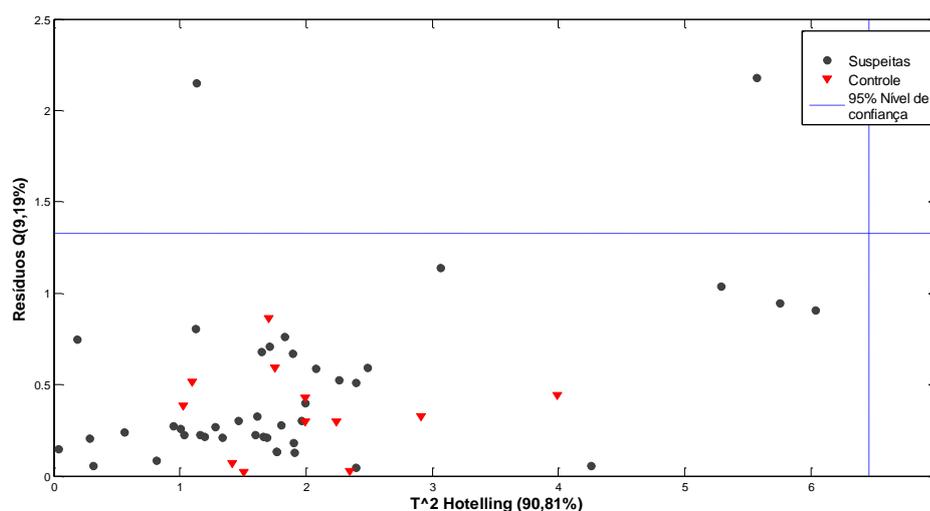


Figura 24: Gráfico de resíduos Q do modelo em função de T^2 de Hotelling (influência) para um modelo com 2 CPs. As linhas pontilhadas indicam limites de confiança de 95% para os parâmetros.

O modelo PCA para os dados físico-químicos das peças de carne analisadas na Operação Vaca Atolada não apresentou amostras com comportamento anômalo, ou seja, simultaneamente com altos resíduos e alta influência, conforme pode ser verificado na Fig. 24. Uma amostra com alto resíduo foi pouco modelada e não deve ser incluída no modelo, assim como uma amostra com alta influência afeta significativamente o modelo construído, pois a influência é a medida da variação de cada amostra dentro do modelo PCA. Logo, este foi considerado um bom modelo para uma análise preliminar dos dados analíticos.

4.1.1.2 Modelo PCA para os dados de ATR-FTIR

Os espectros MIR registrados para as 55 amostras são mostrados na Fig. 25. Visualmente, é difícil notar uma discriminação entre as amostras suspeitas e controle. As absorvâncias dos espectros na região do infravermelho médio foram agrupadas em uma matriz de dados para análise por PCA. Por apresentarem variáveis de mesma natureza e distribuição, os dados foram centrados na média. Antes disso, eles foram pré-processados usando os mesmos métodos que serão descritos na seção 4.1.2.2 (PLS-DA).

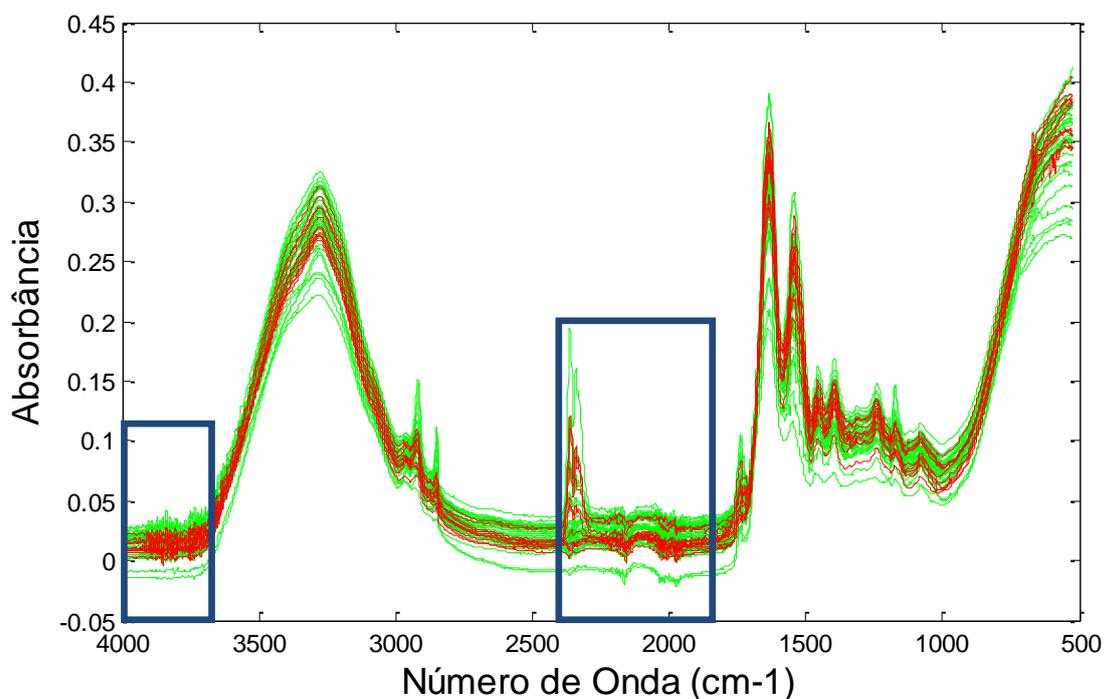


Figura 25: Espectros MIR para as 55 amostras. Em vermelho as amostras controle e em verde as amostras suspeitas. Em azul, as regiões espectrais eliminadas são destacadas.

Para a construção do modelo foram eliminadas as regiões de número de onda entre 1800 e 2400 cm^{-1} (região de absorção de CO_2 e do diamante do ATR⁷⁰) e acima de 3700 cm^{-1} (região contendo somente ruído espectral). No modelo construído, as três primeiras CPs explicaram 94,17% da variância total, enquanto a 4ª CP explicou apenas 3,07%. Portanto, considerou-se que a observação das três primeiras CPs é suficiente para a interpretação dos dados.

Na Fig. 26 é apresentado o gráfico de escores de CP1 *versus* CP2, com a identificação dos tipos de corte de carne bovina das amostras analisadas. Este gráfico foi construído a fim de verificar se a variância explicada pelo modelo poderia se referir ao tipo de corte de carne e não à adulteração. Embora se percebam algumas tendências, como a presença de amostras de chã de fora em valores positivos de CP1 e de amostras de contra filé em valores negativos de CP1, claramente não se observa agrupamentos de amostras por tipo de carne, indicando que o modelo não distingue tipos de corte.

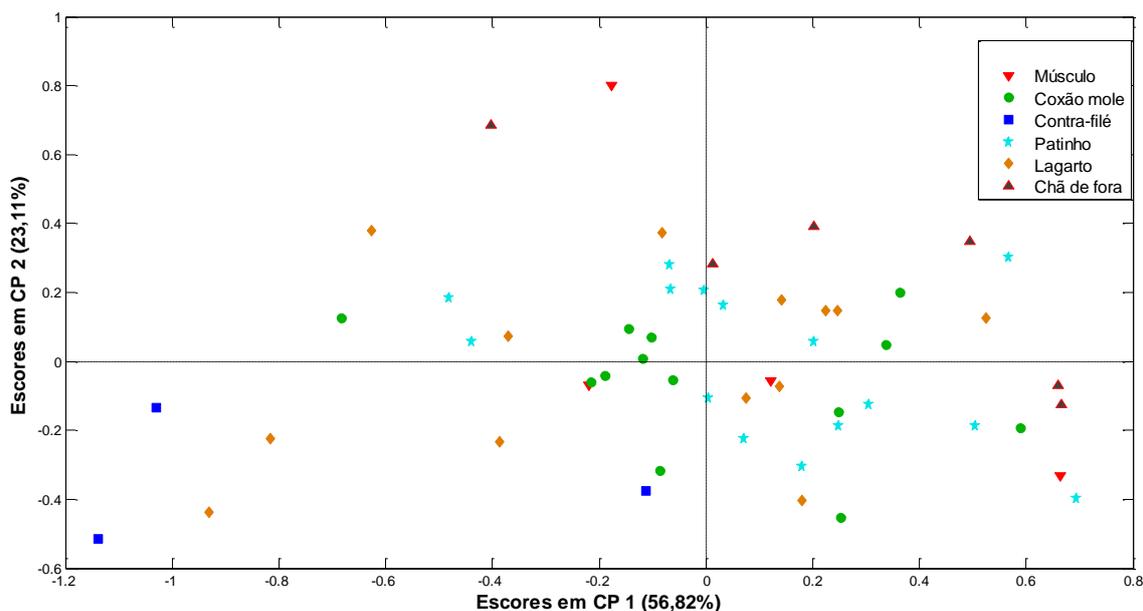


Figura 26: Escores de CP1 x CP2 para dados ATR-FTIR de amostras – Tipos de carne.

Na Fig. 27 é apresentado o gráfico de escores de CP1 *versus* CP2 para o mesmo conjunto de dados, com a identificação das amostras controle como triângulos e das amostras suspeitas como círculos. Embora as amostras controle se encontrem menos dispersas no gráfico, não se percebe um agrupamento claro delas. A terceira CP (não mostrada) não apresentou

nenhuma discriminação significativa. De um modo geral, o modelo PCA não apresentou capacidade de discriminação para estas amostras.

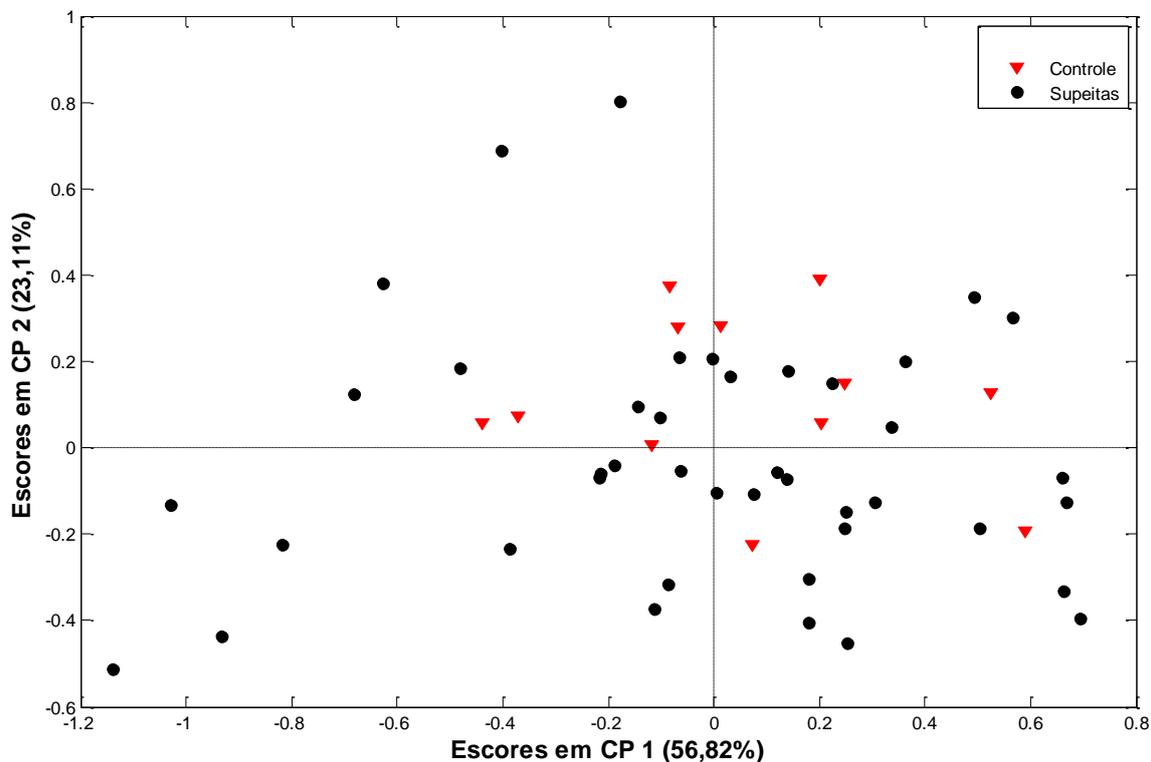


Figura 27: Escores de CP1 x CP2 para dados ATR-FTIR.

Comparando-se as Figs. 27 e 22, nota-se claramente que os resultados físico-químicos possuem maior capacidade discriminante em relação à adulteração das amostras do que os espectros MIR. Isto já era esperado, pois as variáveis físico-químicas foram escolhidas justamente buscando esta discriminação, enquanto os espectros vibracionais fornecem informações menos seletivas, que estão ligadas à composição química total das amostras, a qual está sujeita a uma variabilidade muito maior em função de várias características, tais como o tipo de corte da carne e a origem do animal. Por outro lado, a espectroscopia MIR é uma técnica mais simples, rápida e barata, que pode ser usada de maneira não destrutiva, não consome reagentes nem gera resíduos. Desta forma, existe o interesse em se combinar dados de MIR com um número menor de variáveis físico-químicas, gerando economia e rapidez nas análises. Além disso, serão desenvolvidos na sequência modelos de classificação supervisionada, para os quais o acréscimo da informação

sobre a origem das amostras, questionadas ou controle, pode representar um grande ganho na sua capacidade preditiva.

4.1.2 Modelos PLS-DA

As amostras foram divididas em dois conjuntos de dados, sendo 38 para treinamento e 17 para teste. O conjunto de treinamento deve apresentar amostras que são representativas de toda a variabilidade incluída no modelo, enquanto o conjunto de teste é usado para testar sua eficácia de maneira robusta, a partir de amostras que foram omitidas na sua construção. A seleção das amostras nos dois conjuntos foi realizada utilizando o algoritmo de seleção de Kennard-Stone (KS) separadamente em cada classe. Desta forma, as amostras suspeitas de fraude (43) foram separadas em 30 para treinamento e 13 para teste. As amostras controle (12) foram divididas em 8 para treinamento e 4 para teste. O número de amostras escolhidas em cada classe foi definido previamente como aproximadamente dois terços para o conjunto de treinamento. Em cada modelo, o algoritmo de KS foi aplicado em particular. Portanto, as amostras de treinamento e teste não são necessariamente as mesmas nos vários modelos apresentados a seguir. A definição de amostras adulteradas como classe 1 e amostras controle como classe 0 foi arbitrária.

4.1.2.1 Modelo PLS-DA para os dados Físico-Químicos

Após a separação das amostras nos conjuntos de treinamento e teste, os dados foram autoescalados e submetidos ao processo de reamostragem por validação cruzada. Para o modelo PLS-DA com os dados físico-químicos foram escolhidas 3 VLs, as quais explicam 94,16% da variância em X e 36,58% em Y . As previsões do modelo são apresentadas na Fig. 28.

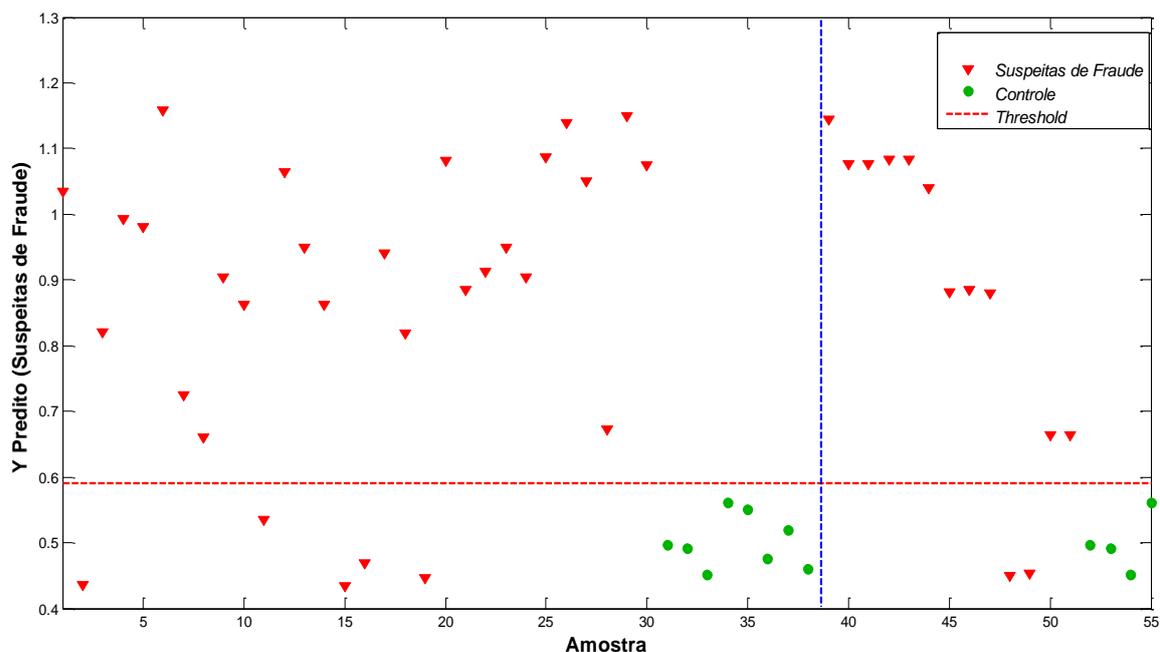


Figura 28: Previsões para o modelo PLS-DA com os dados físico-químicos. A linha tracejada vermelha indica o limite de decisão (*threshold*) e a linha tracejada azul indica a separação dos conjuntos treinamento e teste.

As primeiras 38 amostras pertencem ao conjunto treinamento e as demais ao conjunto teste. O *threshold* foi estimado próximo a 0,6, logo, acima deste valor as amostras são previstas como adulteradas e abaixo deste valor as amostras são previstas como não adulteradas. Percebe-se nesta figura, que no conjunto de treinamento 5 amostras foram previstas como não adulteradas em 30 amostras suspeitas de fraude (erros falso-negativos), enquanto nenhuma amostra controle foi prevista como adulterada. Já o conjunto de teste apresentou 2 amostras previstas como não adulteradas em 13 questionadas, enquanto nenhuma amostra controle foi prevista como adulterada. Algumas figuras de mérito para este modelo são apresentadas na Tabela 4. Tais resultados demonstram uma boa capacidade preditiva do modelo construído.

