



UNIVERSIDADE ESTADUAL DO SUDOESTE DA BAHIA – UESB
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA E CIÊNCIA DE
ALIMENTOS

DETECÇÃO DA ADIÇÃO DO LEITE DE VACA AO LEITE DE BÚFALA POR
MEIO DA TÉCNICA ESPECTROSCÓPICA FTIR-ATR E QUIMIOMETRIA

LARISSA KAULY ROSA SILVA

ITAPETINGA
BAHIA - BRASIL
2017

UNIVERSIDADE ESTADUAL DO SUDOESTE DA BAHIA – UESB
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA E CIÊNCIA DE
ALIMENTOS

DETECÇÃO DA ADIÇÃO DO LEITE DE VACA AO LEITE DE BÚFALA POR
MEIO DA TÉCNICA ESPECTROSCÓPICA FTIR-ATR E QUIMIOMETRIA

LARISSA KAULY ROSA SILVA

Dissertação apresentada à Universidade Estadual do Sudoeste da Bahia, como parte das exigências do Programa de Pós-Graduação em Engenharia e Ciência de Alimentos, Área de Concentração em Ciência de Alimentos, Linha de Pesquisa em Tecnologia de Alimentos, para obtenção do título de "Mestre".

Orientador (a): DSc. Sibelli Passini Barbosa Ferrão

Co-orientador: DSc. Leandro Soares Santos

ITAPETINGA
BAHIA - BRASIL
2017

637.127 Silva, Larissa Kauly Rosa

S581d Detecção da adição do leite de vaca ao leite de búfala por meio da técnica espectroscópica FTIR-ATR e quimiometria. / Larissa Kauly Rosa Silva
. - Itapetinga: UESB, 2017.
88p.

Dissertação apresentada à Universidade Estadual do Sudoeste da Bahia, como parte das exigências do Programa de Pós-Graduação em Engenharia e Ciência de Alimentos, Área de Concentração em Ciência de Alimentos, Linha de Pesquisa em Tecnologia de Alimentos, para obtenção do título de "Mestre". Sob a orientação da Prof^a. D.Sc. Sibelli Passini Barbosa Ferrão e coorientação do Prof. D.Sc. Leandro Soares Santos.

1. Leite de búfala – Fraude - Multivariada. 2. Adulteração em leite de búfala - Espectroscopia. 3. Produtos lácteos - Controle de qualidade e segurança. I. Universidade Estadual do Sudoeste da Bahia. Programa de Pós-Graduação em Engenharia e Ciência de Alimentos. II. Ferrão, Sibelli Passini Barbosa. III. Santos, Leandro Soares. IV. Título.

CDD(21): 637.127

Catálogo na fonte:

Adalice Gustavo da Silva – CRB/5-535
Bibliotecária – UESB – Campus de Itapetinga-BA

Índice Sistemático para Desdobramento por Assunto:

1. Leite de búfala – Fraude - Multivariada.
2. Adulteração em leite de búfala - Espectroscopia
3. Produtos lácteos - Controle de qualidade e segurança



Universidade Estadual do Sudoeste da Bahia
Programa de Pós-Graduação
Mestrado em Engenharia e Ciência de Alimentos



Áreas de Concentração: Engenharia de Alimentos
Ciência de Alimentos

DECLARAÇÃO DE APROVAÇÃO

Título: DETECÇÃO DA ADIÇÃO DO LEITE DE VACA AO LEITE DE BÚFALA POR MEIO DA TÉCNICA ESPECTROSCÓPICA FTIR-ATR E QUIMIOMETRIA.

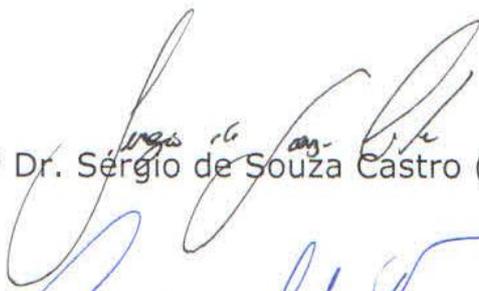
Autor (a): LARISSA KAULY ROSA SILVA

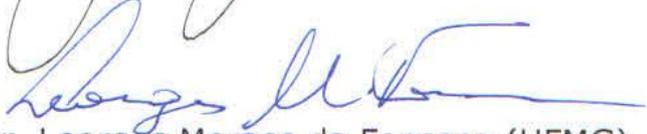
Orientador (a): Prof.^a Dr.^a Sibelli Passini Barbosa Ferrão

Coorientador (a): Prof.^o Dr. Leandro Soares Santos

Aprovado como parte das exigências para obtenção do Título de **MESTRE EM ENGENHARIA E CIÊNCIA DE ALIMENTOS, ÁREA DE CONCENTRAÇÃO: CIÊNCIA DE ALIMENTOS**, pela Banca Examinadora.


Prof.^a Dr.^a Sibelli Passini Barbosa Ferrão (UESB)


Prof.^o Dr. Sérgio de Souza Castro (UESB)


Prof.^o Dr. Leorges Moraes da Fonseca (UFMG)

Itapetinga-BA, 22 de fevereiro de 2017.

*“A persistência é o menor
caminho para o êxito”.*

Charles Chaplin

À Deus
À minha mãe
À minha família

Dedico!

AGRADECIMENTOS

Agradeço primeiramente a Deus, pela vida, pela saúde, pela proteção diária, pela concessão de valiosas oportunidades e por me conferir sabedoria para que eu possa conquistar meus sonhos e objetivos.

À minha mãe maravilhosa, pelo amor, pelo carinho, por sempre se fazer presente, mesmo distante e por ser a minha base, a minha inspiração e o meu maior incentivo na realização dos meus sonhos.

À minha família, meu presente, minha alegria, que sempre esteve ao meu lado, me apoiando, acreditando nas minhas aspirações e me mostrando que sou capaz de alcançá-las. Em especial, à Tia Camila, Tio Eri, Henrique, Alice e Maria, um agradecimento mais que especial. Vocês me proporcionaram essa conquista e conseguiram torná-la ainda mais brilhante, único...Tia Mil, sem palavras para agradecer a mãe e amiga que você é pra mim!

Ao meu par, meu amor, “meu filho” Ramon, pelo amor, companheirismo e por sempre estar me apoiando, me incentivando e me tranquilizando nos momentos mais difíceis. Você foi primordial para que eu chegasse até aqui. Te amo!

À Universidade Estadual do Sudoeste da Bahia (UESB), pela estrutura de laboratórios e pelos funcionários que possibilitaram o desenvolvimento deste trabalho.

À minha querida orientadora, professora Dr^a. Sibelli Ferrão, pela orientação e contribuição intelectual com tanta dedicação e disponibilidade e, principalmente, por ter acreditado e confiado em mim. Com você aprendi a ter foco e tranquilidade para realizar o que nos propomos a fazer.

Ao meu co-orientador, professor Dr. Leandro Soares pela disponibilidade e paciência, e pela valiosa contribuição, ampliando meus conhecimentos e enriquecendo esta pesquisa.

Aos membros da banca examinadora por aceitarem fazer parte da avaliação deste trabalho.

Ao Grupo de Estudos em Leite (GEL).

Aos colegas de Mestrado, em especial as minhas amigas nutris Jéssica e Maria José, grandes amigas que levarei para minha vida.

À Fapesb pela concessão da bolsa.

À todos aqueles que de alguma maneira contribuíram para a realização deste trabalho.

Meus sinceros agradecimentos!

LISTA DE TABELAS

Tabela 1: Atribuições da absorção dos compostos orgânicos das amostras de leite.....	42
Tabela 2: Arquitetura das redes neurais artificiais testadas.....	49
Tabela 3: Número de observações e o percentual de classificação por grupo no treinamento (T) e validação (V).....	66
Tabela 4: Configuração das Redes Neurais Artificiais testadas, Erro Quadrático Médio (RMSE), e taxas de classificação (%)......	67
Tabela 5: Configurações das redes neurais testadas para a predição e os coeficientes preditivos.....	70

LISTA DE FIGURAS

Figura 1: Esquema ilustrativo do espectrômetro Infravermelho com Transformada de Fourier (FTIR).....	27
Figura 2: Funcionamento do neurônio artificial.....	38
Figura 3: Espectros obtidos por espectrômetro FTIR-ATR das amostras do tratamento referência búfala (TRB) e tratamento referência vaca (TRV).....	52
Figura 4: Espectros obtidos por espectrômetro FTIR-ATR das amostras do tratamento referência búfala (TRB), tratamento referência vaca (TRV) e das amostras e das amostras de leite de búfala adicionadas de leite de vaca (T10 ao T90).....	54
Figura 5: <i>Scree Plot</i> dos autovalores associados aos componentes principais.....	57
Figura 6: Gráfico de dispersão das amostras do tratamento referência búfala (TRB) e do tratamento referência vaca (TRV) em relação aos componentes principais CP1, CP2 e CP3..	58
Figura 7: Gráficos de dispersão das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV) e das amostras de leite de búfala adicionadas de leite de vaca (T10 ao T90) em relação aos componentes principais CP1, CP2 e CP3.....	60
Figura 8: Gráfico de médias dos escores das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV) e das amostras de leite de búfala adicionadas de leite de vaca (T10 ao T90) em relação ao CP2 e CP3.....	61
Figura 9: Dendograma das amostras do tratamento referência búfala (TRB – B), do tratamento referência vaca (TRV – V).....	63
Figura 10: Gráfico do erro quadrático médio (RMSE) em função do número de iterações nas etapas de treinamento e validação da rede neural de melhor arquitetura.....	69
Figura 11. Valores reais versus valores preditos pela rede das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV), e das amostras de leite de búfala adicionadas de leite de vaca (T10 ao T90).....	71

LISTA DE SÍMBOLOS

%	Porcentagem
>	Maior
<	Menor
γ	Gama
°C	Graus Celsius
®	Marca registrada
L/min	Litros por minuto
g	Gramas
mL	Mililitro
Z	Nova variável
σ	Desvio padrão
μ	Média
λ	Lambda
I	Matriz Identidade
S	Matriz de Variância e Covariância
a	Autovalor
β_0	Constante da função linear discriminante
β_1	Constante das variáveis da função linear discriminante
C	Carbono
H	Hidrogênio
O	Oxigênio
N	Nitrogênio

LISTA DE SIGLAS

UESB	Universidade Estadual do Sudoeste da Bahia
BA	Bahia
CLAE	Cromatografia Líquida de Alta Eficiência
RIISPOA	Regulamento de Inspeção Industrial e Sanitária de Produtos de Origem Animal
TRB	Tratamento Referência Búfala
TRV	Tratamento Referência Vaca
ACP	Análise de Componentes Principais
CP	Componentes Principais
AD	Análise Discriminante
RNA	Redes Neurais Artificiais
FTIR	Infravermelho com Transformada de Fourier
ATR	Reflexão Total Atenuada
AU	Unidades de absorvância
SAS	Statistical Analysis System
JavaNNS	Java Neural Network Simulator
MLP	Multilayers perceptrons
RMSE	Raiz do erro quadrado médio
EST	Extrato Seco Total

SUMÁRIO

RESUMO	14
ABSTRACT.....	15
1. Introdução.....	16
2. Revisão de Literatura.....	18
2.1. Bubalinocultura: aspectos gerais e produção leiteira	18
2.2. Características do leite de búfala x vaca	18
2.3. Adulteração em leite de búfala.....	20
2.4. Métodos analíticos para detecção de adulteração.....	22
2.5. Espectroscopia no Infravermelho	24
2.6. Infravermelho com Transformada da Fourier - FTIR	26
2.7. Utilização do FTIR na detecção de adulteração em leite e derivados	28
2.8. Quimiometria	30
2.8.1. Análise de Componentes Principais (ACP).....	32
2.8.2. Análise de Agrupamento (AA) ou Cluster Analysis	34
2.8.3. Análise Discriminante (AD).....	35
2.8.4. Redes Neurais Artificiais (RNA).....	37
3. Objetivos	40
3.1. Objetivo Geral.....	40
3.2. Objetivos Específicos.....	40
4. Material e Métodos.....	41
4.1. Coleta de dados e local do experimento.....	41
4.2. Preparo das amostras de leite de vaca e búfala.....	41
4.3. Aquisição dos espectros	41
4.4. Tratamento Prévio dos Dados	43
4.5. Análise estatística multivariada dos dados	43
4.5.1. Análise de Componentes Principais (ACP).....	44

4.5.2.	Análise de Agrupamento (AA) ou Cluster Analysis	45
4.5.3.	Análise Discriminante (AD)	46
4.5.4.	Redes Neurais Artificiais (RNA)	47
4.5.4.1.	Desenvolvimento e arquitetura da RNA	48
4.5.4.2.	Treinamento e Validação	49
4.5.4.3.	Classificação e Predição	50
5.	Resultados e Discussão.....	52
5.1.	Espectros das amostras referência e das amostras de leite de búfala com adição de leite de vaca.....	52
5.2.	Análise Multivariada dos dados	56
5.2.1.	Análise de Componentes Principais (ACP).....	56
5.2.2.	Análise de Agrupamento (AA) ou <i>Cluster Analysis</i>	62
5.2.3.	Análise Discriminante (AD).....	64
5.2.4.	Redes Neurais Artificiais (RNA)	67
5.2.4.1.	Classificação.....	67
5.2.4.2.	Predição.....	70
6.	Conclusões.....	73
7.	Referências Bibliográficas	74

RESUMO

SILVA, L.K.R. **Detecção da adição do leite vaca ao leite de búfala por meio da técnica espectroscópica FTIR-ATR e quimiometria.** Itapetinga – BA: UESB, 2017. 88 p. (Dissertação – Mestrado em Engenharia e Ciência de Alimentos).*

A detecção de adulteração tornou-se uma questão importante no controle de qualidade e segurança de produtos lácteos, especialmente do leite de búfala, pois a sazonalidade que afeta a sua disponibilidade e o seu preço comparado ao leite de vaca o tornam alvo para possíveis adulterações. Objetivou-se com este estudo detectar a adição de leite de vaca ao leite de búfala por meio da técnica espectroscópica FTIR-ATR e ferramentas quimiométricas. Foram coletadas amostras de leite búfala e vaca por 45 dias sendo preparadas 9 combinações com crescentes níveis de adição de leite de vaca ao leite de búfala (T10 ao T90), e 2 amostras com 100% de leite de búfala (TRB) e vaca (TRV), totalizando 495 amostras. Foram adquiridos 15 picos na região espectral entre 4000 cm^{-1} a 600 cm^{-1} por meio das análises no FTIR-ATR e os valores de absorvância de cada pico gerado foram submetidos à Análise de Componentes Principais (ACP), Análise Discriminante (AD), Análise de Agrupamento (AA) e às Redes Neurais Artificiais (RNA). Foi utilizado SAS® *Studio* e o Java *Neural Network Simulator*. Os resultados revelaram que por meio da ACP, o TRB apresentou comportamento diferente do TRV e possibilitou a separação de amostras com adição de leite de vaca ao leite de búfala apenas a partir de 40%, sendo as variáveis referentes à gordura importantes na separação das amostras. O método de AA apresentou agrupamento satisfatório apenas para as amostras do TRB e TRV. A AD apresentou baixas taxas de classificação no treinamento e validação tanto para o TRB e TRV quanto para amostras com níveis de adição. As RNA se apresentaram mais eficazes na classificação das amostras em relação às demais análises multivariadas. Foram obtidas taxas de classificação de 100% para as amostras TRB, TRV e para a maioria dos grupos de amostras de leite de búfala com adição de leite de vaca e 95,55% para a classificação total, além da boa capacidade de predição com correlações e erros satisfatórios.

Palavras-chave: espectroscopia, fraude, multivariada, produtos lácteos.

* Orientador (a): Sibelli Passini Barbosa Ferrão, DSc., UESB. Co-orientador: Leandro Soares Santos, DSc., UESB.

ABSTRACT

SILVA, L.K.R. **Detection of the addition of cow's milk to buffalo milk by the FTIR-ATR spectroscopic technique and chemometrics.** Itapetinga – BA: UESB, 2017. 88 p. (Dissertation - Master of Engineering and Food Science). *

Detection of tampering has become an important issue in the quality and safety control of dairy products, especially buffalo milk, since the seasonality that affects their availability and price compared to cow's milk makes it a target for possible adulteration. The objective of this study was to detect the addition of cow's milk to buffalo milk by means of the spectroscopic technique FTIR-ATR and chemometric tools. Samples of buffalo and cow milk were collected for 45 days, and 9 combinations were prepared with increasing levels of cow's milk addition to buffalo milk (T10 to T90), and 2 samples with 100% buffalo milk (BER) and cow (TRV), totaling 495 samples. A total of 15 peaks were obtained in the spectral region between 4000 cm^{-1} and 600 cm^{-1} by FTIR-ATR analyzes and the absorbance values of each peak were subjected to Principal Component Analysis (ACP), Discriminant Analysis (DA), Cluster Analysis (CA) and Artificial Neural Networks (ANN). SAS® Studio and the Java Neural Network Simulator were used. The results showed that, through the PCA, the TRB presented a different behavior from the TRV and allowed the separation of samples with cow's milk added to buffalo milk from only 40%, with the fat variables being important in separating the samples. The AA method presented satisfactory grouping only for the TRB and TRV samples. The AD presented low classification rates in training and validation for both TRB and TRV as well as samples with addition levels. The ANNs were more effective in the classification of the samples in relation to the other multivariate analyzes. 100% classification rates were obtained for the TRB, TRV and most of the buffalo milk samples with cow's milk addition and 95.55% for the total classification, in addition to the good prediction ability with correlations and errors.

Keywords: Spectroscopy, fraud, multivariate, dairy products..

*Advisor (a): Sibelli Passini Barbosa Ferrão, DSc., UESB. Co-advisor: Leandro Soares Santos, DSc., UESB.

1. INTRODUÇÃO

A garantia da autenticidade de produtos lácteos tornou-se um problema mundial devido ao surgimento de inúmeros episódios de adulteração, dentre os quais, a mistura de leite de espécies diferentes tem sido recorrente. Por razões econômicas e de saúde pública, torna-se cada vez mais importante a detecção de produtos adulterados no mercado. As flutuações sazonais na sua disponibilidade e o preço mais elevado em relação ao leite de vaca tornam o leite de búfala alvo de adulteração por parte de alguns produtores e laticínios que adicionam leite de maior disponibilidade e menor preço ao leite mais nobre para suprir esta demanda (RODRIGUES et al., 2012; FILHO et al., 2014).

Quando comparado com o leite de outras espécies, o leite de búfala confere aos derivados lácteos propriedades sensoriais peculiares, permitindo a elaboração de produtos com características tecnológicas próprias e maior rendimento industrial, em razão de seu maior nível de extrato seco total (EST) e dos minerais. Embora apresente maior valor comercial em relação ao leite de vaca, e sua produção venha se difundindo no país, pouco se tem feito para regulamentação do padrão de autenticidade e qualidade do leite bubalino e seus derivados lácteos, o que tem dificultado medidas de controle e fiscalização.

A determinação da autenticidade do leite bubalino entregue ao laticínio é de grande importância tornando-se imprescindível o desenvolvimento de métodos analíticos sensíveis, rápidos e eficientes para a detecção da adição de leite de outra espécie ao leite de búfala, que garantam a autenticidade do produto, possibilitando o seu controle e, também, proteção ao consumidor.

O Infravermelho com Transformada de Fourier associado à Reflexão Total Atenuada (FTIR-ATR) é uma tecnologia atraente para a indústria de alimentos, pois medições simples, rápidas e não destrutivas de componentes físicos e químicos podem ser obtidas. Os avanços na análise dos dados por meio do desenvolvimento do método de análise multivariada tornam esta tecnologia oportuna para a triagem rápida de grandes volumes de amostras, ideal para o monitoramento de adulterações em alimentos. (CASSOLI et al., 2011; PAVIA, 2010).

Vale ressaltar que, apesar da importância do leite de búfala para o mercado, a

legislação brasileira ainda não possui um método oficial validado para detecção de adulteração sendo a autenticidade do leite e de produtos lácteos sempre determinada por métodos analíticos mais tradicionais, geralmente laboriosos e de alto custo.

Neste sentido, a aplicação das ferramentas quimiométricas junto à técnica de espectroscopia no infravermelho, por meio da extração e interpretação de informações de conjuntos de dados complexos, vem apresentando uma grande vantagem tanto para a quantificação dos principais componentes do leite quanto para a detecção de adulteração de produtos lácteos. Diante da necessidade de combater práticas fraudulentas na produção e comercialização de leite, este trabalho justifica-se na possibilidade de auxiliar com resultados detalhados na discriminação de produtos adulterados visando a garantia de autenticidade.

2. REVISÃO DE LITERATURA

2.1. Bubalinocultura: aspectos gerais e produção leiteira

A criação de búfalos é originária da Ásia e tem apresentado crescimento mundial de 3,29% ao ano. A bubalinocultura brasileira tem se desenvolvido nas últimas décadas, o que é evidenciado pelo crescimento populacional desses animais de forma significativa em diversas regiões (FAO, 2014).

Segundo o Ministério da Agricultura, Pecuária e Abastecimento, o rebanho brasileiro bubalino está estimado em torno de 1,18 milhão de bubalinos, sendo a região Nordeste, com cerca de 127 mil animais, a terceira maior produtora do País no ranking da criação de bubalinos, com destaque para o Maranhão e a Bahia, que possuem cerca de 84 mil e 26 mil bubalinos. O Norte ocupa primeiro lugar, com 720 mil cabeças (BRASIL, 2015).