Tabela 4: Figuras de mérito do modelo PLS-DA para os dados físico-químicos.

Conjunto/ Parâmetro	Sensibilidade (%)	Especificidade (%)	TFN (%)	TFP (%)
Treinamento	83,3	100,0	16,7	0,0
Teste	84,6	100,0	15,4	0,0

Avaliando-se os Coeficientes de Regressão para as variáveis utilizadas neste modelo (Fig. 29), as variáveis acima de 0 contribuem mais significativamente para a previsão das amostras suspeitas de fraude, enquanto variáveis abaixo de 0 contribuem mais significativamente para a previsão de amostras controle. Percebe-se que o teor de íons Cloreto é claramente o que mais contribui para a previsão das amostras adulteradas. Este fato é extrema relevância, pois o cloreto de sódio (sal refinado iodado) foi um dos sais adicionados com fins de fraude, visando o aumento da WHC da carne. Não se verificou a mesma importância para o íon sódio, provavelmente por este ser um componente natural dos músculos, participando do transporte ativo de substâncias e íons através da membrana plasmática (bomba sódio-potássio), e estar presente naturalmente nas amostras em altas concentrações. Como os dados foram autoescalados, foi atribuído pesos às variáveis de naturezas distintas. Desta forma, o gráfico de coeficientes de regressão mostra as variáveis mais importantes para discriminação de classes, independente da escala natural das variáveis físico-químicas..

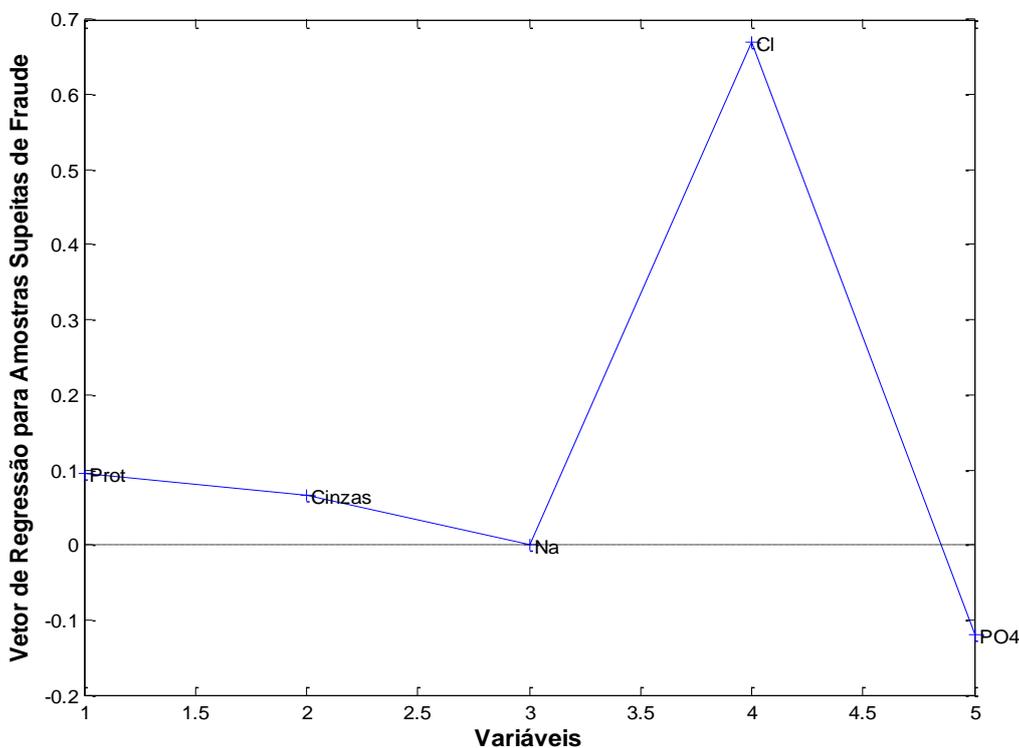


Figura 29: Coeficiente de Regressão para as variáveis físico-químicas.

A avaliação dos coeficientes de regressão nos permite verificar a direção em que cada variável é mais significativa para o modelo, entretanto, os

parâmetros mais robustos para a análise da importância das variáveis no modelo são os VIPscores (Fig. 30). Variáveis com valores de VIPscores maiores que 1 são significativas para as previsões do modelo, enquanto variáveis abaixo de 1 não contribuem significativamente. Neste gráfico, confirma-se que a variável mais importante para o modelo de previsão é o teor de Cl⁻, enquanto as variáveis teor de Proteína Total e de íons Fosfato não foram significativas. No entanto, no modelo PCA (não supervisionado) para os dados físico-químicos, o teor de Proteína foi associado às amostras controle.

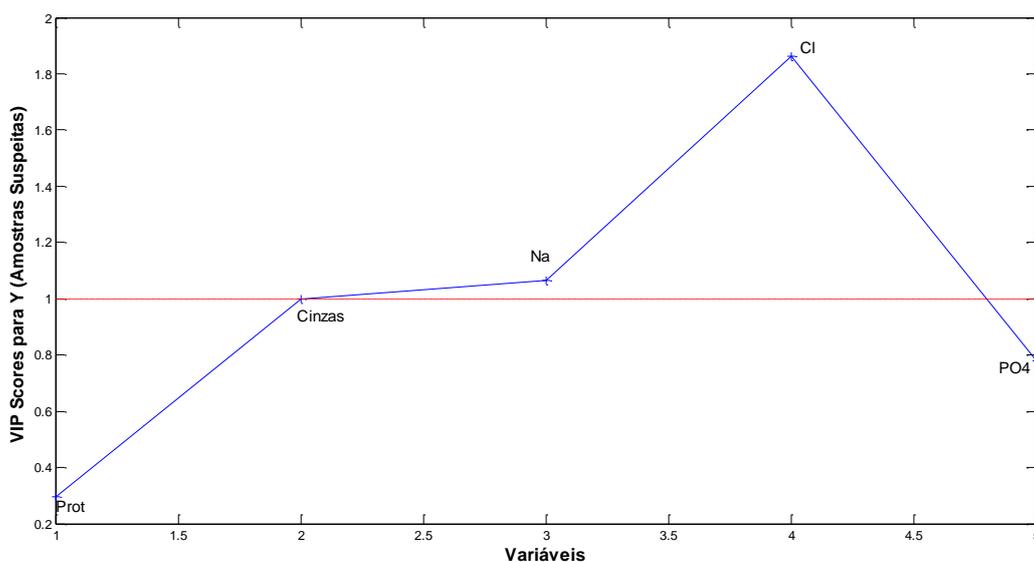


Figura 30: VIPscores para as variáveis físico-químicas. A linha pontilhada indica o valor de 1,0, sugerido como limite de significância.

4.1.2.2 Modelo PLS-DA para os dados ATR-FTIR

Assim como no modelo PCA, para a construção deste modelo também foram previamente eliminadas as regiões espectrais entre 1800 e 2400 cm^{-1} e acima de 3700 cm^{-1} . Após a separação das amostras nos conjuntos de treinamento e teste usando o algoritmo KS, os dados foram pré-processados na seguinte sequência. Primeiro, foi usado o alisamento Savitzky-Golay,⁷¹ com janela de 15 pontos e polinômio de 2º grau (sem derivada), visando a redução de ruídos espectrais. Este método utiliza filtros digitais polinomiais de janela móvel, assumindo que o ruído tem alta frequência em relação ao sinal de interesse, para aumentar a razão sinal/ruído. Em seguida, foi aplicada a correção do espalhamento multiplicativo (MSC - *Multiplicative Scatter Correction*),⁷² visando eliminar desvios de linha-base não lineares, tipicamente

observados em medidas de reflectância de sólidos no infravermelho. Finalmente, os dados centrados na média.

Para este modelo foram escolhidas 4 VLs, as quais explicam 95,70% da variância em **X** e 20,71% em **Y**. As previsões são apresentadas na Fig. 31 O *threshold* foi estimado próximo a 0,7.

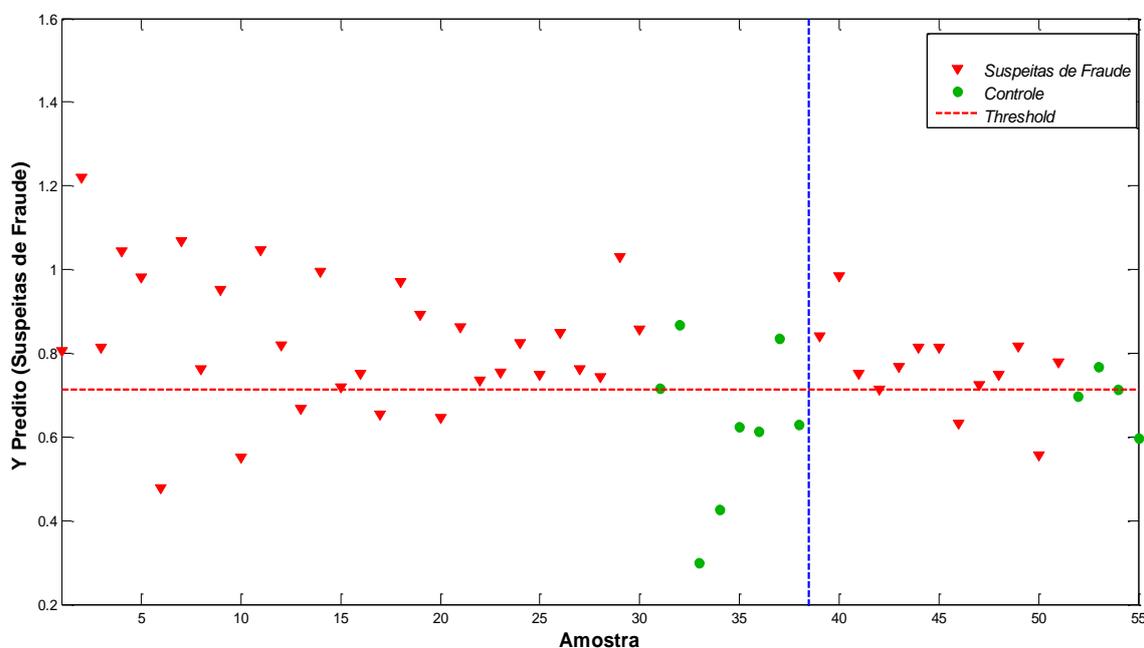


Figura 31: Previsões para o modelo PLS-DA construído com os espectros MIR.

Percebe-se nesta figura que no conjunto de treinamento 5 amostras foram previstas como não adulteradas em 30 possíveis. Para o mesmo conjunto, 3 amostras controle foram previstas como adulteradas (erros falso-positivos) em 8 possíveis. Já o conjunto de teste apresentou 3 amostras previstas como não adulteradas em 13 questionadas, enquanto 1 amostra controle foi prevista como adulterada em 4 possíveis. As figuras de mérito relativas a esta previsão são apresentadas na Tabela 5. Percebe-se, como esperado, que este modelo apresenta uma capacidade preditiva inferior ao modelo anterior, construído apenas com as variáveis físico-químicas. No entanto, os valores de sensibilidade e especificidade foram considerados razoáveis, considerando o total de amostras.

Tabela 5: Figuras de mérito do modelo PLS-DA para os dados ATR-FTIR das amostras reais

Conjunto/ Parâmetro	Sensibilidade (%)	Especificidade (%)	TFN (%)	TFP (%)
Treinamento	83,3	62,5	16,7	37,5
Teste	76,9	75,0	23,1	25,0

Avaliando-se os Coeficientes de Regressão para as variáveis utilizadas neste modelo (Fig. 32), verifica-se que a região entre 3000 e 2750 cm^{-1} apresenta picos de grande intensidade positiva, que estão associados à previsão de amostras adulteradas. Da mesma maneira, também são observados picos positivos em 670, 1100, 1170, 1250, 1460, 1555 e 1660 cm^{-1} , enquanto os picos em torno de 800, 900, 1000, 1125, 1220, 1360, 1430, 1500, 1600, 1630 e 1750 cm^{-1} apresentam intensidades negativas significantes, estando, portanto, associados à previsão de amostras controle.

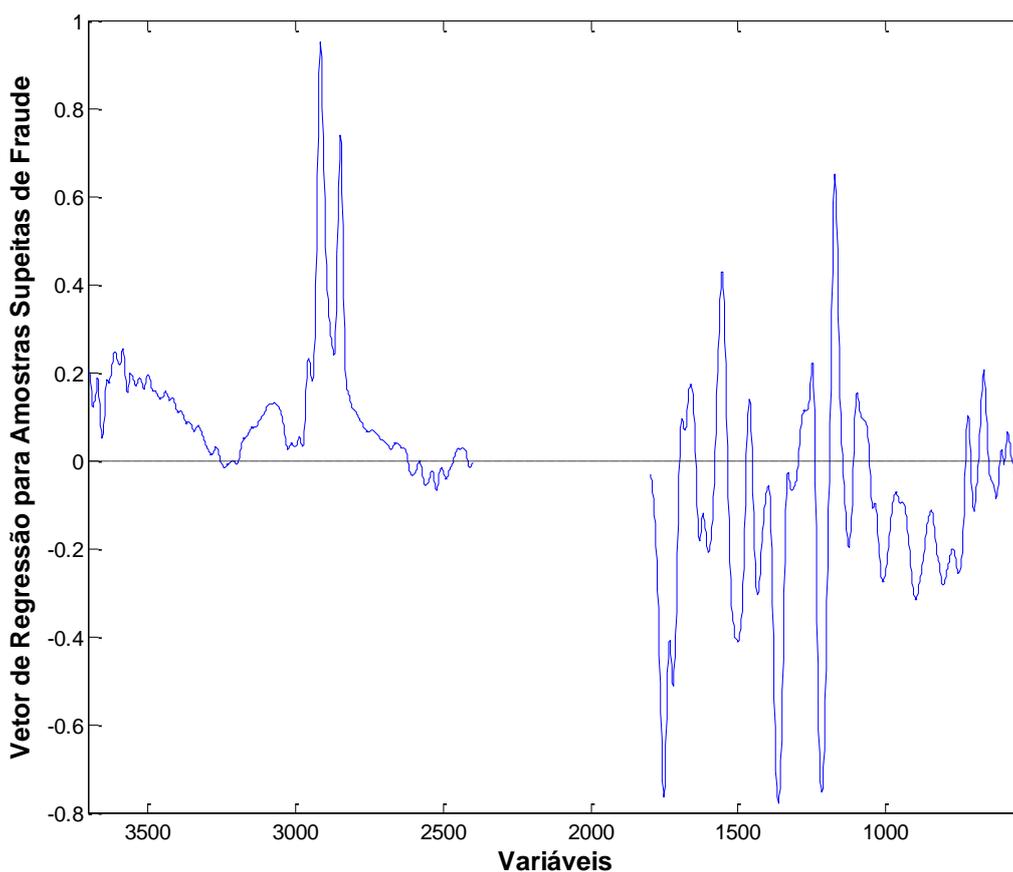


Figura 32: Coeficiente de Regressão para as variáveis de ATR-FTIR.

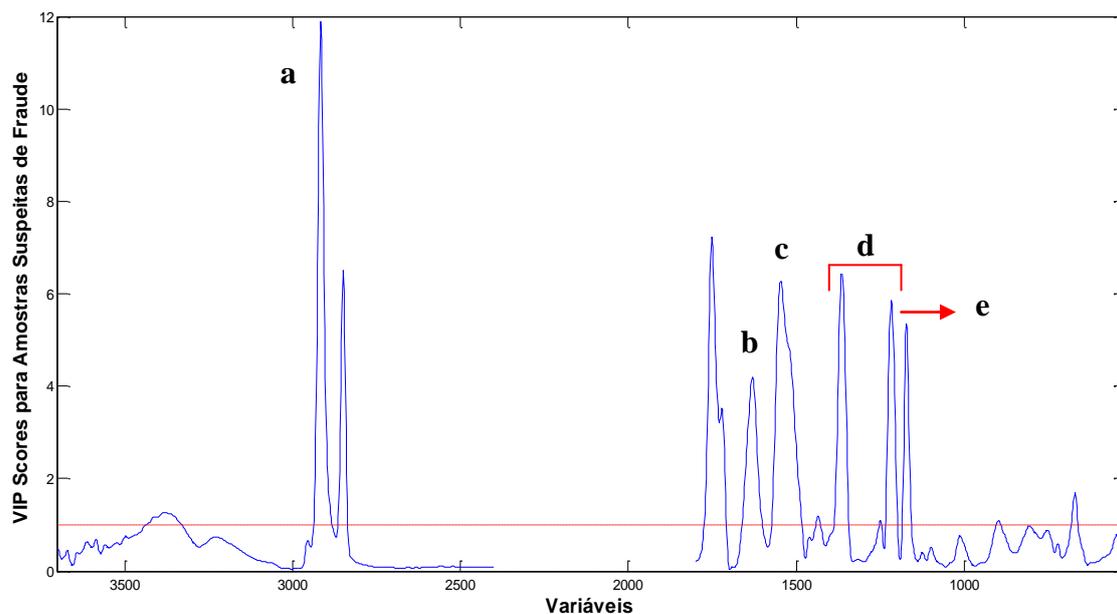


Figura 33: *VIP scores* para o modelo PLS-DA com os espectros ATR-FTIR.

De maneira complementar, analisando-se o gráfico de *VIP scores* (Fig. 33), conclui-se que o pico mais significativo para o modelo está situado em 2916 cm^{-1} (geralmente atribuído à deformação axial assimétrica CH_2 – (a)),³⁹ seguido, em importância, pelos seguintes picos: 1750, 2850, 1365, 1554, 1220, 1170, 1630 e 670 cm^{-1} .