A produção brasileira de leite de búfala destinada à industrialização vem crescendo em média 25% ao ano desde 2001. Estima-se que a produção leiteira de búfala no Brasil seja superior a 92 milhões de litros por ano (IBGE, 2015).

A sazonalidade reprodutiva da espécie é um problema para a produção leiteira, pois interfere negativamente na oferta dos produtos lácteos à indústria, havendo uma queda de, aproximadamente, 60% da produção no período da entressafra. Portanto, enquanto a demanda por derivados se mantém, a disponibilidade desse produto diminui pela escassez de matéria-prima. Devido a essas flutuações sazonais para disponibilidade do leite de búfala, aliado ao preço mais elevado comparado com o leite de vaca, a adição de leite de outras espécies no momento da chegada do leite ao laticínio, bem como, durante o preparo de produtos lácteos de búfala torna-se uma fraude frequente. Por isso, atualmente, são aplicados diversos métodos laboratoriais a fim de se detectar tal adulteração (FILHO et al., 2014).

2.2. Características do leite de búfala x vaca

Entende-se por leite o produto oriundo da ordenha completa, ininterrupta, em condições de higiene, de vacas sadias, bem alimentadas e descansadas. O leite de outros animais deve denominar-se segundo a espécie de que procede (BRASIL, 2011).

O leite de búfala apresenta características próprias que o diferencia sob o ponto

de vista físico-químico e organoléptico de outras espécies, podendo variar conforme o estágio de lactação, raça, alimentação, entre outros fatores. Teores de lipídeos, proteínas, lactose, sólidos totais e sólidos desengordurados são superiores aos do leite de vaca, ocasionando maior rendimento industrial além de conferir grande importância nutricional (VENTURINI et al., 2007; SINDHU e ARORA, 2011).

Segundo dados regionais, o leite de búfala apresenta maior teor dos principais constituintes: em média 4,26-5,8% de gordura, 3,05-4,0% de proteína, 4,27% de lactose (CALDEIRA et al., 2010; PIGNATA et al., 2014). Em contrapartida, o leite de vaca apresenta em média 3,8% de gordura, 2,92% de proteínas e 4,14% de lactose (PIGNATA et al., 2014).

A vitamina A é a que apresenta maior destaque apesar de apresentar menor teor em relação ao leite de vaca, conferindo uma das características mais marcantes ao leite de búfala, que é uma coloração acentuadamente branca e opaca devido à ausência dos pigmentos carotenóides (provitamina A), responsável pela cor amarelada dos produtos derivados do leite bovino (AMARAL et al., 2005).

Em relação ao teor de minerais, o leite é mais rico em cálcio, fósforo e magnésio do que o leite de vaca, porém é mais pobre em potássio e cloro. A relação entre os teores de cálcio e fósforo é 1,71, enquanto que no leite de vaca é de 1,31 (MACEDO et al., 2001; AMARAL et al., 2005; TEIXEIRA et al., 2005).

O leite de búfala é importante tanto para o consumo *in natura* como matéria-prima para elaboração de produtos lácteos e sua grande importância está na transformação em derivados, uma vez que seu elevado teor em extrato seco, incluindo gordura e proteína, possibilita um alto rendimento industrial (RAMOS et al., 2003).

A gordura e as proteínas são os constituintes do leite que apresentam maiores valores econômicos, sendo importantes em programas de pagamento por qualidade na indústria e na produção de alguns derivados lácteos. O teor de gordura é um parâmetro que contribui para desenvolvimento do sabor característico e melhoria da textura, além de ser o componente mais variável do leite, sendo facilmente modificado de acordo com a alimentação, sanidade, idade e raça, clima, entre outros (VENTURINI et al., 2007).

No leite de búfala o teor de gordura geralmente está acima de 5,5%, podendo ser encontrado no Brasil valores de até 10,4% (FARIA et al., 2002; FERNANDES et al., 2005; RANGEL et al., 2011). Entretanto, a concentração total do colesterol do leite de búfala é menor do que a encontrada no leite de vaca (275 mg/100 g contra 330 mg/100

g de gordura), além de apresentar níveis apreciáveis de ácidos graxos essenciais ao organismo, possuindo uma proporção de ácido graxo saturado/ácido graxo poliinsaturado nutricionalmente aceitáveis, enquanto que para o leite de vaca a cadeia de ácidos graxos possui, cerca de 56 % dos ácidos graxos dos saturados e apenas 4% de ácidos poliinsaturados (AMARAL et al., 2005; ARAÚJO, 2011).

O leite bubalino pode ser considerado bastante rico em proteína, com valor médio de 4% (MATTOS, 2007). Dentre as principais proteínas encontradas no leite de ambas as espécies, podem ser citadas as caseínas, albuminas, globulinas e outras, em quantidades menores. Da porção proteica do leite de búfalas, a caseína equivale a 77-79% das proteínas (α_{s1} -caseína, α_{s2} -caseína, β -caseína e κ -caseína), sendo que cerca de 21-23% são de proteínas solúveis do soro (β -lactoglobulina, α -lactoalbumina e albumina). A albumina ou lactoalbumina, conhecida como proteína do soro, possui grande qualidade e valor econômico expressivo na indústria farmacêutica. Além disso, o leite também contém proteínas pouco abundantes, como a lactoferrina, imunoglobulinas, glicoproteínas e enzimas (RICCI e DOMINGUES, 2012; RONCADA et al., 2012).

A utilização do leite de búfala para a elaboração de produtos que, tradicionalmente, são derivados do leite de vaca se justifica quando se compara a qualidade nutricional e tecnológica destes. O leite de búfala em relação ao de vaca, apresenta-se como um produto nobre, sendo, portanto, de elevada qualidade e maior rendimento que os produtos de leite bovino (FARIA et al., 2006).

Diante dessas características, o leite de búfala é uma excelente matéria-prima que pode fornecer ampla variedade de produtos, tais como, queijos frescos e maturados, sorvetes, iogurtes e creme de leite, dentre outros derivados e sua composição permite um rendimento de cerca de 40% superior ao leite bovino na elaboração de alguns derivados (MÉNARD et al., 2010; ANDRIGHETTO, 2011).

2.3. Adulteração em leite de búfala

São considerados matérias-primas ou produtos fraudados aqueles que apresentem modificações propositais de natureza física, química ou biológica que alterem suas características sensoriais ou a sua composição intrínseca, comprometendo seu valor nutritivo. Assim, o produto pode ser considerado fraudado quando for

empregada substância de qualquer qualidade, tipo ou espécie diferente das expressas na formulação original. Dentre os tipos de fraude, tem-se a fraude por adulteração, onde modificações intencionais com a finalidade de lucro afetam as características sensoriais e o valor nutritivo do alimento (BRASIL, 2007).

O Regulamento da Inspeção Industrial e Sanitária de Produtos de Origem Animal (RIISPOA), em seu artigo 542 para leite e derivados, considera impróprio para consumo humano todo leite beneficiado que, entre outras características, esteja fraudado. Em seu artigo 543, é considerado fraudado, adulterado ou falsificado o leite adicionado de quaisquer substâncias de qualidade, tipo e espécies diferentes da composição normal do produto (BRASIL, 1997).

Os tipos mais comuns de adulterações e fraudes do leite incluem adição de água, soro, reconstituintes, neutralizantes, conservantes, leite de espécies diferentes, gorduras não lácteas, alteração da composição e troca de data de validade (DIAS et al., 2009).

A garantia da autenticidade de produtos alimentícios, como aqueles contendo leite e/ou proteínas lácteas, tornou-se um problema global. Assim, faz-se necessário detectar a comercialização de produtos rotulados de maneira fraudulenta e de qualidade inferior no mercado, quer por razões econômicas, quer por razões de saúde pública (VELOSO et al., 2002).

As flutuações na disponibilidade de leite de espécies diferentes da bovina, como de cabra, ovelha e búfala, bem como o preço mais elevado oferecido a estas, levam produtores e indústrias de produtos lácteos a adulterarem o leite e seus derivados tradicionais. Sobretudo, a adição fraudulenta de leite bovino em leite de búfala para produção de derivados que deveriam conter apenas leite de búfala, prejudica a qualidade final do produto além de gerar uma perda nutricional e econômica para o consumidor (EGITO et al., 2006).

A composição química e a produção do leite de búfala são influenciadas por diversos fatores, tais como raça, estágio de lactação, idade, manejo, sanidade, condições climáticas, alimentação e período do ano. Geralmente, a búfala produz cerca de 82% de leite no outono-inverno e apresenta escassez de matéria-prima durante o verão (CZERWENKA et al., 2010; ANDRIGHETTO, 2011). A escassez, o alto valor agregado, a maior oferta do leite de vaca, aliados à dificuldade da identificação desse tipo de fraude devido às semelhanças entre esses produtos são alguns dos possíveis motivos que levam a fraude (RODRIGUES et al., 2012).

Existem legislações vigentes que determinam a obrigatoriedade de informação dos componentes presentes nos rótulos das embalagens de produtos de origem animal e lácteos. A rotulagem deve estar de acordo com a elaboração do produto, sua origem, peso, marca e denominação de venda, sendo obrigatória a lista de ingredientes, a validade e o lote. Portanto, as leis em vigor amparam o consumidor de possíveis alterações provocadas pelo fabricante na rotulagem, porém, ainda é escasso o conhecimento do mesmo sobre a real formulação de produtos e suas possibilidades de adulteração (DIAS, 2009).

A ocorrência de adulterações de qualquer natureza implica, além de abuso contra a economia do consumidor, em danos significativos à saúde pública, contrariando o Código de Defesa do Consumidor e a legislação brasileira vigente (ALMEIDA, 2001).

2.4. Métodos analíticos para detecção de adulteração

O objetivo dos métodos normalmente utilizados para avaliar a autenticidade de leite é identificar marcadores únicos que caracterizem a individualidade do leite de espécies diferentes, a exemplo das proteínas que são amplamente empregadas (GUERREIRO et al., 2013).

Métodos eletroforéticos, cromatográficos e imunológicos, de forma individual ou em conjunto são importantes exemplos de métodos para a separação, identificação e quantificação das proteínas do leite de diferentes espécies (FONTENELE, 2013).

Um dos métodos largamente aplicados na Itália para detecção de fraude em leite é a eletroforese, utilizada na diferenciação do leite de várias espécies e/ou na mistura entre eles, por meio da aplicação de potencial elétrico em um determinado tipo de gel, onde moléculas de cargas positivas e negativas são atraídas para seus polos opostos e moléculas de menor massa molecular terão mais facilidade de migrar que as maiores, sendo possível a separação e identificação das frações protéicas. Uma das mais tradicionais técnicas de eletroforese é a realizada em gel de poliacrilamida (PAGE) na presença do surfactante dodecil sulfato de sódio (SDS) (SDS-PAGE - "Polyacrylamide Gel Electrophoresis in Sodium Dodecyl Sulfate") (GÓMEZ-RUIZ et al., 2004; EGITO et al., 2006; SILVA, 2010).

O método oficial europeu de referência para detectar proteína bovina em produtos lácteos é a focalização isoelétrica (IEF Isoelectric focusing) da γ -caseína, que se baseia na separação de acordo com o ponto isoelétrico (pI) da proteína. No ponto isoelétrico a migração da proteína se estabiliza por apresentar carga líquida nula, ou seja, o número de cargas positivas é similar ao número de cargas negativas (ENNE et al., 2005; LÓPEZ-CALLEJA et al., 2005; LOCCI et al., 2008).

Os métodos de cromatografia líquida de alta eficiência (CLAE) estão entre os mais empregadas na detecção de adulterações em misturas, destacando-se a RP-HPLC ("Reverse-phase High Pressure Liquid Chromatography" ou cromatografia líquida de fase reversa) (SILVA, 2010; FONTENELE, 2013).

Ensaio imunológico como o método ELISA ("Enzyme-Linked Immunosorbent Assay" ou Ensaio de Imunoabsorção ligado à enzima) detecta a concentração de uma proteína de interesse em uma amostra que pode conter muitas proteínas diferentes, porém nem sempre é possível discriminar leite de espécies intimamente relacionadas, como a de vaca e búfala, e também não é aplicável a amostras tratadas termicamente (SILVA, 2010; DALMASSO et al., 2011).

Apesar da importância do leite de búfala para o mercado, a legislação brasileira ainda não possui um método oficial validado para detecção de fraude em leite bubalino, bem como, de outras espécies com adição de leite bovino. Embora a autenticidade do leite e de produtos lácteos sempre tenha sido determinada por métodos analíticos mais tradicionais e sejam considerados como referência, eles requerem análises laboriosas, equipamentos sofisticados, de alto custo, que exigem operadores altamente qualificados e treinados, além de consistirem em processos demorados, que necessitam da aquisição de reagentes químicos, os quais, na maioria dos casos oferecem risco à saúde humana (LIN, 2009; KAMAL e KAROUI, 2015).

O desenvolvimento de tecnologias automatizadas que apresentem um rápido resultado e possua melhor biossegurança torna mais eficiente as fiscalizações para identificação de fraudes em alimentos, além de trazer melhorias à qualidade dos produtos (EGITO et al., 2006).

Assim, por razões éticas e econômicas torna-se imprescindível o desenvolvimento de métodos sensíveis para a detecção de outros tipos de leite em produtos lácteos originados do leite de búfala, que garantam a autenticidade do produto

e ao mesmo tempo proporcionem rapidez, segurança e resultados consistentes (EGITO et al., 2006; RAYNAL-LJUTOVAC et al., 2008).

2.5. Espectroscopia no Infravermelho

Existe uma necessidade de se desenvolver métodos analíticos rápidos, precisos, baratos e eficientes para a detecção de adulteração já que estes possuem muitas vantagens comparadas às técnicas mais tradicionais. Recentemente, mais atenção tem sido dada ao desenvolvimento das técnicas não invasivas, não destrutivas e eficazes, tal como a espectroscopia no infravermelho (DOMINGO et al., 2014).

Técnicas espectroscópicas têm sido cada vez mais utilizadas para análise da presença e de concentrações de substâncias químicas em alimentos. Diferentes técnicas como a espectroscopia eletrônica, de fluorescência, espectroscopia raman, espectroscopia de absorção no infravermelho e as variáveis destas técnicas têm sido cada vez mais utilizadas juntamente com análises estatísticas multivariadas com esta finalidade (FERRÃO et al., 2004; MORGANO et al., 2005).

Esta técnica é considerada rápida, sensível, não invasiva, não destrutiva, de custo relativamente baixo, fornecendo uma grande quantidade de informações com apenas um teste, pois demanda pouco ou nenhum pré-tratamento da amostra. É limpa, uma vez que dispensa o uso de reagentes químicos e evita a geração de resíduos danosos ao ambiente, além dos equipamentos serem relativamente fáceis de manusear (LOHUMI et al., 2013; LOHUMI et al., 2015).

A espectroscopia no infravermelho compreende a interação da radiação eletromagnética com a matéria, sendo um dos seus principais objetivos a determinação de grupos funcionais de uma amostra por meio da variação dos níveis de energia vibracional dos átomos e moléculas, fornecendo informações químicas qualitativas e quantitativas. Isso ocorre por meio da análise da absorção de radiação em diferentes comprimentos de ondas ou da frequência de energia no espectro de grupos químicos funcionais específicos (LUYKX e VAN RUTH, 2008).

Segundo Pavia (2010), as frequências de radiação no infravermelho absorvidas equivalem às frequências vibracionais naturais de uma molécula e a energia absorvida eleva a amplitude dos movimentos vibracionais das suas ligações, ou seja, quando a radiação interage com moléculas orgânicas são gerados movimentos vibracionais e

rotacionais que podem interagir com o campo elétrico da radiação e causar variações na amplitude de um de seus movimentos o que ocasiona na absorção da radiação no infravermelho e formação das bandas/picos.

A energia denominada de infravermelho corresponde à região do espectro eletromagnético situada na faixa de comprimento de onda entre 14.290 cm^{-1} a 200 cm^{-1} , que engloba as regiões do infravermelho próximo, médio e distante. A região denominada infravermelho médio apresenta número de ondas entre 4000 cm^{-1} a 400 cm^{-1} e é mais comumente utilizada pela química orgânica (BARBOSA, 2007; HOLLER, 2009).

A Espectroscopia no Infravermelho Médio (MIR) tem revolucionado os laboratórios de análise de leite, por apresentar grandes vantagens em relação às técnicas anteriores pois fornece informações de frequências e absorção mais intensas na determinação de proteína, gordura e lactose em uma única amostra sem tratamento prévio, rapidez analítica e possibilidade de automatização total do processo (REID et al., 2005; BORIN et al., 2006).

A análise espectroscópica baseia-se na energia vibracional dos átomos de uma molécula e o espectro gerado é registrado por meio de um feixe de luz infravermelha. A energia com que qualquer pico aparece corresponde à frequência de vibração de uma “espécie molecular” presente na amostra, o que permite determinar grupos funcionais, uma vez que, cada grupo tem uma única frequência vibratória (FERRÃO et al, 2007; PAVIA et al., 2010).

Levando em consideração o efeito dos grupos funcionais, um espectro representa uma "impressão digital" única, uma vez que, cada molécula tem seu espectro característico, cujos picos presentes correspondem às frequências de vibrações entre os átomos e a suas respectivas alturas correspondem à quantidade de determinada amostra (análise quantitativa), o que configura a espectroscopia um método viável tanto para identificar diferentes tipos de amostras (análise qualitativa) quanto para confirmar a detecção de uma falsificação (MORGANO et al.,2005; LUO et al., 2006; LUYKX e VAN RUTH, 2008; CASSOLI et al., 2011).

2.6. Infravermelho com Transformada da Fourier - FTIR

O infravermelho com transformada de Fourier (FTIR) é uma tecnologia atraente para a indústria de alimentos, pois medições simples, rápidas e não destrutivas de componentes físicos e químicos podem ser obtidas. Avanços em sua instrumentação juntamente com o desenvolvimento de métodos de análise de dados tornam esta tecnologia ideal para grandes volumes de amostras, triagem rápida e caracterização até dos menores componentes dos alimentos. As técnicas de FTIR já são conhecidas pelas indústrias de alimentos, no entanto, mais recentemente tem-se expandido seu potencial para o monitoramento de adulterações em alimentos (CASSOLI et al., 2011; KAMAL e KAROUI, 2015; LOHUMI et al., 2015).

O instrumento que se obtém o espectro de absorção no infravermelho médio de um composto é chamado de espectrômetro no infravermelho. Dois tipos de espectrômetros no infravermelho médio são bastante utilizados em laboratórios: instrumento dispersivo e o com Transformada de Fourier (FT). Ambos oferecem espectros de compostos em uma faixa comum entre 4.000 cm^{-1} a 400 cm^{-1} , porém, apesar dos dois produzirem espectros praticamente idênticos para um determinado composto, espectros de infravermelho com FT produzem o espectro mais rapidamente e com maior sensibilidade do que os instrumentos dispersivos (PAVIA, 2010).

O avanço da espectroscopia no infravermelho médio, como técnica para análise quantitativa, deve-se à combinação da transformada de Fourier com a nova geometria dos espectrofotômetros e a utilização do interferômetro de Michelson, tornando-os mais rápidos e robustos. O FTIR foi desenvolvido a fim de ultrapassar as limitações encontradas com os instrumentos de dispersão, cuja principal dificuldade era o processo de digitalização lenta (THERMO NICOLET, 2001; HELFER et al., 2006).

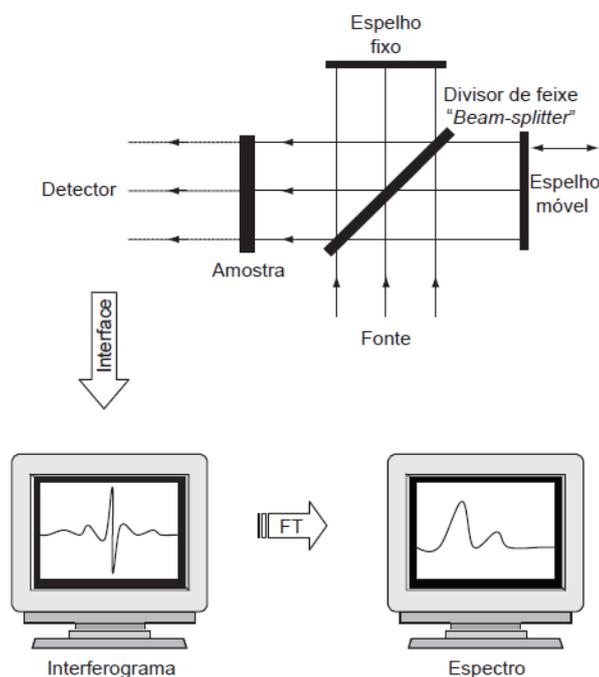
A solução foi desenvolvida com o emprego de um dispositivo óptico muito simples chamado de um interferômetro de Michelson que produz um único tipo de sinal com todas as frequências no infravermelho codificadas, podendo ser medidas em segundos, muito rapidamente, reduzindo assim o tempo de análise de uma amostra para outra (THERMO NICOLET, 2001; HELFER et al., 2006).