Considerando-se os parâmetros do modelo discutidos anteriormente, buscaram-se atribuir as regiões espectrais mais importantes para a discriminação de amostras questionadas, considerando os principais adulterantes encontrados nos locais pela operação da Polícia Federal. Destacam-se, portanto, os picos em 1630 cm^{-1} (b), geralmente característico de resíduos de imina em colágeno tipo II, em 1554 cm^{-1} (c), comumente característico de amida II, e as bandas entre 1200 e 1350 cm^{-1} (d), usualmente atribuídas à região de *fingerprint* para colágeno devido ao tripeptídeo (Gly-Pro-Hyp).⁷³ Destacam-se, também, as bandas em aproximadamente 1690 e 1230 cm^{-1} , que podem ser associadas às vibrações de agregados de folhas-beta de proteínas, provocadas pela adição de NaCl ou KCl à carne²⁰ e uma banda de vibração em torno de 1220 cm^{-1} (e), em geral associada ao grupo éster sulfato da carragena.⁷⁴

4.1.3 Fusão de Dados

4.1.3.1 Modelo PLS-DA de Fusão de Dados Completa

Foi adotada a estratégia, mais simples e comum na literatura, de fusão de dados de baixo nível, ou seja, os dados serão simplesmente combinados e modelados após sofrerem os pré-processamentos adequados. Assim como na seção anterior, as regiões espectrais entre 1800 e 2400 cm^{-1} e acima de 3700 cm^{-1} foram previamente eliminadas. As demais regiões espectrais e as cinco variáveis físico-químicas foram combinadas para a formação de “meta-espectros”, utilizados na construção da matriz de dados. Novamente, as amostras foram divididas em dois conjuntos de treinamento e teste, utilizando o algoritmo KS, de maneira similar à descrita na seção anterior. Os espectros MIR foram previamente pré-processados usando alisamento Savitzky-Golay e MSC, também como descrito na seção anterior. Na sequência, todas as variáveis foram autoescaladas para a fusão de dados, a fim de compensar os efeitos das diferentes naturezas das variáveis fundidas. O modelo PLS-DA foi inicialmente escolhido com 4 VLs, explicando 78,46% da variância em \mathbf{X} e 32,19% em \mathbf{Y} . As previsões são mostradas na Fig. 34.

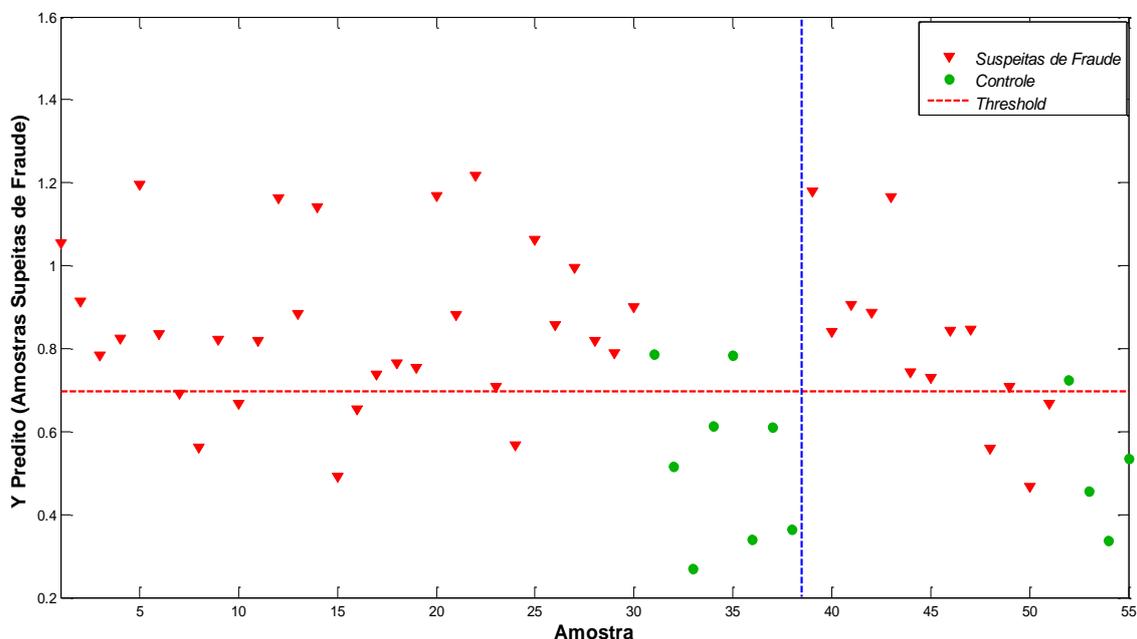


Figura 34: Previsões para o modelo PLS-DA (4 VLs) com fusão dados completa. A linha tracejada em azul separa os conjuntos treinamento e teste.

Neste modelo, para o conjunto de treinamento 6 amostras foram determinadas como falso-negativos em 30 amostras suspeitas, e 2 amostras controle foram consideradas como falso-positivos em 8 possíveis. Já o conjunto de teste apresentou 3 amostras suspeitas como falso-negativos em 13 possíveis e 1 amostra controle dentre 4 como falso-positivo.

Um modelo de fusão de dados utilizando 7 VLs foi testado e apresentou melhores resultados. Este modelo explicou 97,45% da variância em **X** e 58,48% em **Y**, sendo suas previsões apresentadas na Fig. 35.

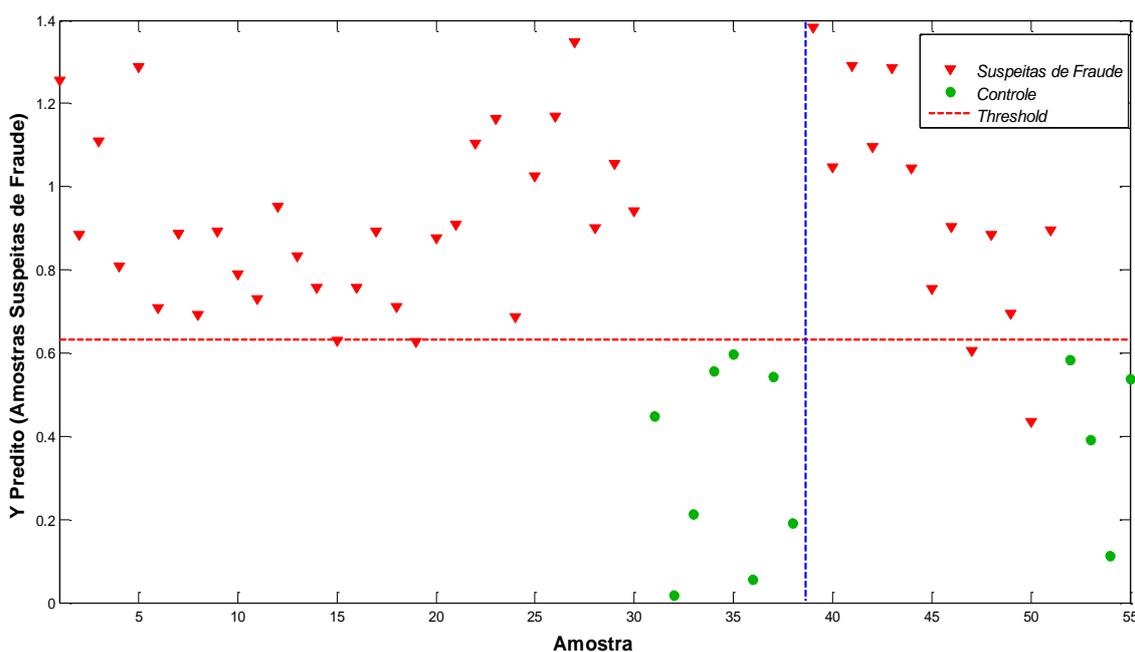


Figura 35: Previsões para o modelo PLS-DA (7 VLs) com fusão dados completa. A linha tracejada em azul separa os conjuntos treinamento e teste.

Neste modelo, para o conjunto de treinamento 2 amostras suspeitas foram determinadas como não adulteradas em 30, enquanto no conjunto de teste, 2 amostras suspeitas foram determinadas como não adulteradas em 13. Nenhuma amostra controle nos dois conjuntos foi prevista erroneamente. As figuras de mérito para este modelo são apresentadas na Tabela 6, indicando que a fusão de dados forneceu melhores previsões que os modelos construídos apenas com os dados físico-químicos ou com os espectros MIR.

Tabela 6: Figuras de mérito para o modelo PLS-DA construído com Fusão de Dados Completa

Conjunto/ Parâmetro	Sensibilidade (%)	Especificidade (%)	TFN (%)	TFP (%)
Treinamento	93,3	100,0	6,7	0,0
Teste	84,6	100,0	15,4	0,0

Assim como no modelo construído somente com os dados físico-químicos, o Cl^- foi a variável com maior *VIPscores* e coeficiente de regressão mais positivo. O comportamento das outras 4 variáveis físico-químicas foi semelhante ao descrito no modelo anterior, sem fusão de dados. A seguir, para facilitar a visualização, serão mostrados apenas os coeficientes de regressão e *VIPscores* relativos aos dados espectrais.

Avaliando-se os Coeficientes de Regressão para as variáveis espectrais (Fig. 36), verifica-se que a região entre 3100 e 2800 cm^{-1} apresentou grande contribuição positiva no modelo, associada à presença de adulterantes na carne, assim como os picos em aproximadamente 2450, 1730, 1170, 1100, 1050, 720 e 670 cm^{-1} . Destaca-se ainda o pico em torno de 605 cm^{-1} , geralmente atribuído à vibração esquelética do anel piranóide de maltodextrina,⁷⁵ um dos adulterantes encontrados na fraude. Picos em torno de 3670, 3650, 3610, 2670, 2610, 2600, 2560, 2550, 2520, 2470, 1770, 1700, 1500, 1430, 1360, 1210, 1130, 1010, 900, 810 e 760 cm^{-1} apresentam intensidades negativas significantes.

Analisando-se o gráfico de *VIPscores* das variáveis espectrais, decidiu-se realizar uma seleção de variáveis, adotando-se um limite de corte de 1,5 (Fig. 37, linha tracejada em vermelho). Este valor de *threshold* é mais rigoroso que o valor de 1,0, normalmente utilizado,⁵⁹ e foi adotado visando selecionar um número menor de regiões espectrais de maior poder discriminante. O objetivo desta seleção de variáveis é construir um novo modelo PLS-DA com menos variáveis, as quais sejam mais seletivas. Além disso, a combinação das regiões espectrais selecionadas com uma única variável físico-química, Cl^- , que é a mais discriminante, buscará gerar um método mais simples e de menor custo. Esse novo modelo será discutido na próxima seção.

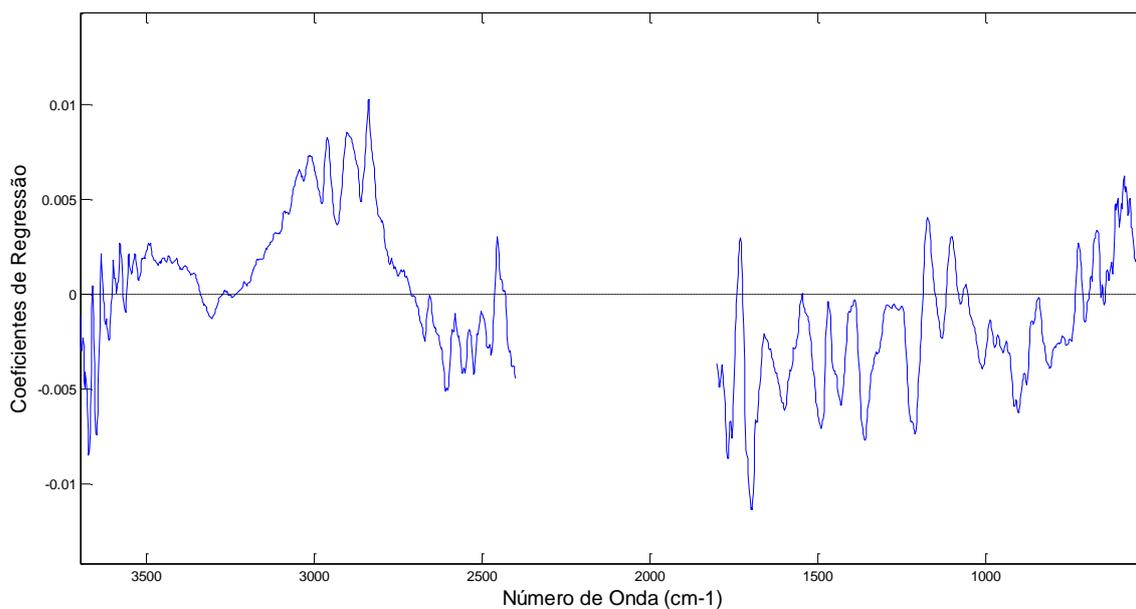


Figura 36: Coeficientes de Regressão das variáveis espectrais do modelo PLS-DA de fusão de dados completo.

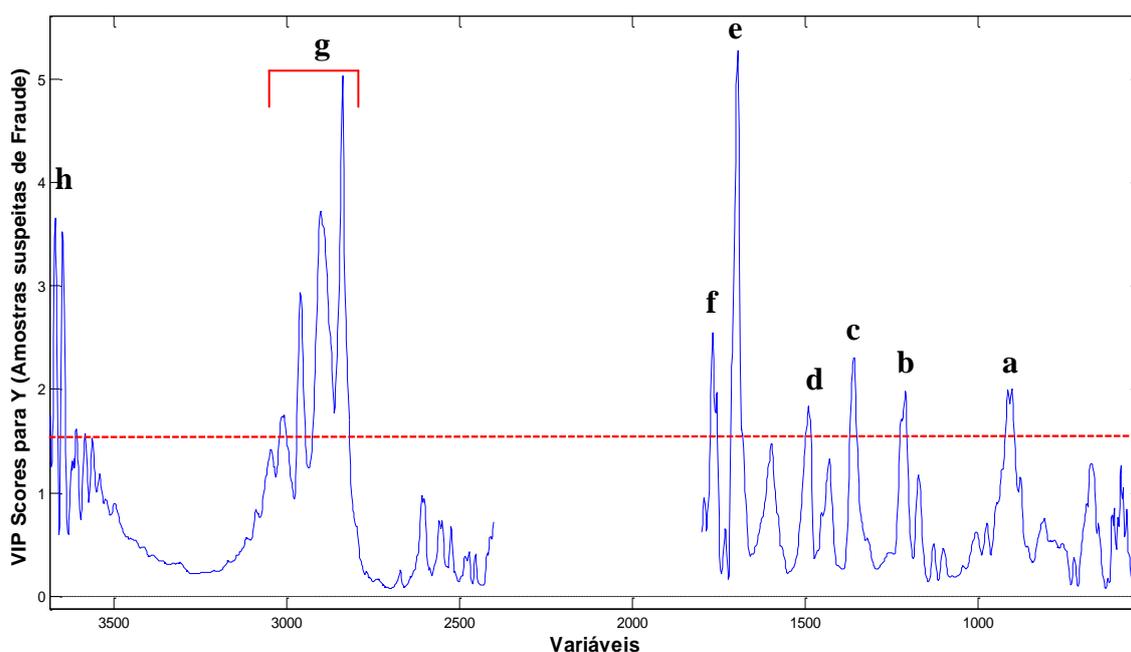


Figura 37: *VIP*scores das variáveis espectrais do modelo PLS-DA de fusão de dados completo.

Portanto, a partir dos *VIP*scores, foram selecionadas oito regiões espectrais (acima do limite de decisão definido em 1,5), nos seguintes intervalos de número de onda: 893-918 cm^{-1} (a), 1204-1230 cm^{-1} (b), 1347-1373 cm^{-1} (c), 1481-1500 cm^{-1} (d), 1677-1717 cm^{-1} (e), 1752-1775 cm^{-1} (f), 2815-2972 cm^{-1} (g) e 3638-3690 cm^{-1} (h). Dentre as regiões significativas selecionadas para este modelo, podem ser feitas algumas atribuições espectrais relacionadas à presença dos adulterantes encontrados. Destacam-se as bandas espectrais em

torno de 1690 e 1230 cm^{-1} , ambas habitualmente associadas às vibrações de agregados de folhas-beta de proteínas provocadas pela adição dos sais NaCl e/ou KCl à carne.²⁰ A segunda região selecionada (b) pode ser também associada à banda de vibração em 1220 cm^{-1} , geralmente atribuída ao estiramento do grupo éster sulfato da carragena.⁷⁴ A banda em 2916 cm^{-1} é usualmente atribuída à deformação axial assimétrica de CH_2 , enquanto que a região entre 1677-1717 cm^{-1} está incluída na faixa de 1600-1700 cm^{-1} , considerada a região mais utilizada no MIR para análise da estrutura secundária de proteínas do colágeno,⁷³ devido à vibração relacionada ao grupo amida I (região e). Finalmente, a região entre 890 e 930 cm^{-1} (região a) comumente está relacionada à presença de tripolifosfato de sódio.⁷⁶

4.1.3.2 Modelo PLS-DA de Fusão de Dados com Seleção de Variáveis

Conforme já mencionado, buscou-se construir um novo modelo PLS-DA de fusão de dados utilizando um menor número de variáveis. Foram combinadas as 8 regiões espectrais do MIR selecionadas no modelo anterior com a variável química de maior poder discriminante, Cl⁻. Os dados foram tratados de maneira similar à descrita para o modelo da seção anterior.

Neste modelo PLS-DA foram utilizadas 6 VLs, as quais explicam 97,31% da variância em **X** e 46,87% em **Y**. As previsões são apresentadas na Fig. 38.