O interferômetro consiste basicamente de dois espelhos planos (um fixo e um móvel) e emprega um divisor de feixes (beam-splitter), que converte o feixe de entrada do infravermelho e o divide em dois feixes ópticos. Um feixe, 50% da radiação

incidente, é transmitido para o espelho móvel (situado a poucos milímetros do divisor de feixes) e o outro feixe é refletido para o espelho fixo. Os dois espelhos refletem os dois feixes e são recombinados quando se encontram no separador de feixes (Figura 1) (COATES, 1998; COATES, 2000).

Quando o feixe se move é gerado um comprimento fixo e o outro é constantemente modificado à medida que acontecem os movimentos do espelho (Figura 1). Portanto, a soma de todas as interações construtivas e destrutivas para cada componente resulta num sinal complexo denominado interferograma, o qual possui uma propriedade única em cada ponto (função resultante da mudança da posição do espelho), sendo possível originar informações sobre diversas frequências no infravermelho para determinada amostra. Ou seja, ao passo que o interferograma é mensurado, todas as frequências são mensuradas simultaneamente resultando em análises extremamente rápidas (COATES, 2000).

Figura 1: Esquema ilustrativo do espectrômetro Infravermelho com Transformada de Fourier (FTIR).



Fonte: HELFER et al. (2006)

Contudo, o sinal do interferograma não pode ser interpretado diretamente, sendo necessário um meio de decodificá-lo. Após a aquisição do interferograma é aplicada uma técnica matemática chamada Transformada de Fourier que converte os dados

obtidos no interferômetro em um espectro que relaciona a intensidade versus frequência (número de onda), cuja transformação é realizada pelo computador que, em seguida, apresenta a informação espectral desejada para análise (Figura 1) (THERMO NICOLET, 2001; MORGANO et al., 2005; HELFER et al., 2006).

O método por refletância pode apresentar acessórios para obtenção de espectros por Refletância Total Atenuada (ATR), Refletância Especular e Refletância Difusa (SKOOG et al., 2002). Atualmente, espectroscopia no infravermelho médio tornou-se mais disponível para a análise de produtos lácteos, principalmente após a introdução de ATR (KAROUI et al., 2010).

O método de refletância difusa é baseado em reflexos aleatórios nas partículas da amostra, contudo, a dificuldade com esta abordagem decorre do considerável efeito do tamanho da partícula, sua forma, densidade e a sua distribuição sobre o espectro, além da lentidão do processo. Isso dificulta a análise da amostra de leite, inclusive devido à heterogeneidade da sua composição, que abrange glóbulos de gordura, micelas de caseína, globulinas e moléculas de lactose (SCHMILOVICH et al., 2000).

No método ATR, a radiação infravermelha não é guiada através da amostra em si, mas através do cristal que entra em contato com a amostra, cujo índice de refração é bastante elevado. O feixe é refletido várias vezes no interior do cristal antes de ser dirigido para o detector, e ao atingir a superfície refletora, o mesmo penetra na amostra até uma determinada profundidade para gerar o espectro (MELLING e THOMPSON, 2002).

2.7. Utilização do FTIR na detecção de adulteração em leite e derivados

A combinação do FTIR e ATR tem sido utilizada com sucesso para investigar componentes do leite e de produtos lácteos bem como seus grupos químicos específicos (COROIAN et al., 2012; JAISWAL et al., 2015). Nas últimas décadas, os métodos espectroscópicos associados às ferramentas quimiométricas têm sido aplicados para avaliar a qualidade de produtos lácteos como alternativa para substituir os procedimentos de referência (ETZION et al. 2004; NICOLAOU et al., 2010; CASSOLI et al., 2011; BOTELHO et al., 2015; JAISWAL et al., 2015).

Estudos recentes avaliaram os principais componentes (proteína, gordura e lactose) de diferentes tipos de leite (vaca, cabra, ovelha) bem como a presença de

determinados adulterantes utilizando FTIR e calibração multivariada para análise quantitativa por meio da correlação de regiões espectrais de absorção de cada espécie (INON et al, 2004; DENG et al., 2006; PAPPAS et al., 2008; NICOLAOU et al., 2010; CASSOLI et al., 2011; SANTOS et al., 2013a; VALENTI et al., 2013; BOTELHO et al., 2015; JAISWAL et al., 2015).

As regiões espectrais e os picos ou bandas geradas pela análise espectroscópica correspondem a possíveis compostos presentes nas amostras de leite e seus respectivos grupos funcionais, bem como as características das moléculas orgânicas. Normalmente, as ligações de ácidos graxos são identificadas em picos surgidos na faixa de comprimento de onda entre 3500 cm^{-1} a 1747 cm^{-1} , advindos de ligações CH_2 e C-O . Os grupos funcionais C=O , N-H e C-N originam os picos referentes às amidas I e II e III, respectivamente, entre as faixas de 1660 cm^{-1} a 1241 cm^{-1} . Os carboidratos, como a lactose, geram bandas envolvendo a presença de ligações C-O/C-C/C-O-O na faixa entre 1150 cm^{-1} a 1000 cm^{-1} de absorbância (COATES, 2000; LEI et al., 2010; NICOLAOU et al., 2010).

Nicolaou et al. (2010) avaliaram a diferenciação de leites das espécies vaca, cabra e ovelha por meio de misturas binárias e terciárias em combinação com métodos estatísticos multivariados e verificaram as diferenças existentes em suas composições com a sobreposição das bandas derivadas dos espectros dos leites, com destaque ao comprimento de onda referente às gorduras (2.927 cm^{-1}) cuja intensidade se apresentou maior para o leite de ovelha.

Diferenças foram detectadas entre os leite de cabra e ovelha por Pappas et al. (2008), em que foi utilizado o FTIR para a obtenção de espectros com a finalidade de identificar e diferenciar os leites desengordurados por meio da detecção de uma região espectral de “impressão digital” de cada espécie. Dos onze picos gerados entre 1840 cm^{-1} a 950 cm^{-1} foram notadas pequenas diferenças entre as amostras de leite, sendo que a principal diferença foi localizada na banda 1745 cm^{-1} , referente à vibração do carbonil do grupo éster referente aos açúcares, a qual foi mais intensa para o leite de cabra que o de ovelha devido ao maior grau de esterificação dos açúcares.

Assim, o FTIR tem sido utilizado para a diferenciação de leites oriundos de distintas espécies animais e também de diferentes tipos de queijos com a determinação da sua composição, diferenciação quanto à textura, aspectos sensoriais, vida de prateleira, variações sazonais e identificação das regiões onde foram produzidos por

meio da análise da maturação, acidez, grau de deformação, gordura, dentre outros aspectos (COROIAN et al., 2012).

2.8. Quimiometria

A estatística univariada tem sido aplicada a processos químicos experimentais, contudo a sua utilização tornou-se limitada em estudos que pretendem avaliar, simultaneamente, muitas variáveis. Assim, a análise multivariada vem sendo introduzida no tratamento de dados químicos, obtendo rapidamente espaço nos estudos científicos e dando origem a uma nova disciplina, a Quimiometria. Este é um campo de pesquisa interdisciplinar que tem como ferramenta a aplicação da estatística multivariada, métodos matemáticos e computacionais, para investigar, interpretar, classificar e maximizar as informações extraídas de um conjunto de dados de origem química, cujas principais áreas são o planejamento e otimização de processos experimentais, reconhecimento de padrões e a calibração multivariada (SENA et al., 2000; SOUZA et al., 2011; SOUZA E POPPI, 2012).

Análise multivariada se refere a todas as técnicas estatísticas que, simultaneamente, analisam múltiplas medidas sobre indivíduos ou objetos sob investigação. Assim, qualquer análise simultânea de mais do que duas variáveis pode ser considerada, a princípio, como multivariada (MINGOTI, 2007; HAIR et al., 2010)

Na área de planejamento de experimentos busca-se encontrar quais as variáveis que mais afetam um determinado processo, assim como a interação entre elas. No reconhecimento de padrões, a partir de uma vasta gama de informações (medidas químicas ou espectrais, por exemplo) sobre uma série de objetos, pretende-se encontrar agrupamentos de amostras (objetos) que são similares entre si e, assim, detectar tendências nos dados. Na calibração multivariada, busca-se estabelecer um modelo que relacione uma série de medidas (químicas ou espectrais) realizadas em amostras com uma determinada propriedade (SOUZA E POPPI, 2012).

Segundo Patras et al. (2011), a utilização do raciocínio multivariado é uma técnica empregada para caracterizar e controlar a qualidade de produtos alimentícios, com representação relativamente simples de conjuntos de dados complexos, o que é importante para a compreensão adequada das relações existentes entre as inúmeras variáveis analisadas em um alimento.

Quando utilizada para avaliar dados instrumentais, a Quimiometria muitas vezes resulta em uma avaliação mais precisa e mais rápida da composição de um produto alimentício, ou até mesmo das propriedades físicas e sensoriais. Por exemplo, a composição (gordura, fibra, umidade, carboidratos) dos produtos lácteos pode ser mensurada utilizando-se a espectroscopia no infravermelho e avaliada por quimiometria (INON et al., 2004; SOUZA et al., 2011).

Métodos quimiométricos tem sido uma ferramenta útil na avaliação da qualidade, identidade e controle de parâmetros dos processamentos de produtos lácteos (WATKINS e WIJESUNDERA, 2006; KASEMSUMRAN et al., 2007; RODRIGUEZ-NOGALES e VASQUEZ, 2007; FAYE et al., 2008; SACCO et al., 2009; SOLA-LANARRANGA e NAVARRO-BLASCO, 2009). O emprego da técnica de espectroscopia no infravermelho com o auxílio de ferramentas quimiométricas mostrou-se efetivo para a quantificação dos principais nutrientes do leite (gordura, proteína, carboidrato, extrato seco total) (WU et al., 2007; WU et al., 2008).

A Quimiometria também vem sendo empregada conjuntamente aos métodos analíticos ricos em dados, como a espectroscopia no infravermelho, espectrometria de massa ou cromatografia, para identificar a autenticidade de produtos lácteos, evidenciando a sua capacidade para a determinação da autenticidade e/ou a detecção de falsificação em comparação aos procedimentos de referência (MOORE et al., 2012; CAPUANO et al., 2014).

Os métodos de análise multivariada, em se tratando do emprego de técnicas espectroscópicas no infravermelho, possibilitam manipular dados de absorbância espectral associados a mais de uma frequência ao mesmo tempo, sendo possível modelar propriedades químicas e físicas de amostras a partir de seus dados espectroscópicos (NICOLAOU et al., 2010; BOTELHO et al., 2015; JAISWAL et al., 2015).

Nesse sentido, os métodos multivariáveis utilizam, simultaneamente, dois ou mais valores de absorbâncias de um determinado comprimento de onda de uma amostra, enquanto que nos métodos clássicos de análise univariável, se utiliza somente a absorbância de uma frequência. A precisão dos métodos univariáveis é dependente da capacidade para identificar uma única banda isolada para cada componente. Já os métodos multivariáveis podem ser utilizados igualmente quando as informações espectrais estão sobrepostas de vários componentes através de várias regiões espectrais

selecionadas (PATRAS et al., 2011).