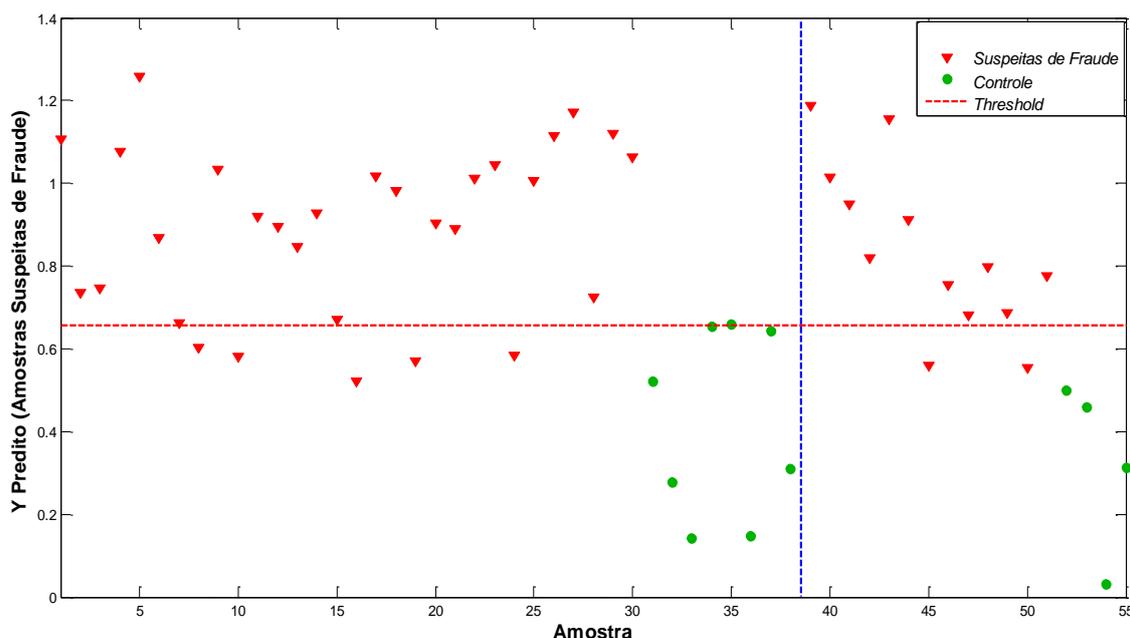


Figura 38: Previsões para o modelo PLS-DA com fusão dados e seleção de variáveis.

Para o conjunto de treinamento, 5 amostras suspeitas foram classificadas como não adulteradas em 30, enquanto apenas 1 amostra controle foi classificada como adulterada. Para o conjunto de teste, 2 amostras suspeitas foram classificadas como não adulteradas em 13, enquanto nenhuma amostra controle foi incorretamente classificada. As figuras de mérito relativas a este modelo são apresentadas na Tabela 7. Observa-se que este modelo apresentou boas taxas de sensibilidade e especificidade, embora a sua capacidade preditiva tenha sido inferior ao modelo de fusão de dados completo, mostrado na seção anterior. Uma comparação em maior profundidade para os diversos modelos PLS-DA construídos para a detecção de amostras de carne adulteradas será feita na próxima seção.

Tabela 7: Figuras de mérito para o modelo PLS-DA construído com Fusão de Dados e seleção de variáveis.

Conjunto/ Parâmetro	Sensibilidade (%)	Especificidade (%)	TFN (%)	TFP (%)
Treinamento	83,3	87,5	16,7	12,5
Teste	84,6	100,0	15,4	0,0

4.1.4 Comparação dos modelos PLS-DA para as amostras reais

Para as amostras reais, apreendidas na operação “Vaca Atolada”, foram construídos quatro modelos PLS-DA: usando apenas os dados físico-químicos (FQ), apenas os espectros MIR, fusão de dados completa e fusão de dados com seleção de variáveis (8 regiões espectrais mais teor de cloreto). Para a comparação global desses modelos, a Tabela 8 apresenta os números de falso-negativos (FN) e falso-positivos (FP) na previsão fornecida por cada um deles, tanto para o conjunto de treinamento, quanto para o conjunto de teste. Além disso, são mostrados os valores da taxa de eficiência (EFR), uma figura de mérito mais rigorosa, que leva em conta tanto a sensibilidade quanto à especificidade qualitativas e está definida na equação 13 (p. 32).

Tabela 8: Comparação dos 4 modelos PLS-DA construídos para as amostras reais.

Modelo	Treinamento			Teste			N° VLs	Variância explicada em X (%)	Variância explicada em Y (%)
	EFR	FN	FP	EFR	FN	FP			
FQ	83,3%	5	0	84,6%	2	0	3	94,16	36,58
MIR	45,8%	5	3	51,9%	3	1	4	95,70	20,71
Fusão completa	93,3%	2	0	84,6%	2	0	7	97,45	58,48
Fusão com sel. de var ^s	70,8%	5	1	84,6%	2	0	6	97,31	46,87

Comparando-se os resultados da Tabela 8, percebe-se claramente que o modelo gerado com os dados físico-químicos é melhor que o modelo construído apenas com dados espectroscópicos (MIR) por apresentar maior eficiência para os conjuntos treinamento e teste utilizando um número de variáveis latentes menor. Entretanto, o modelo com dados físico-químicos utilizou cinco variáveis, cujas medidas consumiram grandes quantidades de insumos, geraram resíduos e demandaram análises morosas, laboriosas, destrutivas e de maior custo para sua construção. Dados de ATR-FTIR são obtidos de forma rápida, limpa e não destrutiva, e permitem a identificação de regiões espectrais mais significativas para a predição.

O modelo de fusão de dados completa, combinando as cinco variáveis físico-químicas e toda a região espectral do MIR (com exceção de regiões de interferência), forneceu os melhores resultados no geral. Este modelo forneceu valores de EFR próximos ou acima de 85%, apresentando apenas 4 amostras suspeitas classificadas como falso-negativos e todas as amostras controle corretamente classificadas. Em contrapartida, o modelo construído apenas com os dados físico-químicos classificou 7 amostras suspeitas no total como falso-negativos.

O modelo de fusão de dados com seleção de variáveis apresentou EFR para o conjunto de teste idêntica ao modelo de fusão completa. Entretanto, ele apresentou 3 falso-negativos e 1 falso-positivo a mais para o conjunto de treinamento. Embora tenha mostrado eficiência inferior ao modelo de fusão completa, a elaboração do modelo de fusão de dados com seleção de variáveis

se justifica por utilizar apenas uma variável físico-química, o teor de cloreto. A economia de tempo, custos e reagentes gerada pela eliminação da necessidade de obter as medidas para as outras quatro variáveis, proteínas, cinzas, sódio e fosfato, compensa a menor eficiência classificatória do modelo.

Em conclusão, o modelo de Fusão de Dados com seleção de variáveis foi considerado o de melhor razão custo/benefício, sendo apropriado como método de verificação inicial para a detecção sistemática do tipo de fraude em estudo em amostras de carne bovina *in natura*.

4.2 Amostras Simuladas

As amostras simuladas foram preparadas pela injeção de quatro tipos de misturas adulterantes, conforme descrito na seção 3.2.1 e de acordo com o planejamento mostrado no Anexo B. Do total de amostras produzidas em nesta seção, 51 peças de carne foram utilizadas para a obtenção de líquido resultante do descongelamento (purga) e outras 18 peças de carne foram utilizadas para obtenção de carne descongelada sem purga, totalmente triturada e homogeneizada. Devido à pequena quantidade de amostras de carne, estas não foram utilizadas para a construção de modelos. Portanto, nesta aplicação, os modelos PLS-DA foram construídos apenas com as amostras de purga.

4.2.1 Modelos PLS-DA

Como nas aplicações anteriores envolvendo espectros MIR as regiões entre 1800 e 2400 cm^{-1} e acima de 3700 cm^{-1} foram previamente eliminadas. Os pré-processamentos usados foram os mesmos: alisamento Savitzky-Golay, MSC, e dados centrados na média. Os espectros registrados no modo de ATR são mostrados na Fig. 39. Como o número de amostras controle obtido foi muito pequeno, apenas três, não faz sentido dividir as amostras em conjuntos de treinamento e de teste. Como alternativa para avaliar a robustez dos modelos com amostras independentes, adotou-se a estratégia de mostrar os resultados obtidos por validação cruzada, ao invés de apresentar a classificação de amostras de um conjunto de teste. Conforme já mencionado, usou-se validação cruzada do tipo *leave-one-out*, na qual uma amostra é retirada de cada vez para ser prevista por um modelo elaborado com as restantes. Portanto, os resultados mostrados serão para a previsão das amostras não incluídas na construção dos modelos, ou seja, a previsão será feita de maneira robusta e não viciada. Inicialmente, elaborou-se um modelo contendo os 4 tipos de amostras adulteradas. Na sequência, construíram-se quatro modelos independentes contendo apenas um tipo de adulteração, com a intenção de melhor observar as alterações espectrais causadas por cada tipo de mistura adulterante.

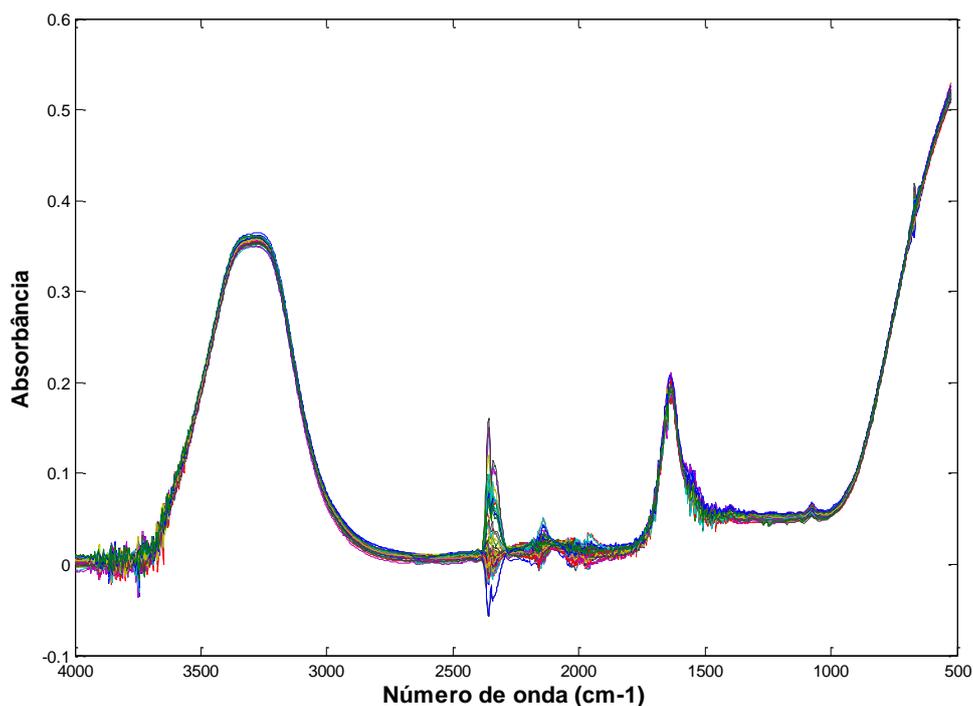


Figura 39: Espectro MIR das 51 amostras simuladas.

4.2.1.1 Modelo PLS-DA com os 4 tipos de adulteração

Tentou-se inicialmente elaborar um modelo PLS2-DA que fosse capaz de classificar as amostras controle e os 4 tipos de adulteração especificamente (5 classes). Para este modelo, não se conseguiram boas previsões para as classes específicas de cada tipo de adulterante. Ou seja, constatou-se que o modelo previa bem as amostras controle, mas não conseguia discriminar cada tipo de adulteração. De certo modo, isto já era esperado, pois as misturas usadas possuem adulterantes em comum, dificultando sua discriminação. Dessa forma, apresenta-se a seguir um modelo PLS1-DA, que foi capaz de discriminar perfeitamente as amostras controle das amostras adulteradas, independente do tipo. Uma observação importante é que nesta aplicação, ao contrário da anterior (amostras reais), a classe controle foi arbitrariamente definida como 1 e a classe das amostras adulteradas como zero.

Neste modelo PLS-DA foram utilizadas 6 VLs, as quais explicam 93,10% da variância em **X** e 32,07% em **Y**. As previsões são mostradas na Fig. 40.

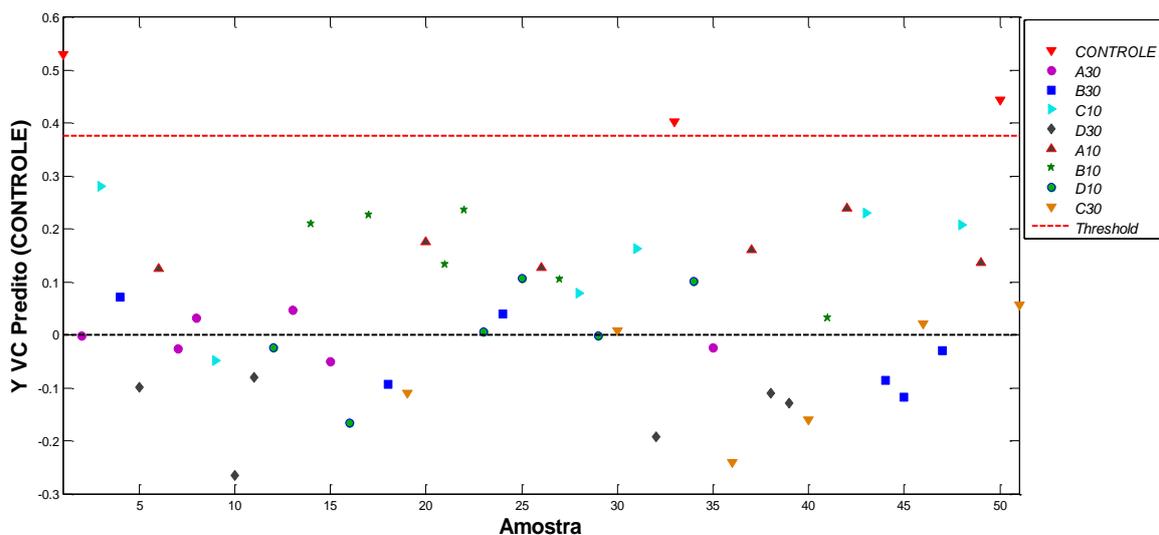


Figura 40: Previsões de validação cruzada do modelo PLS-DA construído com os espectros MIR para as amostras simuladas.

Observando a figura 40, verifica-se que nenhuma amostra injetada foi prevista como não adulterada, enquanto nenhuma amostra controle foi prevista como adulterada. Portanto, este modelo teve 100% de eficiência. Analisando-se os Coeficientes de Regressão (Fig. 41), verificam-se bandas negativas entre 3400 e 3050 cm^{-1} (a), atribuídas ao estiramento da ligação O-H de água; entre 1600-1700 cm^{-1} (b), geralmente atribuída à vibração de deformação simétrica no plano (tesoura) da ligação H-O-H da água;³⁹ em 1157 cm^{-1} (c), comumente atribuída a vibrações de deformação da ligação COH de maltodextrina;⁷⁵ bandas entre 850 e 950 cm^{-1} e entre 1000 e 1150 cm^{-1} (d), em geral características de tripolifosfato de sódio,⁷⁶ e uma banda em aproximadamente 1700 cm^{-1} (e), usualmente descrita como banda característica de amida I de colágeno.⁷³ Estas regiões negativas contribuem mais significativamente para a identificação de amostras como injetadas/adulteradas. Avaliando-se as bandas positivas, verifica-se uma grande região ruidosa entre 1800 e 1300 cm^{-1} e importantes bandas em torno de 2940 cm^{-1} , 2920 cm^{-1} , 1634 cm^{-1} e na região entre 1220 e 1350 cm^{-1} .

Avaliando-se o gráfico de *VIP*scores (Fig. 42), observa-se que as bandas mais significativas para a discriminação encontram-se na região entre 1750 e 1500 cm^{-1} , seguida pelos picos em 670 cm^{-1} , 3175 cm^{-1} e 1076 cm^{-1} . De maneira coerente, as regiões em torno de 3200 e 1600-1700 cm^{-1} estão associadas ao maior teor de água das amostras adulteradas. Na sequência,

serão construídos modelos específicos para a discriminação de cada tipo de mistura adulterante.

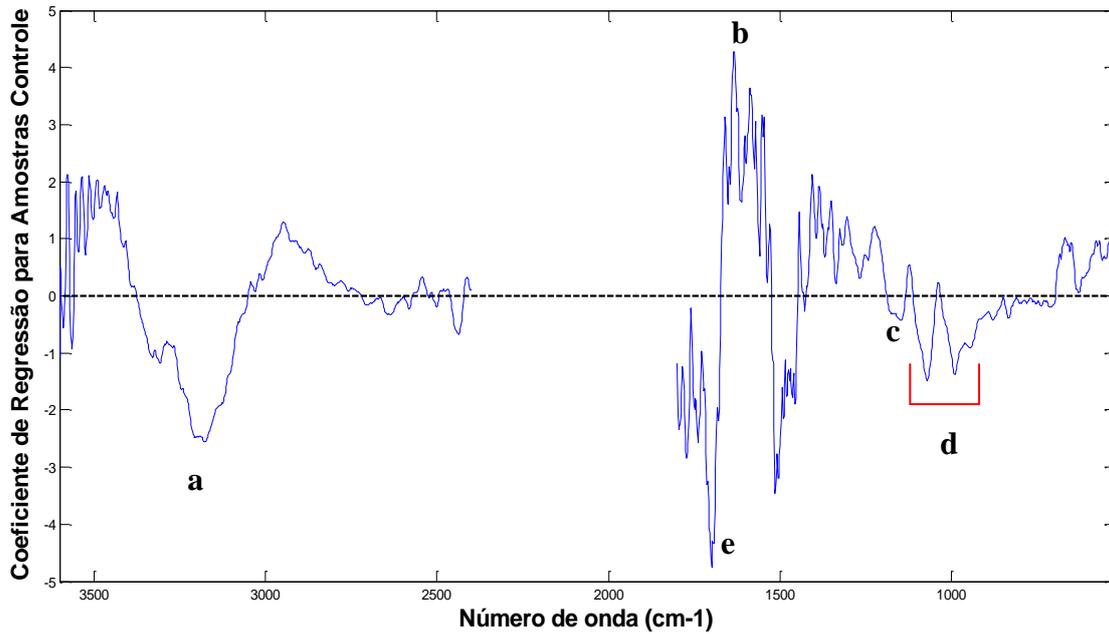


Figura 41: Coeficientes de Regressão do modelo PLS-DA para as amostras simuladas.