Dentre as ferramentas quimiométricas associadas à espectroscopia no infravermelho importantes para identificar a autenticidade do leite de búfala tem-se a Análise de Componentes Principais (ACP), Análise de Agrupamento (AA), Análise Discriminante (AD) e Redes Neurais Artificiais (RNA).

2.8.1. Análise de Componentes Principais (ACP)

A Análise de Componentes Principais (ACP) é uma das técnicas multivariadas mais importantes utilizadas na quimiometria e é a base para diversos métodos de reconhecimento de padrões, classificação e calibração multivariada, uma vez que, dentre outros fatores, constrói a base onde se fundamentam a maioria dos outros métodos multivariados de análise de dados. Normalmente, a ACP é utilizada com o objetivo de visualizar a estrutura dos dados, encontrar similaridades entre amostras, detectar amostras anômalas (“outliers”) e reduzir a dimensionalidade do conjunto de dados (LYRA et al., 2010; SOUZA E POPPI, 2012).

Essa técnica fundamenta-se essencialmente na transformação do conjunto original de variáveis em outro, os componentes principais (CP), por meio das combinações lineares das variáveis originais. Esse novo conjunto formado apresenta dimensão equivalente ao conjunto original, ou seja, se for analisado um conjunto de p variáveis relacionadas, teremos p componentes principais ortogonais, que devem ser mutuamente não correlacionados, por consequência da ortogonalidade (PEREIRA e PEREIRA, 2004; DA SILVA e PADOVANI, 2006; FERREIRA, 2011; KAMAL E KAROUI, 2015).

O método reduz a dimensionalidade por meio da representação do conjunto de dados no novo sistema de eixos (CP), permitindo a visualização da natureza multivariada dos dados em poucas dimensões. No espaço original, as amostras são pontos localizados em um espaço n -dimensional, sendo n igual ao número de variáveis. Com a redução de dimensionalidade proporcionada pela ACP, as amostras passam a ser pontos localizados em espaços de dimensões reduzidas definidos pelas CPs, por exemplo, bi- ou tridimensionais (SOUZA E POPPI, 2012).

Cada CP gerado contém uma quantidade máxima de variação em relação aos dados originais, que não são explicadas pelos demais componentes, uma vez que estes

não apresentam correlação. Assim, as informações totais dos dados originais estão contidas nos CP gerados, onde cada componente retém uma porcentagem da variância original (BORGOGNONE et al., 2001; PACHECO et al., 2013).

Os CP são obtidos em ordem decrescente de máxima variância. Cada CP sucessivo contém menos da variabilidade total do conjunto de dados inicial e, dessa forma, o componente principal 1 é a combinação linear de maior variância, detendo mais informação dos dados originais que o componente principal 2, que por sua vez tem mais informação que o componente principal 3, e assim sucessivamente. Cada CP é uma combinação linear de todas as variáveis originais, sendo independentes entre si e estimados com o propósito de reter, em ordem de estimação, o máximo de informação, em termos da variação total contida nos dados (NETO e MOITA, 1998; JOHNSON e WICHERN, 1999; BORGOGNONE et al., 2001; FERREIRA, 2011).

A ACP foi utilizada no estudo de Borin et al. (2006) para avaliar no leite em pó a presença de adulterantes comuns (amido, soro de leite ou sacarose) usando espectroscopia no infravermelho próximo. Foi verificada a separação das amostras de leite em pó adulteradas com base no tipo de adulterante sendo possível notar uma clara separação entre os grupos.

Nicolaou et al. (2010) fez uso da ACP para a separação de amostras adulteradas de leite de ovelha com leite de cabra cujas combinações variaram de 0% a 100%, aumentando a cada 5%. Foi possível separar as amostras adulteradas em função do nível de adulteração, contudo, foi observado que as amostras adulteradas com 10% a 40% possuíam comportamento semelhante na análise de dispersão. Os grupos com 50% a 65% apresentavam comportamento parecido e os que possuíam entre 70% e 100% apresentaram o mesmo comportamento. Nota-se com este estudo que amostras com menor nível de adição do leite de cabra ao leite de ovelha apresentaram dificuldade em sua identificação e separação, algo que não é interessante para a determinação da autenticidade do produto.

Carvalho et al. (2015) trabalharam com a detecção de adulteração de leite em pó com soro de leite empregando diferentes níveis de adulteração por meio da ACP. A análise de componentes principais foi capaz de separar amostras adulteradas de acordo com a concentração do soro presente no leite. No entanto, amostras adulteradas com níveis mais baixos de soro (1%, 2%, 4%, 6%) se comportaram de forma semelhante às amostras sem adulteração, sendo possível observar também que a detecção de

adulteração em leite em pó é mais complexa em menores níveis de adição do adulterante.

Jaiswal et al. (2015) avaliaram a detecção de extrato de soja em leite de búfala utilizando a técnica espectroscópica FTIR-ATR no qual a ACP demonstrou diferenças claras na faixa de comprimento de onda $1680-1058\text{ cm}^{-1}$, referentes às proteínas entre o extrato de soja e o leite bubalino. Este fato ajudou na interpretação das diferenças e semelhanças entre amostras por meio de padrões de agrupamento destas evidenciando que amostras com composições distintas apresentam comportamento também distinto na dispersão gráfica dos componentes principais em estudo.

2.8.2. Análise de Agrupamento (AA) ou Cluster Analysis

O método que também se apresenta útil na classificação de amostras com diferentes características é a Análise de Agrupamentos ou Análise de Conglomerados, ou *Cluster Analysis* que tem como objetivo agrupar os dados de forma a identificar semelhanças entre os objetos de estudo, ou seja, dividir os elementos de uma população em grupos de forma que aqueles pertencentes a um mesmo grupo sejam similares entre si e os elementos em grupos diferentes sejam heterogêneos em relação às características que neles foram medidas (variáveis) (HELFER et al., 2006; DOMINGO et al., 2014).

O método mais comum é a classificação hierárquica, em que os objetos são agrupados por meio da semelhança de uma classificação taxonômica e representados num gráfico com uma estrutura em árvore denominada dendograma que agrupa amostras em função da similaridade cuja escala varia de zero (amostras sem similaridades) a um (amostras similares) (SEIDEL et al., 2008).

As técnicas hierárquicas aglomerativas partem do princípio de que no início do processo de agrupamento tem-se n conglomerados, ou seja, cada elemento do conjunto de dados observado é considerado como sendo um conglomerado isolado. Em cada passo do algoritmo, os elementos amostrais vão sendo agrupados, formando novos conglomerados até o momento em que todos os elementos considerados estão num único grupo (MINGOTI, 2005).

Para proceder esta classificação é necessário definir matematicamente o conceito de proximidade, ou seja, a distância entre dois objetos. A técnica de agrupamento começa com o cálculo da matriz de similaridade ou dissimilaridade, também chamada

de matriz de proximidade, entre os elementos observados. A intensidade do agrupamento é uma propriedade não só dos dados utilizados mas, também, do processo sob o qual esses dados são submetidos (RECHE, 2013).

O processo de agrupamento envolve basicamente duas etapas: a estimativa de uma medida de proximidade e a adoção de uma técnica de agrupamento. Entre as medidas mais usuais para estabelecer o conceito de distância entre n e m , destacam-se o coeficiente de correlação de Pearson, que é uma medida de similaridade, e as distâncias Euclidianas e de Mahalanobis, que são medidas de dissimilaridade. Medida de similaridade indica que quanto maior for esta medida, maior a semelhança entre os grupos estudados, enquanto que, medida de dissimilaridade sugere que quanto maior for a sua medida menor será a semelhança entre os grupos (HELFER, et al., 2006; SEIDEL et al., 2008).

A distância euclidiana entre amostras ou variáveis é calculada e transformada em uma matriz cujos elementos são índices de similaridade que variam de 0 a 1, quanto menor for a distância, maior será a similaridade dos elementos comparados. Enquanto que, para o coeficiente de correlação, quanto maior for o valor obtido mais parecidos serão os objetos formados. A distância de Mahalanobis é muito utilizada quando existem repetições dentro das unidades amostrais e principalmente quando os caracteres são correlacionados e quando existe mais de dois para o processo de agrupamento (MINGOTI, 2005).

2.8.3. Análise Discriminante (AD)

A Análise Discriminante (AD) é uma técnica estatística de classificação que determina uma classe ou classes de amostras conhecidas que são muito semelhantes a uma amostra desconhecida, ou seja, determina um conjunto de observações em classes pré-definidas. O método deve especificar duas classes de amostras conhecidas e podem ser usados padrões múltiplos para descrever cada classe. Quando se trata de dados espectrais podem ser usadas múltiplas regiões do espectro para a análise (JIN e AN, 2011).

Este método é utilizado quando se quer estabelecer relações entre uma variável dependente e variáveis independentes, e tem por finalidade criar índices que diferenciam ou discriminam os grupos. Por consequência, se permite a identificação de

grupos similares e a classificação de novos casos, onde a inserção é feita no grupo que o caso tem maior probabilidade de pertencer, com base nas características do novo objeto (MAROCO, 2003).

É possível obter uma função discriminante, que é definida como uma fronteira que separa regiões de alocação, onde o conjunto de valores é separado em grupos através de combinações lineares de variáveis independentes. A partir dessas combinações lineares é possível classificar os elementos em cada uma das categorias da variável dependente com alto poder de discriminação (MAROCO, 2003; HAIR JR. et al., 2010; PESTANA e GAGEIRO, 2014).

Essa função tem também a propriedade de minimizar as probabilidades de má classificação, quando as populações são normalmente distribuídas com média e variância conhecidas e os grupos podem ser visualizados em um espaço multidimensional, muitas vezes menor que o original. Os coeficientes das funções discriminantes indicam a contribuição das variáveis originais para cada função discriminante (REGAZZI, 2000; SCHEEREN, 2000).

A análise discriminante linear, também conhecida como o critério de Fisher, é a técnica mais conhecida desta classe e se caracteriza por maximizar a razão entre as somas dos quadrados entre grupos e a soma de quadrados dentro dos grupos, sendo uma combinação linear das variáveis originais e se caracterizando por produzir separação máxima entre duas populações (MINGOTI, 2005; DAS e NENADIC, 2008).

O algoritmo da análise discriminante pode ser usado para separar uma variedade de amostras entre espectros de amostras desconhecidas, com base em um modelo de calibração para determinar qual classe de padrões é muito semelhante ao desconhecido. O resultado é o nome da classe que é muito semelhante ao espectro desconhecido e uma medida da distância entre ele e cada classe constituinte do modelo (HELFER et al., 2006).

Os estudos de Cozzi et al. (2009), Karoui et al. (2011), Ottavian et al. (2012) e Andueza et al. (2013) relataram que a espectroscopia associada à análise discriminante poderia ser considerada uma ferramenta confiável para discriminar e identificar a autenticidade de leites e queijos de acordo com diferentes regiões agrícolas, período de produção e sistema de alimentação dos animais (pastagens).