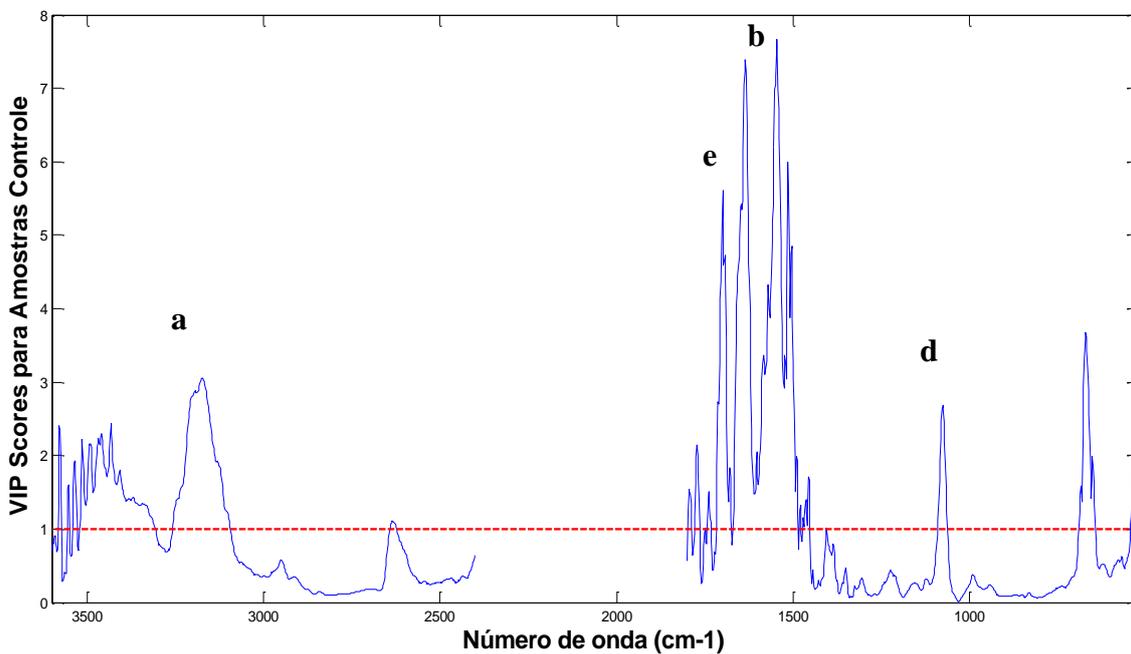


Figura 42: VIP Scores do modelo PLS-DA para as amostras simuladas.

4.2.1.2 Modelo PLS-DA para amostras adulteradas por injeção de água

Este modelo foi elaborado com 15 amostras, sendo 3 amostras controle e 12 adulteradas por injeção de água (mistura A). A injeção de água

corresponde ao tipo de adulteração mais rudimentar a ser feita nas amostras de carne. O grupo de 12 amostras adulteradas foi composto por 6 adulterações num nível inferior, injeção de uma massa da mistura correspondendo a 10 % m/m da carne (legenda A10 na Fig. 43), e outras 6 num nível superior, injeção de uma massa da mistura correspondendo a 30 % m/m da carne (A30). Os outros três tipos de adulterações, com os quais foram elaborados os modelos descritos nas seções 4.2.1.3-5, foram planejados de maneira similar (Anexo B).

Neste modelo foram utilizadas 5 VLs, as quais explicam 95,25% da variância em **X** e 65,75% em **Y**. As previsões são apresentadas na Fig. 43.

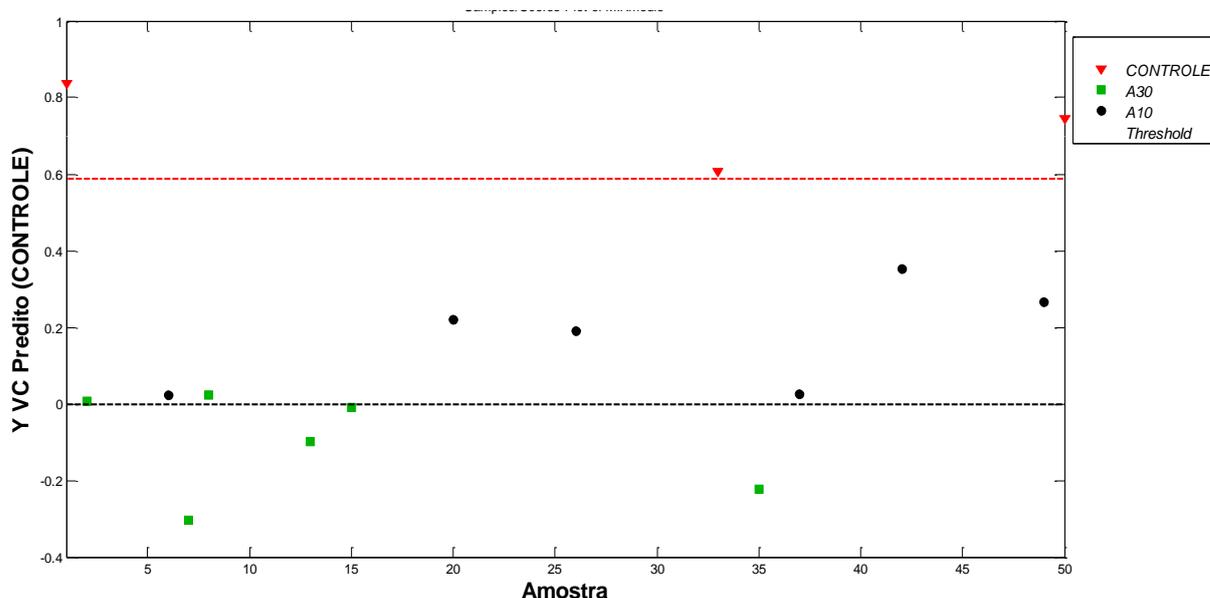


Figura 43: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção de água.

Observa-se nesta figura que nenhuma amostra controle foi prevista incorretamente e todas as amostras adulteradas foram corretamente classificadas. Além disso, verifica-se uma tendência de agrupamento das amostras com injeções a 10% em valores positivos de Y predito, enquanto as amostras com injeções a 30% tendem a se localizar em valores negativos de Y predito, indicando coerentemente que as amostras adulteradas num nível superior estão mais distantes em sua composição das amostras controle.

Avaliando-se os Coeficientes de Regressão para a classe controle (Fig. 44) observam-se bandas positivas acima de 3400 cm^{-1} , entre as regiões $3100\text{-}2800\text{ cm}^{-1}$, $1700\text{-}1500\text{ cm}^{-1}$, $1450\text{-}1050\text{ cm}^{-1}$, $690\text{-}654\text{ cm}^{-1}$ e abaixo de 580 cm^{-1} . Bandas negativas situam-se nas regiões de $3400\text{-}3100\text{ cm}^{-1}$, $2800\text{-}2400$

cm^{-1} , $1800\text{-}1700\text{ cm}^{-1}$, $1520\text{-}1460\text{ cm}^{-1}$, $1050\text{-}850\text{ cm}^{-1}$, $780\text{-}690\text{ cm}^{-1}$ e $640\text{-}580\text{ cm}^{-1}$.

Avaliando-se o gráfico de *VIP*scores para as amostras controle (Fig. 45), observa-se que as bandas mais significativas estão em aproximadamente 1635 cm^{-1} , 1550 cm^{-1} e 670 cm^{-1} . Considerando-se em conjunto os *VIP*scores e os coeficientes de regressão, destaca-se, como já discutido na seção anterior, a presença de vibrações entre $3100\text{-}3400\text{ cm}^{-1}$ e $1600\text{-}1700\text{ cm}^{-1}$ (a), geralmente atribuídas à molécula de água, associadas às amostras adulteradas.

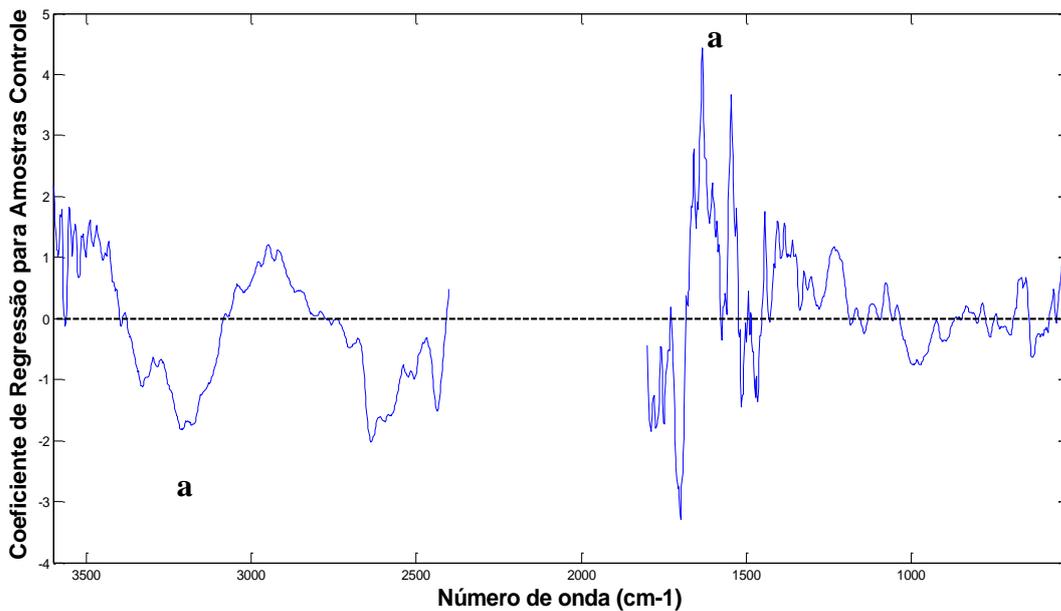


Figura 44: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção de água.

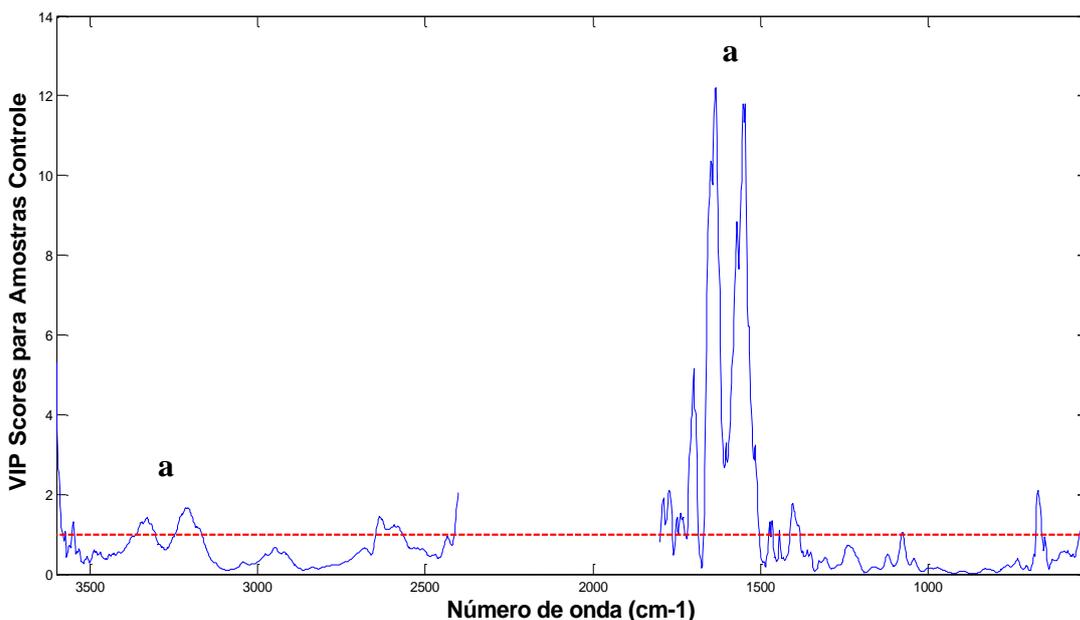


Figura 45: *VIP*scores do modelo PLS-DA para as amostras adulteradas por injeção de água.

4.2.1.3 Modelo PLS-DA Modelo PLS-DA para amostras adulteradas por injeção de tripolifostafo

O conjunto de dados analisado neste modelo é composto por 3 amostras controle e 12 amostras adulteradas pela injeção de soluções de tripolifosfato de sódio (mistura B), sendo 6 adulteradas no nível superior de 30% m/m (B30) e 6 no nível inferior de 10% m/m (B10). O tripolifostafo foi uma das principais substâncias encontradas nos frigoríficos autuados na operação “Vaca Atolada”.

Neste modelo PLS-DA foram utilizadas 3 VLs, as quais explicam 83,96% da variância em **X** e 56,47% em **Y**. As previsões são apresentadas na Fig. 46.

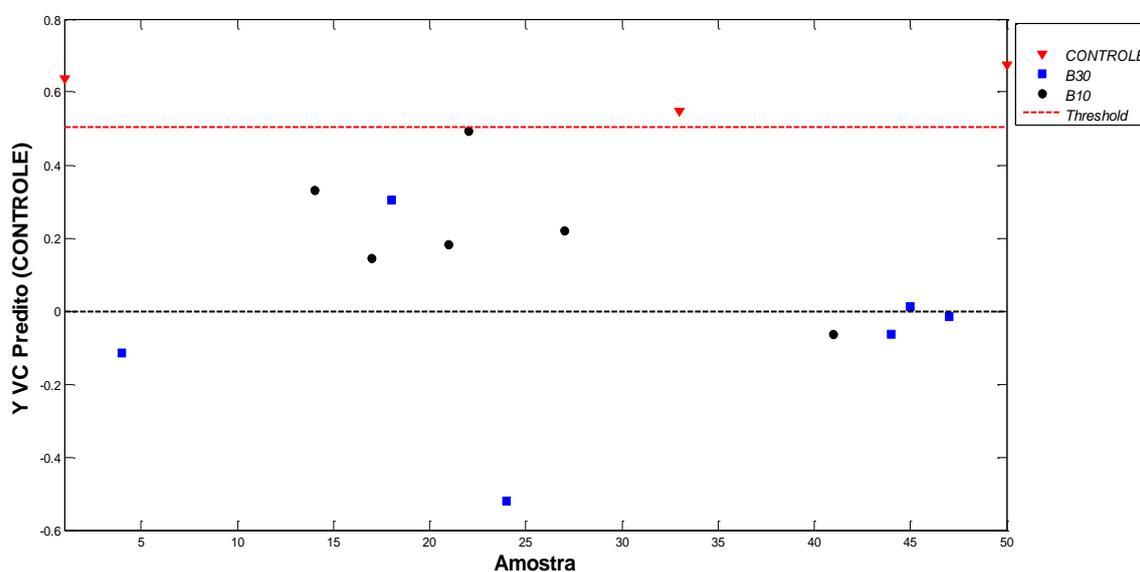


Figura 46: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção de solução de tripolifosfato de sódio.

Assim como para o modelo anterior, observou-se uma eficiência de 100%, sem a ocorrência de nenhum falso-positivo ou falso-negativo. Também se observou novamente a tendência de agrupamento das previsões para amostras menos adulteradas próximas às amostras controle, com uma exceção.

Avaliando o gráfico de coeficientes de regressão (Fig. 47), observa-se na região negativa (que contribuem mais significativamente para a identificação das amostras injetadas) bandas entre $3400\text{-}3100\text{ cm}^{-1}$, $2690\text{-}2550\text{ cm}^{-1}$, $1800\text{-}1670\text{ cm}^{-1}$, $1200\text{-}700\text{ cm}^{-1}$, e, especialmente, a banda entre $1150\text{-}1000\text{ cm}^{-1}$ (a), característica do tripolifosfato de sódio⁷⁶ adicionado às amostras simuladas.

As bandas positivas mais significativas para a previsão das amostras controle estão nas regiões entre 3100 e 2700 cm^{-1} , 1660 e 1200 cm^{-1} e 700 a 640 cm^{-1} .

Avaliando o gráfico de *VIPscores* (Fig. 48) para este modelo, destacam-se as bandas em 1635 cm^{-1} , 1545 cm^{-1} , 670 cm^{-1} , 1075 cm^{-1} , 1700 cm^{-1} , 3580 cm^{-1} e 2630 cm^{-1} . Percebe-se que uma banda específica do tripolifosfato é a quarta banda de maior importância no modelo. Bandas em torno de 1635 cm^{-1} (b) são geralmente atribuídas a resíduos de iminas de estruturas de colágeno, que é um componente natural da carne bovina.⁷³ As bandas de coeficientes de regressão negativo em torno de 3100-3400 cm^{-1} e 1600-1700 cm^{-1} , esta última a de maior valor de *VIPscores*, são indicativas da absorção de maior teor de água pelas amostras adulteradas.

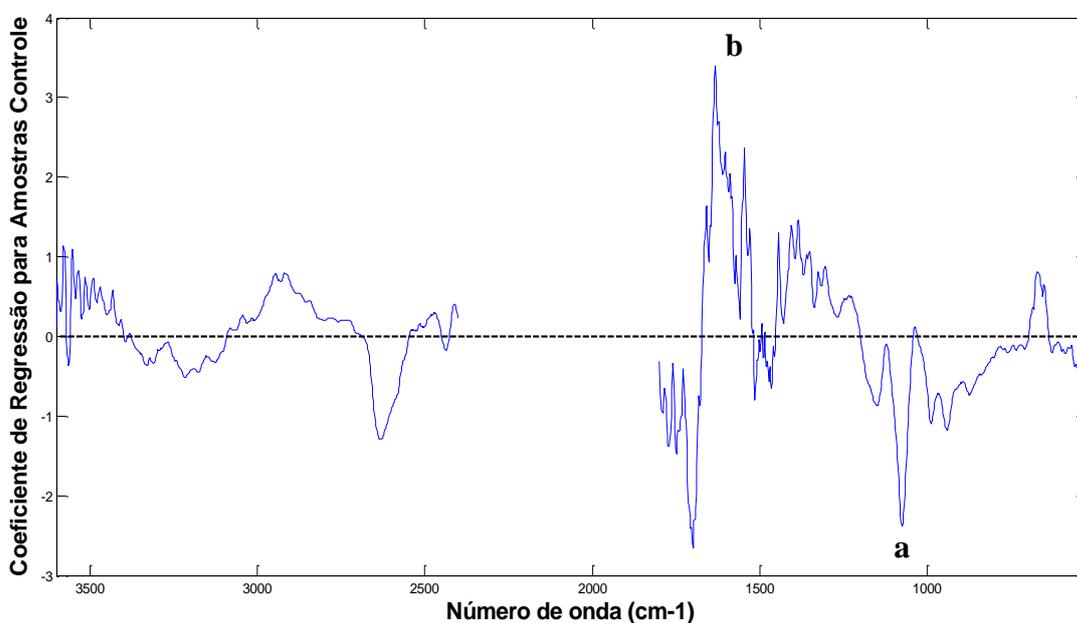


Figura 47: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção de solução de tripolifosfato de sódio.

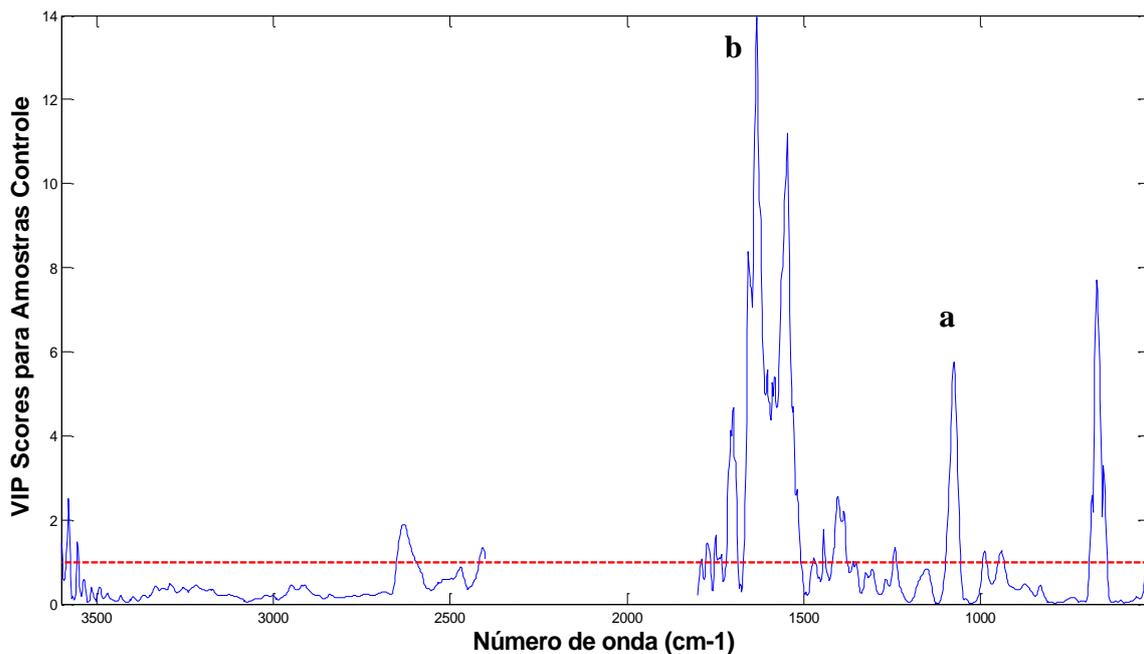


Figura 48: *VIP*scores do modelo PLS-DA para as amostras adulteradas por injeção de solução de tripolifosfato de sódio.

4.2.1.4 Modelo PLS-DA Com Separação em classes Controle e Grupo C

Neste modelo, as amostras de carne foram adulteradas por injeção de uma mistura comercial contendo os principais adulterantes encontrados na operação “Vaca Atolada”: tripolifosfato de sódio, maltodextrina, sal refinado iodado (NaCl), pirofosfato ácido de sódio e carragena. Esta foi denominada mistura C. Novamente, o modelo foi construído com apenas 15 amostras: 3 amostras controle e 12 amostras adulteradas em dois níveis de adição, 10% (C10) e 30% (C30). No modelo PLS-DA foram utilizadas 4 VLs, as quais explicam 92,72% da variância em **X** e 79,49% em **Y**. As previsões são mostradas na Fig. 49.

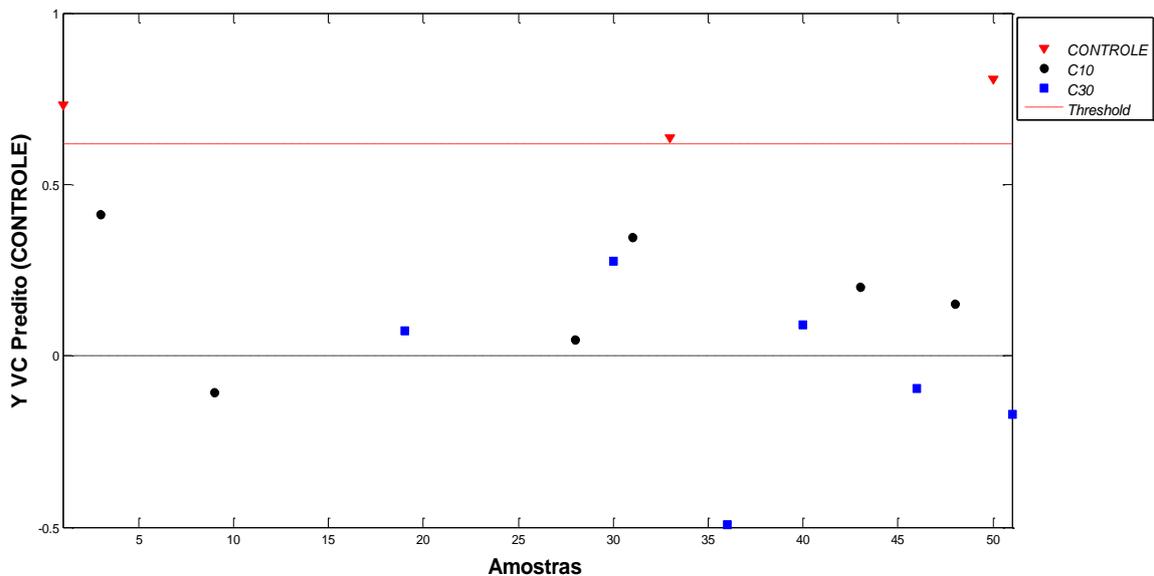


Figura 49: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção da mistura C.

Assim como para os modelos anteriores, a eficiência deste foi de 100%. No entanto, a tendência das amostras adulteradas em se agruparem em função do nível de adição não foi tão nítida.

Avaliando-se os Coeficientes de Regressão para este modelo (Fig. 50) observam-se bandas positivas entre as regiões 3080 e 2740 cm^{-1} , 2500 e 2450 cm^{-1} , 1665 e 1525 cm^{-1} , 1450 a 1200 cm^{-1} , 1130 e 1100 cm^{-1} , 1150 a 1010 cm^{-1} e abaixo de 700 cm^{-1} . Bandas negativas situam-se nas regiões 3460 a 3080 cm^{-1} , 2740 a 2500 cm^{-1} , 1800 a 1670 cm^{-1} , 1520 a 1450 cm^{-1} , 1200 a 1125 cm^{-1} , 1100 a 1050 cm^{-1} , 1000 a 850 cm^{-1} e 810 a 750 cm^{-1} . Destaca-se a banda negativa entre 1150 e 1000 cm^{-1} (a), usualmente atribuída ao tripolifosfato de sódio,⁷⁶ que foi adicionado às amostras simuladas e bandas próximas a 1080 e 1155 cm^{-1} (b), habitualmente atribuídas ao estiramento de CO e à deformação axial de COH da maltodextrina.⁷⁵ Novamente, identificam-se coeficientes negativos para as bandas características de vibrações de água (c), além do pico em torno de 1220-1230 cm^{-1} (d), região na qual geralmente ocorrem vibrações relacionadas tanto ao efeito da adição de sais na configuração das proteínas da carne,²⁰ quanto à presença do adulterante carragena.⁷⁴

Avaliando-se o gráfico de *VIP*scores para as amostras controle (Fig. 51), observa-se que as bandas mais significativas foram 1550 cm^{-1} , 673 cm^{-1} , 1633 cm^{-1} , 1693 cm^{-1} e 1400 cm^{-1} . Bandas em torno de 1635 cm^{-1} (e) são, em geral,

atribuídas a resíduos de iminas e em torno de 1552 cm^{-1} (f) comumente associadas à amida II de estruturas de colágeno, que é um componente natural da carne bovina.⁷³

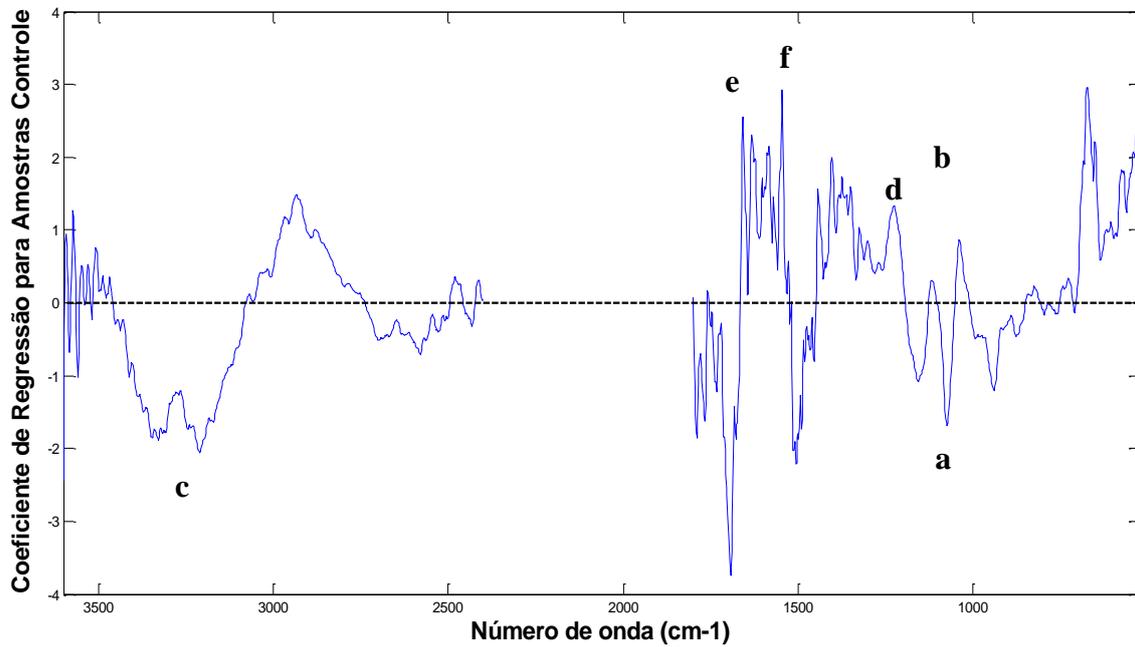


Figura 50: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção da mistura C.

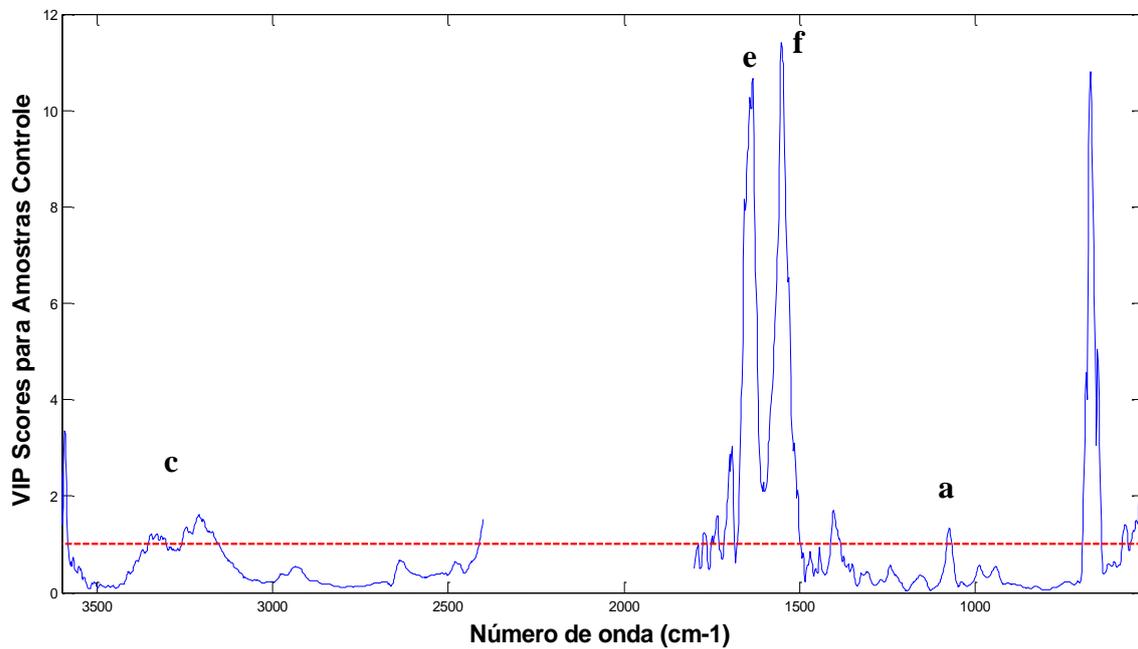


Figura 51: VIP Scores do modelo PLS-DA para as amostras adulteradas por injeção da mistura C.

4.2.1.5 Modelo PLS-DA para amostras adulteradas por injeção da mistura D

A mistura adulterante usada neste último modelo (mistura D) diferiu da usada no modelo anterior apenas pela presença a mais de colágeno.

No modelo PLS-DA foram utilizadas 3 VLs, as quais explicam 84,00% da variância em **X** e 72,81% em **Y**. As previsões são mostradas na Fig. 52.

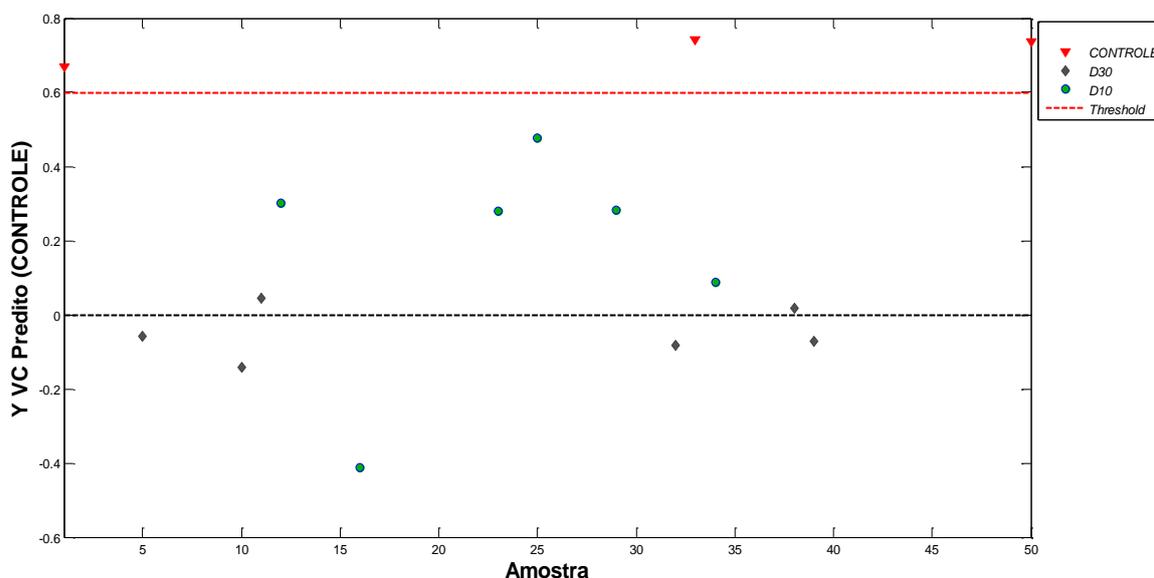


Figura 52: Previsões de validação cruzada do modelo PLS-DA construído com espectros MIR para as amostras adulteradas por injeção da mistura D.

Assim como em todos os outros modelos anteriores, a eficiência preditiva foi de 100%. O agrupamento das amostras adulteradas em função do nível de adição foi nítido, com exceção de uma amostra.