2.8.4. Redes Neurais Artificiais (RNA)

Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no sistema nervoso, com capacidade de aquisição e manutenção de conhecimento, além de representarem um conjunto de unidades de processamento caracterizado por neurônios artificiais, que são interligados por um grande número de interconexões. As características mais atrativas das RNA está associada com a elevada habilidade em mapear sistemas não-lineares, aprendendo os comportamentos envolvidos a partir de informações obtidas (SILVA et al., 2010)

O uso das redes neurais em relação aos modelos empíricos revela vantagens pois este é um método capaz de capturar relações funcionais entre variáveis de entrada e saída a partir de uma base de dados, representando uma ferramenta valiosa para os sistemas de processamento de alimentos, envolvendo estimativa, previsão e classificação (SINGH et al., 2009).

As RNA possuem uma estrutura interconectada composta por muitos elementos cuja unidade básica para o processamento de informações é o neurônio artificial. Este pode receber uma ou mais entradas transformando-as em saídas sendo capazes de realizar cálculos em paralelo para o processamento de dados (BAS e BOYACI, 2007). Cada entrada tem um peso associado que determina sua intensidade de influência no dado de saída. Na definição da arquitetura de uma RNA, consideram-se o número de camadas, o número de neurônios e a conexão entre os neurônios (DANESHVAR et al., 2006; ABER et al., 2009).

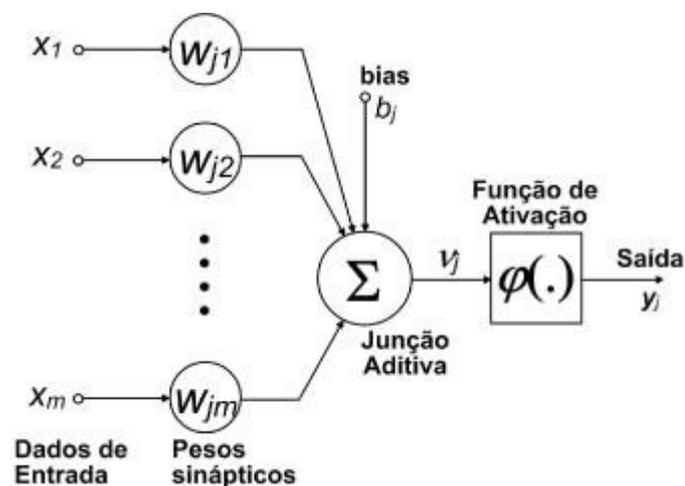
As RNA podem lidar com múltiplas variáveis dependentes e independentes ao mesmo tempo, sem o conhecimento prévio da relação funcional que possa existir entre elas, além disso, também podem funcionar como modelos preditivos que descrevem a relação funcional entre as variáveis de entrada e variáveis de saída de um sistema (ZHOU e THERDTHAI, 2010).

Os neurônios artificiais na camada de entrada recebem sinais que são passados para as camadas intermediárias, por meio de conexões matemáticas. Cada neurônio processa uma parte das informações, o que ocorre em paralelo, e automaticamente as distribui através das diferentes unidades que compõem a rede (pesos sinápticos). As unidades de processamento transferem as somas ponderadas de informações por uma função de ativação que contém um grande número de sinais informativos, os quais são

convertidos em valores definidos durante a operação de treinamento e teste (Figura 2) (SINGH et al., 2009; ELMOLLA et al., 2010; KHATAEE et al., 2010).

Um dos aspectos mais relevantes das redes neurais artificiais está na capacidade de aprendizagem a partir de apresentações de amostras padrões que refletem o comportamento do sistema. Em seguida, após a rede ter aprendido o relacionamento entre as entradas e saídas, esta é capaz de generalizar soluções. A rede será então capaz de produzir uma saída próxima daquela esperada a partir de quaisquer sinais inseridos em suas entradas (SILVA et al., 2010).

Figura 2: Funcionamento do neurônio artificial.



Fonte: FIORIN et al., 2011

Onde: X_1, X_2, \dots, X_m são os sinais de entrada; $W_{j1}, W_{j2}, \dots, W_{jm}$ os pesos sinápticos, valores que servirão para ponderar cada variável de entrada; Σ , é a combinação linear ou função de soma ou junção aditiva, a qual agrega todos os sinais de entrada; bias (b_j), o limiar de ativação, que especifica o patamar apropriado para gerar de disparo em direção à saída do neurônio; v_j , potencial de ativação, o resultado produzido pela diferença do valor produzido entre a junção aditiva e o limiar de ativação; $\varphi(\cdot)$, Função de ativação, que limita a saída do neurônio dentro de um intervalo de valores razoáveis e y_j , o sinal de saída (SILVA et al., 2010).

O processo de treinamento de uma rede neural consiste na aplicação de passos

ordenados que sejam necessários para a sintonização dos pesos sinápticos e limiares de seus neurônios (função de ativação), tendo como objetivo final a generalização de soluções a serem produzidas pelas suas saídas, cujas respostas são representativas do sistema que estão sendo mapeadas. Desta forma, os valores previstos são subsequentemente comparados com o valor real e as conexões entre as unidades de processamento com seus devidos pesos. A rede mede o erro entre o valor obtido e o valor desejado, usando o algoritmo para ajustar os pesos das interconexões e reduzindo o erro nas determinações subsequentes. Pela utilização desta estrutura de rede, o programa continua a fazer alterações nas conexões internas, através do processo de tentativa e erro, até atingir um determinado nível de certeza (RAI et al., 2005).

Na composição de redes neurais são utilizados algoritmos de aprendizagem que operam de maneira diferente. Além disso, as redes diferem também em sua arquitetura, ou organização dos neurônios. As redes de multicamadas comumente se utilizam do algoritmo *Backpropagation*, no qual o aprendizado se dá por um processo de retropropagação do erro. O *Resilient Propagation* uma versão aperfeiçoada do algoritmo *Backpropagation*, , atualmente também é bem difundido e tem por objetivo tornar o processo de convergência mais eficiente (SILVA et al., 2010).

As redes neurais artificiais têm sido utilizadas com sucesso por pesquisadores para classificar amostras de leite quanto à sua composição e também em relação a detecção de adulterações. Valente et al. (2014) aplicaram as RNA para classificar amostras de leite adulteradas quanto à adição de soro de queijo com diferentes concentrações apresentando 95% de acertos na classificação.

Balestrieri et al (2001) realizaram a classificação de amostras de leite de cabra, ovelha e búfalo por meio da RNA de acordo com a composição química de cada espécie e apresentou erro de 0,199 e 96% de classificação correta, estudo que se assemelhou ao presente estudo que propõe detectar e classificar a adição de leite de vaca ao leite de búfala utilizando a RNA. Nazário et al.(2009) propuseram uma arquitetura de rede neural para detectar adulteração de amostras de leite por adição de água, obtendo 96,3% de taxa de classificação, valor satisfatório para a detecção de adulteração.

3. OBJETIVOS

3.1. Objetivo Geral

Detectar a adição do leite de vaca ao leite de búfala por meio da análise espectroscópica FTIR-ATR e quimiometria

3.2. Objetivos Específicos

- Diferenciar o leite de búfala e vaca por meio da espectroscopia e quimiometria
- Detectar níveis de adição do leite vaca ao leite de búfala por meio da espectroscopia e quimiometria;
- Separar, agrupar e classificar amostras de leite de búfala, leite de vaca e de leite de búfala adicionado de leite de vaca por meio da Análise de Componentes Principais (ACP), Análise de Agrupamento e Análise Discriminante (AD).
- Criar uma Rede Neural Artificial para classificar amostras de leite de búfala adicionadas e não adicionadas de leite de vaca de acordo com o nível de adição e realizar a predição de amostras de leite desconhecidas.

4. MATERIAL E MÉTODOS

4.1. Coleta de dados e local do experimento

Amostras de leite fluido foram coletadas por quarenta e cinco dias consecutivos, entre os meses de abril e junho de 2016. O leite bovino oriundo de vacas mestiças Holandês x Zebu e o leite bubalino de fêmeas bubalinas Murrah foram adquiridos em um laticínio localizado no município de Itapetinga-BA. A obtenção das amostras de leite foi realizada no período da manhã de cada dia de coleta após seu recebimento no laticínio e logo em seguida, as análises foram realizadas no Laboratório de Processamento de Leite e Derivados da Universidade Estadual do Sudoeste da Bahia (UESB), Campus de Itapetinga - BA.

4.2. Preparo das amostras de leite de vaca e búfala

A partir das amostras de leite de vaca e de búfala, em cada dia de coleta, foram preparadas combinações de 1mL com quantidades variáveis e crescentes de leite de vaca adicionados ao leite de búfala, em sucessivos acréscimos de 10% até 90%, perfazendo 9 níveis de adição (T10, T20, T30, T40, T50, T60, T70, T80, T90), além das amostras de referência, exclusivamente com leite de búfala (tratamento referência búfala – TRB) e de vaca (tratamento referência vaca – TRV). Para cada amostra foram realizadas 45 repetições perfazendo um total de 495 unidades experimentais.

As amostras de referência (TRB e TRV) e amostras de leite de búfala adicionadas de leite de vaca de cada repetição foram armazenadas em eppendorfs e congeladas a -20°C por 48 horas e, posteriormente, liofilizadas por 24 horas no liofilizador de bancada FreeZone da marca Labconco® 4,5 L, a -50°C e bomba de vácuo de 86L/min, sendo obtidas cerca de 1g de cada amostra. Depois de liofilizadas, as amostras foram utilizadas para obtenção dos espectros por meio da análise no Infravermelho com Transformada de Fourier (FTIR).

4.3. Aquisição dos espectros

Os espectros foram obtidos através do FTIR-ATR - Infravermelho com Transformada de Fourier por meio da Reflexão Total Atenuada, no equipamento Agilent Cary® 630, sob a faixa do infravermelho médio, sendo empregado o

comprimento de onda de 4000 a 600 cm^{-1} , com resolução de 4 cm^{-1} , 64 scans e leitura através do cristal de diamante ATR.

As amostras foram colocadas sobre a superfície do diamante ATR para a realização da leitura. Antes de cada digitalização foi coletado o branco, onde foi feita a leitura sem a amostra na superfície. O tempo de aquisição dos espectros foi de aproximadamente 30 segundos.

Foram utilizados como variáveis os valores de comprimentos de onda dos picos adquiridos na faixa de 4000 a 600 cm^{-1} e para a análise estatística foi empregada a absorbância máxima destes picos que variavam entre -0,5 e 0,5 unidades de absorbância (AU).