Avaliando o gráfico de coeficientes de regressão (Fig. 53), observa-se na região a negativa (que contribuem mais significativamente para a identificação das amostras injetadas) bandas entre 3420 e 3080 cm^{-1} , 2870 e 2550 cm^{-1} , 1800 e 1670 cm^{-1} , 1580 a 1570 cm^{-1} , 1480 a 1460 cm^{-1} , 1190 a 1110 cm^{-1} , 1010 a 850 cm^{-1} , abaixo de 630 cm^{-1} . Ressalta-se a banda na região de 3309 cm^{-1} (a), geralmente atribuída à absorção de amida de colágeno, assim como a banda entre 1600 e 1700 cm^{-1} (b), comumente relacionada a estruturas secundárias de proteínas de colágeno, especialmente em 1675 cm^{-1} (c), habitualmente atribuída à vibração da ligação entre folhas beta (β -turns) do colágeno.⁷³ Outra região negativa importante para caracterização de amostras

injetadas está entre 890 e 930 cm^{-1} (d), usualmente atribuída a tripolifosfato de sódio.⁷⁶

As bandas positivas mais significativas para a previsão das amostras controle estão nas regiões acima de 3440 cm^{-1} , entre 3080 e 2850 cm^{-1} , 1670 e 1580 cm^{-1} , 1570 e 1520 cm^{-1} , 1450 e 1200 cm^{-1} , 1130 e 1010 cm^{-1} e 850 a 630 cm^{-1} .

Avaliando o gráfico de *VIP*scores (Fig. 54) para este modelo, destacam-se as bandas em 1545 cm^{-1} , 1633 cm^{-1} , 3592 cm^{-1} e 1400 cm^{-1} .

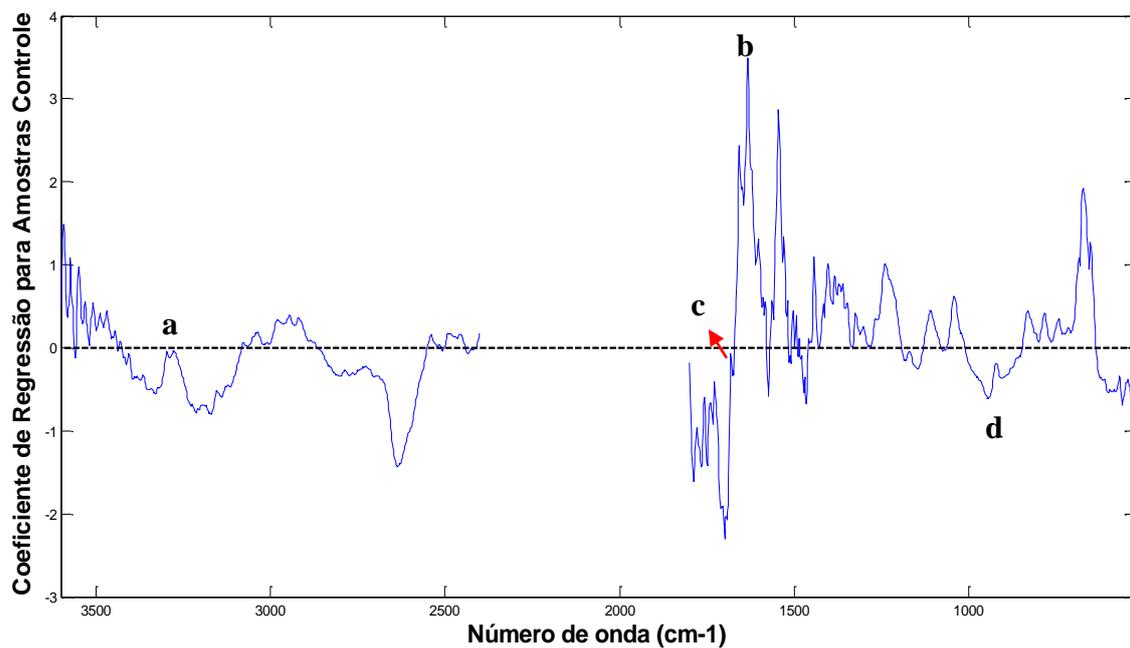


Figura 53: Coeficientes de Regressão do modelo PLS-DA para as amostras adulteradas por injeção da mistura D.

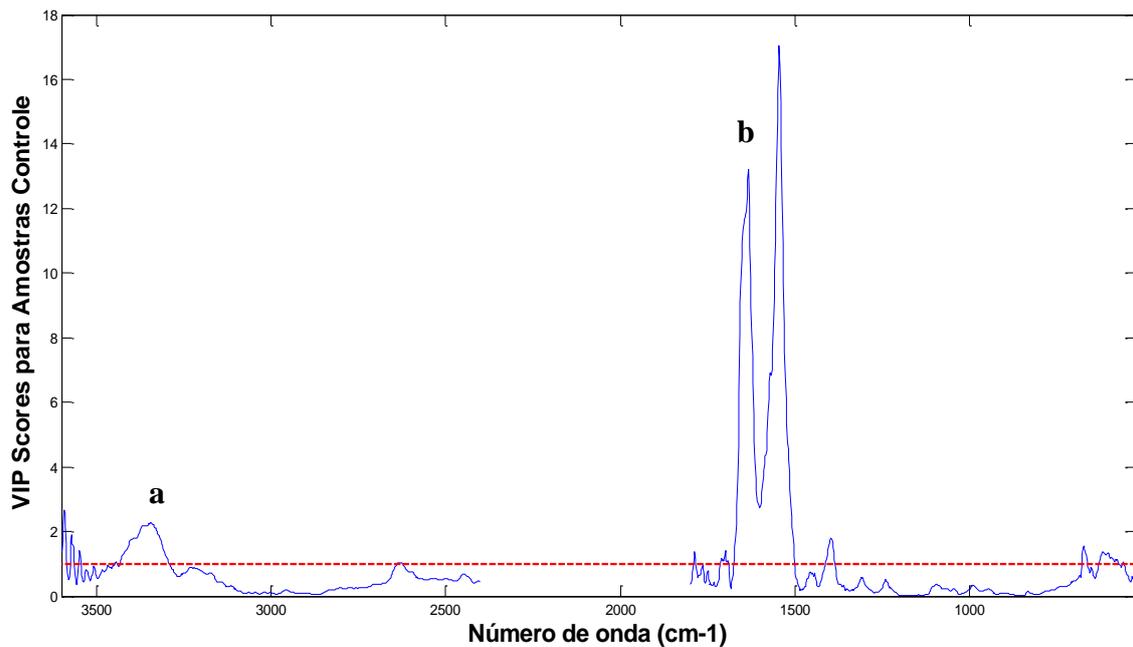


Figura 54: *VIP*scores do modelo PLS-DA para as amostras adulteradas por injeção da mistura D.

4.2.2 Conclusão parcial

Nesta segunda aplicação foi possível desenvolver modelos PLS-DA usando espectros MIR com eficiência classificatória total na discriminação de três amostras de purga, oriundas de amostras de carne controle, em relação a amostras propositadamente adulteradas com várias substâncias que buscaram simular as fraudes encontradas em situações reais, embora tenha sido utilizado um pequeno universo amostral.

5 Conclusão

Esta Dissertação buscou desenvolver modelos quimiométricos para a detecção e caracterização de fraudes em amostras de carne bovina *in natura*, utilizando medidas físico-químicas (proteínas, cinzas, cloreto, sódio e fosfato) e espectros de infravermelho médio. O tipo de adulteração estudado envolveu a injeção de sais e outras substâncias, como colágeno, maltodextrina e carragena, em amostras de carne, de acordo com ocorrências encontradas em situações reais.

A utilização de métodos analíticos para identificação dos constituintes da carne fraudada é necessária, pois outros fatores podem ter levado aos indícios visuais de grande quantidade de líquido exsudado nas extremidades da peça. Além disso, devido à presença de diferentes graus de adulteração, em concentrações mais baixas não é possível a visualização desta fraude. Para tanto, a análise multivariada surge como ferramenta para identificação de padrões e classificação, podendo ser utilizada como análise preliminar (*screening*) de amostras suspeitas.

Numa primeira aplicação, foram construídos modelos a partir de amostras de carne apreendidas em uma operação real da Polícia Federal. A oportunidade de construir modelos com amostras reais foi um dos aspectos de maior relevância deste trabalho. Inicialmente, modelos PCA usando apenas os dados físico-químicos foram construídos de maneira semi-supervisionada, estimando uma elipse de confiança no nível de 90% em torno das amostras controle, sendo capazes de detectar 74% das amostras suspeitas como adulteradas. Modelos PCA construídos com espectros MIR não foram capazes de propiciar uma separação visual das amostras entre controle e adulteradas.

Na sequência, modelos de classificação supervisionada PLS-DA foram construídos buscando melhorar a capacidade de detecção de fraudes. O modelo utilizando apenas os dados físico-químicos apresentou 85% de eficiência (figura de mérito que considera tanto a ocorrência de falso-negativos quanto falso-positivos). O modelo PLS-DA construído apenas com os espectros MIR apresentou uma eficiência bem inferior (52%). No entanto, a fusão dos dados físico-químicos com os espectros MIR permitiu elaborar o modelo PLS-

DA de melhor capacidade preditiva, apresentando uma eficiência de 93% e deixando de detectar apenas 4 amostras adulteradas (4 erros falso-negativos).

Considerando-se as vantagens da espectroscopia MIR sobre os métodos usados nas determinações físico-químicas, tais como maior rapidez, menor custo, análise limpa e não destrutiva, construiu-se um último modelo PLS-DA utilizando seleção de variáveis a fim de se obter uma melhor razão custo/benefício. Das cinco variáveis físico-químicas, foi escolhida apenas a mais discriminante, que foi o teor de cloreto. A partir dos espectros MIR inteiros, foram selecionadas 8 regiões espectrais mais discriminatórias utilizando o critério de *VIPscores* acima de 1,5. O modelo PLS-DA de fusão do teor de cloreto com as regiões mais discriminantes do espectro MIR utilizou apenas 191 das 1808 variáveis originais usadas no modelo de fusão completa. Embora, tenha sido observada uma queda da capacidade preditiva em relação ao modelo de fusão completa, este modelo foi considerado satisfatório, apresentando uma taxa de eficiência de 85%. A queda na capacidade preditiva foi compensada pelo ganho de custo na geração dos dados, relacionado à economia de não utilizar as determinações de teores de proteínas, cinzas, sódio e fosfato. Além disso, a seleção de variáveis permitiu caracterizar o efeito das fraudes nos espectros MIR, através das atribuições das bandas selecionadas a vibrações específicas relacionadas aos adulterantes.

Na segunda aplicação, tentou-se reproduzir as fraudes usando amostras que foram geradas de maneira controlada, com um mínimo de variação, já que a carne adulterada foi originária do mesmo animal. Foram preparados quatro tipos de adulteração, por injeção de água, solução comercial de tripolifosfato de sódio e de outras duas misturas contendo adulterantes, como carragena, maltodextrina, NaCl e colágeno. Ao invés da carne, foram analisadas por espectroscopia MIR amostras de purga, o líquido extraído da carne. Em função do pequeno número de amostras controle, apenas três, os modelos foram baseados nas previsões de validação cruzada. Todos os modelos foram capazes de discriminar as amostras controle dos diversos tipos de fraude. Além disso, foi feita uma caracterização espectroscópica, indicando quais as bandas relacionadas à injeção de cada tipo de mistura adulterante.

Este estudo abre oportunidades para novas pesquisas sobre fraudes em carnes e seus diferentes tipos utilizando técnicas analíticas cada vez mais

sensíveis e métodos quimiométricos mais complexos. Na continuidade deste projeto, pretende-se aplicar técnicas analíticas que gerem dados de segunda ordem na detecção de novas fraudes utilizando um número de amostras maior. Dentre estas técnicas, destacam-se a cromatografia gasosa bidimensional abrangente (GCxGC), na análise de ácidos graxos extraídos da carne, e a espectroscopia de imagem vibracional no infravermelho,²² na análise de carne. Essas técnicas permitem extrair uma quantidade maior de informação dos dados, mas demandam também o uso de métodos quimiométricos mais complexos. Este será o foco do projeto de Doutorado já aprovado no Programa de Pós-Graduação em Química da UFMG pela autora desta Dissertação, cujo título é “Tratamento Quimiométrico de Dados Bidimensionais na Detecção e Caracterização de Fraudes em Carnes Bovinas *in natura*”.

6 Referências Bibliográficas

- 1 Guenther, R. Controle sanitário dos alimentos. São Paulo: Livraria Atheneu, 1992.
- 2 Camargo, R. D.; Fonseca, H.; Grainer, N.; Filho, L.G.P.; Caruso, L.G.B; Andrade, M.O; Nogueira, J.N.; Cantarelli, P.R.; Lima, U.A., Oliveira, A.J.; Meirao, L.S.. Tecnologia dos produtos agropecuários – Alimentos. São Paulo: Livraria Nobel, 1984.
- 3 Valle, E. R. d. Mitos e Realidades sobre o consumo de carne bovina. Disponível em: < <http://www.cnpqg.embrapa.br/publicacoes/doc/doc100/> >. Acesso em: 02/04/2015.
- 4 Tolentino, V. Oficina mostra importância do consumo de carne bovina. 2013. Disponível em: < <http://www.emater.pr.gov.br/modules/noticias/article.php?storyid=3449> >. Acesso em: 02/04/2015.
- 5 Evangelista, J. Tecnologia de Alimentos. São Paulo: Editora Atheneu, 2005.
- 6 Hargin, K. D. Authenticity issues in meat and meat products. Meat Science, 43, S277-S289, 1996.
- 7 DPF. Notícia: Justiça condena acusados de adulterar leite presos na operação Ouro Branco. Uberaba/MG, 2012. Disponível em: < <http://www.dpf.gov.br/agencia/noticias/2012/marco/justica-condena-acusados-de-adulterar-leite-presos-na-operacao-ouro-branco> >. Acesso em: 02/04/2015
- 8 Empresas de transporte de leite adulteram produto no RS, 2013. Disponível em: < <http://veja.abril.com.br/noticia/economia/empresas-de-transporte-de-leite-adulteram-produto-no-rs/> >. Acesso em 02/04/2015.
- 9 DPF. Notícia: PF deflagra Operação Vaca Atolada em BH, Contagem e Ibirité. Belo Horizonte/MG, 2012. Disponível em: < <http://www.dpf.gov.br/agencia/noticias/2012/junho/a-pf-deflagrou-hoje-15-06-a-operacao-vaca-atolada-em-bh-contagem-e-ibirite> >. Acesso em: 02/04/2015.
- 10 Nakagawa, F. Escândalo da venda ilegal de carne de cavalo avança na Europa., Londres, 2013. Disponível em: < <http://economia.estadao.com.br/noticias/geral,escandalo-da-venda-ilegal-de-carne-de-cavalo-avanca-na-europa-imp-,996412> >. Acesso em: 03/04/2015.

- 11 Alamprese, C.; Casale, M.; Sinelli, N.; Lanteri, S.; Casiraghi, E. Detection of minced beef adulteration with turkey meat by UV-vis, NIR and MIR spectroscopy. *LWT-Food Science and Technology*, 53, 225-232, 2013.
- 12 Iwobi, A.; Sebah, D.; Kraemer, I.; Losher, C.; Fischer, G.; Busch, U.; Huber, I. A multiplex real-time PCR method for the quantification of beef and pork fractions in minced meat. *Food Chemistry*, 169, 305-313, 2015.
- 13 Zajac, A.; Hanuza, J.; Dyminska, L. Raman spectroscopy in determination of horse meat content in the mixture with other meats. *Food Chemistry*, 156, 333-338, 2014.
- 14 Mamani-Linares, L. W.; Gallo, C.; Alomar, D. Identification of cattle, llama and horse meat by near infrared reflectance or transreflectance spectroscopy. *Meat Science*, 90, 378-385, 2012.
- 15 Rohman, A.; Siswindari; Erwanto, Y.; Man, Y.B.C.. Analysis of pork adulteration in beef meatball using Fourier transform infrared (FTIR) spectroscopy. *Meat Science*, 88, 91-95, 2011.
- 16 Ballin, N. Z. Authentication of meat and meat products. *Meat Science*, 86, 577-587, 2010.
- 17 Grundy, H. H.; Reece, P.; Sykes, M.D.; Clough, J.A.; Audsley, N.; Stones, R.. Method to screen for the addition of porcine blood-based binding products to foods using liquid chromatography/triple quadrupole mass spectrometry. *Rapid Communications in Mass Spectrometry*, 22, 2006-2008, 2008.
- 18 L59. Commission Recommendation (EC) concerning a coordinated programme for the official control of foodstuffs. *Official Journal of the European Union*, 27-39, 2005.
- 19 Cheng, Q.; Sun, C. W. Factors affecting the water holding capacity of red meat products: A review of recent research advances. *Critical Reviews in Food Science and Nutrition*, 48, 137–159, 2008.
- 20 Perisic, N., Afseth, N.K.; Ofstad, R.; Narum, B.; Kohler, A. Characterizing salt substitution in beef meat processing by vibrational spectroscopy and sensory analysis. *Meat Science*, 95, 576-585, 2013.
- 21 Perisic, N., Afseth, N.K.; Ofstad, R.; Kohler, A. Monitoring Protein Structural Changes and Hydration in Bovine Meat Tissue Due to Salt Substitutes by Fourier Transform Infrared (FTIR) Microspectroscopy. *Journal of Agricultural and Food Chemistry*, 59, 10052-10061, 2011.

- 22 Perisic, N., Afseth, N.K.; Ofstad, R.; Scheel, J.; Kohler, A. FTIR Imaging for Structural Analysis of Frankfurter Sausages Subjected to Salt Reduction and Salt Substitution. *Journal of Agricultural and Food Chemistry*, 61, 3219-3228, 2013.
- 23 Ayadi, M. A.; Kechaou, A.; Makni, I.; Attia, H. Influence of carrageenan addition on turkey meat sausages properties. *Journal of Food Engineering*, 93, 278-283, 2009.
- 24 MAPA. Instrução Normativa SDA 20/1999. Secretaria de Defesa Agropecuária. Brasil: Ministério da Agricultura, Pecuária e Abastecimento, 1999.
- 25 MAPA. Bovinos e bubalinos. Brasil, 2015.
- 26 EMBRAPA. Conhecendo a carne que você consome: Qualidade da carne bovina. Campo Grande: Embrapa Gado de Corte, 1999.
- 27 Alencar, F. A.. Potencial das carnes salgadas para o estado do Mato Grosso do Sul. Disponíveis em: <http://cloud.cnpqg.embrapa.br/origem2012/files/2012/11/PotencialCarnesSalgadas_FlodoaldoAlencar.pdf>. Acesso em 02/04/2015.
- 28 Oliveira, A. L. Maciez da carne bovina. Belo Horizonte: Cadernos técnicos - Escola de Veterinária - UFMG, Belo Horizonte, 7-18, 2000.
- 29 Machado, M. M. Efeito do congelamento e estocagem sobre a qualidade da carne bovina. (Dissertação de Mestrado). Escola de Veterinária, UFMG, Belo Horizonte, 2009.
- 30 IBGE. Indicadores IBGE Estatística da Produção Pecuária: IBGE - Instituto Brasileiro de Geografia e Estatística, 2015.
- 31 UNICAMP. Tabela brasileira de composição de alimentos - Tabela TACO, 4ª edição. Campinas: NEPA - UNICAMP, 2011.
- 32 Embrapa. Noções de Ciência da Carne. Disponível em: <<http://www.cnpqg.embrapa.br/publicacoes/doc/doc77/03nocoescarne.html#3.3.2>>. Acesso em 03/02/2015.
- 33 Borrás, E.; Ferré, J.; Boqué, R.; Mestres, M.; Aceña, L.; Busto, O.. Data fusion methodologies for food and beverage authentication and quality assessment – A review. *Analytica Chimica Acta*, 2015, no prelo.
- 34 Fonseca, P. B. R.; Filho, A. S. Histórico da Inspeção dos Alimentos de Origem Animal no Brasil. Disponível em: <<http://imanet.ima.mg.gov.br/nova/Curiosidades/Historiainspecao1.pdf>>. Acesso em: 02/04/2015.

- 35 Cleofas, J. Regulamento de Inspeção Industrial e Sanitária de produtos de Origem Animal - RIISPOA. Ministério da Agricultura, Pecuária e Abastecimento. Brasil, 1952. Disponível em: < http://www.agricultura.gov.br/arg_editor/file/Aniamal/MercadoInterno/Requisitos/RegulamentoInspecaoIndustrial.pdf >. Acesso em: 02/04/2015.
- 36 BRASIL. LEI Nº 8.078, DE 11 DE SETEMBRO DE 1990. Código de defesa do Consumidor. Código Civil. Brasil: Presidência da República, 1990.
- 37 Solomons, T. W. G.; Fryhle, C. B.; Snyder, S. A. Organic Chemistry. New York: Wiley, 2013.
- 38 Herschel, W. Experiments on the refrangibility of the invisible rays of the sun. Philosophical Transaction of the Royal Society, 90, 255-283, 1800.
- 39 Barbosa, L. C. A. Espectroscopia no Infravermelho na caracterização de compostos orgânicos. Viçosa: Editora UFV, 2007.
- 40 Skoog, D. A.; James, H. F.; Nieman, T. A. Princípios de Análise Instrumental. Porto Alegre: Bookman, 2002.
- 41 Radiação eletromagnética, 2012. Disponível em: < <http://clickeaprenda.uol.com.br/portal/mostrarConteudo.php?idPagina=28679> >
- 42 Pasquini, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. Journal of the Brazilian Chemical Society, 14, 198-219, 2003.
- 43 Stuart, B. Infrared Spectroscopy: fundamentals and applications. New York: Wiley, 2004.
- 44 Cienfuegos, F.; Vaitsman, D. Análise Instrumental. Rio de Janeiro: Interciência, 2000.
- 45 Neto, B. B.; Scarminio, I. S.; Bruns, R. E. 25 anos de Quimiometria no Brasil. Química Nova, 29, 1401-1406, 2006.
- 46 Bruns, R. E.; Faigle, J. F. G. Quimiometria. Química Nova, 8, 84-99, 1985.
- 47 Bruns, R. E. Interview. Brazilian Journal of Analytical Chemistry, 2, X-XVI, 2012.
- 48 International Chemometrics Society (ICS). Disponível em: < <http://www.namics.nysaes.cornell.edu/> >. Acesso em 15/02/2015.
- 49 Brereton, R. G. Applied Chemometrics for Scientists. England: Wiley, 2007.

- 50 Destefanis, G.; Barge, M. T.; Brugiapaglia, A.; Tassone, S. The use of principal component analysis (PCA) to characterize beef. *Meat Science*, 56, 255-259, 2000.
- 51 Bergo, M. C. J. Transferência de calibração na discriminação de mogno e espécies semelhantes utilizando NIRS e PLS-DA. (Dissertação de Mestrado). Instituto de Química, Universidade de Brasília, Brasília, 2014.
- 52 Sun, D. W. *Infrared Spectroscopy for Food Quality Analysis and Control*. Nova York: Academic Press, 2009.
- 53 Barker, M.; Rayens, W. Partial least squares for discrimination. *Journal of Chemometrics*, 17, 166-173, 2003.
- 54 Wise, B. W.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W.; Koch, R. S. *PLS-Toolbox 4.0 for use with Matlab™*: Manson: EigenVector Research Inc, 2006.
- 55 Botelho, B. G.; Reis, N.; Oliveira, L. S.; Sena, M. M. Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA. *Food Chemistry*, 181, 31-37, 2015.
- 56 Gondim, C. S.; Junqueira, R. G.; Souza, S. V. C. Trends in implementing the validation of qualitative methods of analysis: *Revista Instituto Adolfo Lutz*, 70, 433-447, 2011.
- 57 Kennard, R. W.; Stone, L. A. Computer aided design of experiments. *Technometrics*, 11, 137-148, 1969.
- 58 Brown, C. D.; Green, R. L. Critical factors limiting the interpretation of regression vectors in multivariate calibration. *TrAC-Trends in Analytical Chemistry*, 28, 506-514, 2009.
- 59 Chong, I. G.; Jun, C. H. Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, 78, 103-112, 2005.
- 60 Casale, M.; Oliveri, P.; Casolino, C.; Sinelli, N.; Zunin, P.; Armanino, C.; Forina, M.; Lanteri, S. Characterisation of PDO olive oil Chianti Classico by non-selective (UV-visible, NIR and MIR spectroscopy) and selective (fatty acid composition) analytical techniques. *Analytica Chimica Acta*, 712, 56-63, 2012.
- 61 Bro, R.; Nielsen, H. J.; Savorani, F.; Kjeldahl, K.; Christensen, I. J.; Brunner, N.; Lawaetz, A. J. Data fusion in metabolomic cancer diagnostics. *Metabolomics*, 9, 3-8, 2013.

- 62 Brady, M. Foreword. *International Journal of Robotics Research*, 7, 2-4, 1988.
- 63 Godinho, M. S.; Blanco, M. R.; Neto, F. F. G.; Liao, L. M.; Sena, M. M.; Tauler, R.; Oliveira, A. E. Evaluation of transformer insulating oil quality using NIR, fluorescence, and NMR spectroscopic data fusion. *Talanta*, 129, 143-149, 2014.
- 64 Casale, M.; Casolino, C.; Oliveri, P.; Forina, M. The potential of coupling information using three analytical techniques for identifying the geographical origin of Liguria extra virgin olive oil. *Food Chemistry*, 120, 643-643, 2010.
- 65 Pizarro, C.; Rodriguez-Tecedor, S.; Perez-del-Notario, N.; Esteban-Diez, I.; Gonzalez-Saiz, J. M. Classification of Spanish extra virgin olive oils by data fusion of visible spectroscopic fingerprints and chemical descriptors. *Food Chemistry*, 138, 915-922, 2013.
- 66 Biancolillo, A.; Bucci, R.; Magri, A. L.; Magri, A. D.; Marini, F. Data-fusion for multiplatform characterization of an italian craft beer aimed at its authentication. *Analytica Chimica Acta*, 820, 23-31, 2014.
- 67 Vera, L.; Acena, L.; Guasch, J.; Boque, R.; Mestres, M.; Busto, O. Discrimination and sensory description of beers through data fusion. *Talanta*, 87, 136-142, 2011.
- 68 Zotte, A. D.; Ottavian, M.; Concollato, A.; Serva, L.; Martelli, R.; Parisi, G. Authentication of raw and cooked freeze-dried rainbow trout (*Oncorhynchus mykiss*) by means of near infrared spectroscopy and data fusion. *Food Research International*, 60, 180-188, 2014.
- 69 Cozzi, G.; Ferlito, J.; Pasini, G.; Contiero, B.; Gottardo, F. Application of Near-Infrared Spectroscopy as an Alternative to Chemical and Color Analysis To Discriminate the Production Chains of Asiago d'Allevio Cheese. *Journal of Agricultural and Food Chemistry*, 57, 11449-11454, 2009.
- 70 Küpper, L.; Heise, H. M.; Butvina, L. N. Novel developments in mid-IR fiber-optic spectroscopy for analytical applications: *Journal of Molecular Structure*. 563-564, 173-181, 2001.
- 71 Savitzky, A.; Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36, 1627-1639, 1964.
- 72 Rinnan, A.; Van denBerg, F.; Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra: *TrAC Trends in Analytical Chemistry*, 28, 1201-1222, 2009.

- ⁷³ Cao, H.; Xu, S. Y. Purification and characterization of type II collagen from chick sternal cartilage. *Food Chemistry*, 108, 439-445, 2008.
- ⁷⁴ Volery, P.; Besson, R.; Schaffer-Lequart, C. Characterization of commercial carrageenans by Fourier transform infrared spectroscopy using single-reflection attenuated total reflection. *Journal of Agricultural and Food Chemistry*, 52, 7457-7463, 2004.
- ⁷⁵ Smrckova, P. et al. Hydrolysis of wheat B-starch and characterisation of acetylated maltodextrin. *Carbohydrate Polymers*, 98, 43-49, 2013.
- ⁷⁶ Khanmohammadi, M.; Ashori, A.; Kargosha, K.; Garmarudi, A. B. Simultaneous determination of sodium tripolyphosphate, sodium sulfate and linear alkylbenzensulfonate in washing powder by attenuated total reflectance: Fourier transform infrared spectrometry. *Journal of Surfactants and Detergents*, 10, 81-86, 2007.

ANEXO A

Dados Físico-Químicos das amostras reais.

Amostra	Proteína total (g/100g)	Cinzas (g/100g)	Sódio (mg/100g)	Cloreto (mg/100g)	Fosfato (mg/100g)
1-S	17,10	1,50	328,00	176,00	146,00
2-S	18,40	1,50	328,00	176,00	146,00
3-S	16,50	1,50	328,00	176,00	146,00
4-S	15,60	1,50	328,00	176,00	146,00
5-S	17,60	1,60	309,70	151,00	117,40
6-S	17,00	1,60	309,70	151,00	117,40
7-S	16,70	1,60	309,70	151,00	117,40
8-S	16,90	1,60	309,70	151,00	117,40
9-S	17,60	1,60	296,80	153,50	125,70
10-S	14,80	1,60	296,80	153,50	125,70
11-S	18,00	1,60	296,80	153,50	125,70
12-S	17,60	1,60	296,80	153,50	125,70
13-C	11,70	1,20	101,00	40,00	69,00
14-C	21,40	1,05	82,00	17,00	94,00
15-C	8,50	1,15	89,00	54,00	86,00
16-C	18,90	1,03	106,00	34,00	72,00
17-C	21,30	1,10	87,00	2,20	84,00
18-C	10,20	1,00	56,00	32,00	59,00
19-C	21,70	1,10	102,00	18,00	63,00
20-C	11,30	1,10	95,00	53,00	103,00
21-C	21,20	1,07	85,00	3,00	100,00
22-C	9,50	0,82	80,00	33,00	71,00
23-C	10,10	1,20	171,00	40,00	111,00
24-C	21,40	1,10	112,00	7,00	80,00
25-S	5,20	1,79	412,00	187,00	251,00
26-S	4,50	1,79	412,00	187,00	251,00
27-S	17,50	1,52	286,00	106,00	104,00
28-S	16,90	1,52	286,00	106,00	104,00
29-S	17,40	1,52	286,00	106,00	104,00
30-S	6,50	1,79	412,00	187,00	251,00
31-S	16,70	1,52	286,00	106,00	104,00
32-S	21,30	1,04	89,00	12,00	97,00
33-S	7,90	1,20	112,00	77,00	96,00
34-S	21,00	1,04	89,00	12,00	97,00
35-S	21,80	1,04	89,00	12,00	97,00
36-S	9,60	1,20	112,00	77,00	96,00
37-S	8,30	1,20	112,00	77,00	96,00
38-S	8,30	1,20	112,00	77,00	96,00
39-S	8,30	1,30	166,00	150,00	114,00
40-S	4,80	1,30	246,00	187,00	241,00
41-S	15,20	1,30	228,00	116,00	131,00

42-S	16,60	1,40	229,00	119,00	124,00
43-S	14,80	1,10	178,00	112,00	87,00
44-S	17,10	1,20	171,00	97,00	92,00
45-S	6,80	1,10	172,00	160,00	156,00
46-S	7,10	1,20	192,00	153,00	147,00
47-S	20,90	1,04	94,00	8,00	57,00
48-S	8,30	0,91	67,00	32,00	94,00
49-S	21,60	1,10	74,00	6,00	93,00
50-S	9,80	1,33	97,00	32,00	50,00
51-S	5,90	1,60	388,00	143,00	175,00
52-S	6,30	1,50	381,00	144,00	164,00
53-S	7,20	1,70	384,00	149,00	178,00
54-S	19,00	1,20	191,00	71,00	85,00
55-S	16,60	1,40	279,00	95,00	101,00

ANEXO B

Dados do planejamento para a preparação das amostras simuladas.

AMOSTRA	DENTIÇÃO	MASSA INICIAL (g)	TRATAMENTO	MASSA FINAL (g)	PESO (%)
1	0	2950,00	controle-C	2950,00	100,00
2	2	2895,00	Mistura A (30%)	3700,00	127,81
3	2	2230,00	Mistura C (10%)	2850,00	127,80
4	2	2850,00	Mistura B (30%)	3885,00	136,32
5	2	3170,00	Mistura D (30%)	4720,00	148,90
6	0	3340,00	Mistura A (10%)	3935,00	117,81
7	2	2930,00	Mistura A (30%)	3830,00	130,72
8	2	3040,00	Mistura A (30%)	3840,00	126,32
9	2	3465,00	Mistura C (10%)	4385,00	126,55
10	2	3225,00	Mistura D (30%)	4970,00	154,11
11	2	2940,00	Mistura D (30%)	4430,00	150,68
12	0	2930,00	Mistura D (10%)	3795,00	129,52
13	0	3640,00	Mistura A (30%)	3955,00	108,65
14	0	2975,00	Mistura B (10%)	3760,00	126,39
15	2	3540,00	Mistura A (30%)	4360,00	123,16
16	0	3040,00	Mistura D (10%)	3950,00	129,93
17	2	3560,00	Mistura B (10%)	4485,00	125,98
18	2	3610,00	Mistura B (30%)	4565,00	126,45
19	2	3605,00	Mistura C (30%)	4875,00	135,23
20	0	2855,00	Mistura A (10%)	3290,00	115,24
21	2	2655,00	Mistura B (10%)	3400,00	128,06
22	2	2915,00	Mistura B (10%)	3580,00	122,81
23	2	3095,00	Mistura D (10%)	4005,00	129,40
24	2	3305,00	Mistura B (30%)	4610,00	139,49
25	2	3090,00	Mistura D (10%)	4205,00	136,08
26	2	2970,00	Mistura A (10%)	3490,00	117,51
27	2	3750,00	Mistura B (10%)	4605,00	122,80
28	2	3250,00	Mistura C (10%)	4135,00	127,23
29	2	3425,00	Mistura D (10%)	4315,00	125,99
30	2	4245,00	Mistura C (30%)	5560,00	130,98
31	2	3160,00	Mistura C (10%)	3900,00	123,42
32	2	3085,00	Mistura D (30%)	4615,00	149,59
33	2	3445,00	controle-C	3445,00	100,00
34	2	3030,00	Mistura D (10%)	3940,00	130,03
35	2	3395,00	Mistura A (30%)	4300,00	126,66
36	0	2765,00	Mistura C (30%)	4310,00	155,88
37	2	2800,00	Mistura A (10%)	3245,00	115,89
38	2	3350,00	Mistura D (30%)	4790,00	142,99
39	2	2885,00	Mistura D (30%)	4270,00	148,01
40	0	2635,00	Mistura C (30%)	4060,00	154,08
41	0	2960,00	Mistura B (10%)	3780,00	127,70

42	2	2785,00	Mistura A (10%)	3180,00	114,18
43	2	3130,00	Mistura C (10%)	3930,00	125,56
44	2	3315,00	Mistura B (30%)	4660,00	140,57
45	2	3095,00	Mistura B (30%)	4410,00	142,49
46	2	3310,00	Mistura C (30%)	4745,00	143,35
47	2	3150,00	Mistura B (30%)	4455,00	141,43
48	2	4145,00	Mistura C (10%)	5185,00	125,09
49	2	2940,00	Mistura A (10%)	3505,00	119,22
50	2	3380,00	controle-C	3380,00	100,00
51	2	4495,00	Mistura C (30%)	6135,00	136,48
Identificação das misturas utilizadas na simulação das amostras					
Mistura A:	água.				
Mistura B:	tripolifosfato de sódio.				
Mistura C:	tipolifosfato de sódio, maltodextrina, sal refinado iodado, pirofosfato ácido de sódio e carragena.				
Mistura D:	tipolifosfato de sódio, maltodextrina, sal refinado iodado, pirofosfato ácido de sódio, carragena e colágeno.				
OBS: Bovinos com até dois dentes incisivos possuem até 24 meses.					