Foram identificadas na literatura a faixa de comprimentos de onda onde apareceram os principais picos, os grupos funcionais possivelmente identificados nas amostras estudadas e as suas respectivas atribuições e utilizadas como referência para a análise espectral e discussão dos resultados (Tabela 1).

Tabela 1: Atribuições da absorção dos compostos orgânicos das amostras de leite.

Comprimento de onda (cm^{-1})	Grupo Funcional	Atribuição	Referência
3278	O-H	Gordura	Barbosa, 2007; Lei, 2010; Subramanian et al., 2011; Santos et al., 2013a
2919 - 2922	CH ₂	Gordura	Santos et al., 2013b
2851 - 2855	CH ₂	Gordura	Santos et al., 2013b
2081 - 2115	CH ₂	Gordura	Santos et al., 2013b
1742 - 1746	C=O	Gordura	Choudhury et al, 2005; Barbosa, 2007; Santos et al., 2013b
1645 - 1647	C=O	Proteína (Amida I)	Zaleska et al., 2002; Barbosa, 2007; Pappas et al., 2008; Santos et al., 2013; Jaiswal et al., 2014
1540 - 1559	N-H	Proteína (Amida II)	Zaleska et al., 2002; Pappas et al., 2008; Santos et al., 2013

1450 - 1459	CH	Gordura	Zaleska et al., 2002; Pappas et al., 2008
1376 - 1399	CH	Gordura	Zaleska et al., 2002; Pappas et al., 2008
1240 - 1251	N-H / C-H	Proteína (Amida III)	Zaleska et al., 2002; Pappas et al., 2008; Lei, 2010; Jaiswal et al., 2014
1147 - 1149	C-O / C-C / C-O-O	Carboidrato	Nicolaou et al., 2010
1019 - 1133	C-O / C-C / C-O-O	Carboidrato	Lei, 2010; Nicolaou et al., 2010
890 - 893	CH	Carboidrato	Santos et al., 2013
768 - 780	CH ₂	Gordura	Barbosa, 2007; Lei, 2010
699 - 703	N-H	Proteína	Barbosa, 2007; Lei, 2010

4.4. Tratamento Prévio dos Dados

Após a obtenção dos espectros, as amostras foram organizadas em conjunto de dados numéricos, os quais foram manipulados por meio do Microsoft Office Excel 2007®.

Foram realizadas as determinações de desvio padrão e coeficiente de variação. Os dados foram padronizados pela Equação 1, que consistiu na implementação de uma nova variável (Z) com média zero e desvio padrão igual a 1, eliminando as diferenças entre as unidades de medidas das variáveis estudadas.

$$Z = \frac{X - \mu}{\sigma} \quad \text{Eq. 1}$$

Onde:

Z: variável padronizada; X: variável; μ : média; σ : desvio padrão

4.5. Análise estatística multivariada dos dados

Para realização das análises estatísticas multivariadas foi utilizado o programa estatístico *Statistical Analysis System (SAS)® Studio*.

4.5.1. Análise de Componentes Principais (ACP)

A Análise de Componentes Principais foi dividida em duas etapas. A primeira teve como objetivo verificar as amostras de referência e a segunda avaliar a formação de grupos à medida que foram adicionadas amostras fraudadas.

Para o primeiro caso, os dados referentes à TRB e TRV foram padronizados de acordo com a Equação 1. A partir dos dados padronizados foi obtida a matriz de variância e covariância (S), a qual foi fatorada numa matriz diagonal, cujos elementos são iguais a $a_{ii} < \lambda_i$, $i = 1, 2, \dots, 11$. A matriz diagonal foi obtida resolvendo a equação característica da matriz S, Equação 2, contendo autovalores com $\lambda_1 > \lambda_2 > \dots > \lambda_{11}$.

$$|S - \lambda I| = 0 \quad \text{Eq. 2}$$

Onde: S corresponde à matriz de variância e covariância; λ_i : Autovalor e I: Matriz Identidade.

Para cada autovalor (λ_i) foi determinado um autovetor a^*_i a partir da solução do da Equação 3:

$$[S - \lambda I]a^*_i = \emptyset \quad \text{Eq. 3}$$

Onde: S corresponde à matriz de variância e covariância; λ_i : Autovalor e I: Matriz Identidade; a^*_i : o autovetor e \emptyset : conjunto vazio ou 0.

Os autovetores obtidos foram normalizados (a_i) por meio da Equação 4.

$$a_i = \frac{1}{\sqrt{a_{i1}^2 + a_{i2}^2 + \dots + a_{i11}^2}} \begin{bmatrix} a^*_{i1} \\ a^*_{i2} \\ \vdots \\ a^*_{i11} \end{bmatrix} = \frac{a^*_i}{a_i} \quad \text{Eq. 4}$$

Onde: a_i : Autovetor normalizado; $a_{i1}, a_{i2} \dots a_{i11}$, autovetores normalizados correspondentes aos grupos estudados; $a^*_{i1}, a^*_{i2} \dots a^*_{i11}$, autovetores não normalizados correspondentes aos grupos estudados.

Tomando-se os elementos do vetor a_i determinados como os coeficientes CP_i , temos que o i -ésimo CP é obtido pela Equação 5:

$$CP_i = a_{i1}z_1 + a_{i2}z_2 + \dots + a_{i,11}z_{11} \quad \text{Eq. 5}$$

Onde: $i = 1, 2, \dots, n$, em que: CP_i : i -ésimo componente principal; a_{i1}, a_{i2}, a_{in} , elementos do i -ésimo autovetor (a_i) normalizado, associado ao i -ésimo autovalor (λ_i) do i -ésimo componente principal (CP_i).

Para a escolha do número de componentes principais foi considerado o critério de fatores interpretáveis e o critério de Kaiser, os quais selecionaram os primeiros k , componentes principais que retêm juntos a maior proporção de variância dos atributos originais (acima de 70% da variância) e autovalor maior que um, respectivamente.

Após a obtenção do número de componentes principais foram confeccionados gráficos de dispersão dos coeficientes de correlação entre as variáveis e os CP, no espaço bi e tridimensional a fim de verificar a formação de agrupamentos de indivíduos similares mediante exames visuais.

4.5.2. Análise de Agrupamento (AA) ou Cluster Analysis

Inicialmente foi necessário definir matematicamente o conceito de proximidade, ou seja, a distância entre os grupos (TRB, TRV e T10 a T90) sendo realizado o cálculo da matriz de dissimilaridade entre os elementos observados.

O processo de agrupamento envolveu basicamente duas etapas: a estimação de uma medida de proximidade e a adoção de um método de agrupamento. Entre as medidas para estabelecer o conceito de distância entre n e m , foi utilizada distância de Mahalanobis uma medida de dissimilaridade, dada pela Equação 6:

$$D_{ii}^2 = (\bar{X}_i - \bar{X}_{i'})' S^{-1} (\bar{X}_i - \bar{X}_{i'}) \quad \text{Eq. 6}$$

Onde: D_{ii} : Distância de Mahalanobis; \bar{X}_i e $\bar{X}_{i'}$, Vetores aleatórios; S: Matriz de covariância.

O método utilizado para o processo de agrupamento foi o de *Ward*, o qual apresentou melhor coeficiente para o agrupamento dos dados. Segundo Hair et al. (2005), o método de Ward consiste em um procedimento de agrupamento cuja medida de similaridade é calculada como a soma de quadrados entre os agrupamentos feita sobre todas as variáveis, tendendo a resultar em agrupamentos de tamanhos aproximadamente iguais devido à sua minimização de variação interna. Em cada estágio, combinam-se os dois agrupamentos que apresentarem menor aumento na soma global de quadrados dentro dos agrupamentos.

Para a determinação do número de grupos formados foram avaliados o coeficiente de correlação cofenético (CCC), Estatística Pseudo F e Pseudo T² e Soma dos Quadrados (RSQ), e plotado o dendograma que associou em seu eixo vertical o nível de dissimilaridade e em seu eixo horizontal as amostras estudadas. As linhas verticais, partindo das amostras agrupadas tem altura correspondente ao nível de semelhança. Quanto maior a altura de cada grupo maior o erro que distancia um grupo do outro. Como os algoritmos de agrupamento não apresentam solução para determinação do número ideal de grupos formados, estes foram definidos por meio do exame visual do dendograma.

4.5.3. Análise Discriminante (AD)

A análise discriminante consistiu em obter funções matemáticas capazes de classificar uma amostra de leite em uma das populações π_i ($i=1$ – búfala, 2 – vaca, 3 – amostra com baixo nível de adição de leite de vaca ao leite de búfala, 4 – amostra com médio nível de adição de leite de vaca, 5 – amostra com nível alto de adição de leite de vaca), buscando minimizar a probabilidade de má classificação, isto é, minimizar a probabilidade de classificar erroneamente uma amostra em uma população π_1 , quando realmente pertence à população π_2, π_3, π_4 ou π_5 .

Realizou-se inicialmente, a parte exploratória da análise, onde foi estabelecido o número de grupos (5) e identificou-se as características que foram utilizadas com o objetivo de traçar o perfil geral de cada grupo visando realizar a discriminação ou separação das amostras, possibilitando a diferenciação das mesmas. Esta etapa consistiu no treinamento dos dados, em que foi utilizada a Função Discriminante de Anderson.

Todos os dados foram aleatorizados, utilizando a função aleatória do Microsoft

Excel®, em seguida organizados de acordo com os seus respectivos grupos, constando amostras padrões (TRB e TRV) e amostras de búfala adicionadas de leite de vaca. A aleatorização e distribuição dos dados ocorreram em cada grupo, sendo que para a fase de treinamento e validação, foram utilizados 80% e 20% dos dados, respectivamente, constando 36 amostras para treinamento e 9 amostras para validação de modo que o número de dados fosse igual para todos os grupos, totalizando 225 amostras para este experimento.

A partir dos dados originais foram estimadas médias e matrizes de covariâncias das populações estudadas e obtidas as funções discriminantes de Anderson para o processo de classificação das amostras. A função discriminante de Anderson consistiu em uma combinação linear das variáveis originais utilizadas no estudo a fim de maximizar a razão entre as classes e minimizar a dispersão dentro de cada grupo obtido, conforme se observa na Equação 7.

$$Y = \hat{\beta} + \hat{\beta}_1 * \gamma_{3278} + \hat{\beta}_2 * \gamma_{2851} + \dots + \hat{\beta}_{15} * \gamma_{700}$$

Eq.7

Onde: Y: Função linear discriminante; $\hat{\beta}_0$: Constante da função linear discriminante; $\hat{\beta}_1, \hat{\beta}_2 \dots \hat{\beta}_n$, são os coeficiente das variáveis utilizadas no estudo.

Após as fases anteriormente descritas, realizou-se a etapa de validação. A partir das equações lineares obtidas na fase de treinamento, os novos dados de validação foram aplicados a essas equações a fim de gerar as taxas de classificações e verificar a capacidade de generalização desses modelos.

4.5.4. Redes Neurais Artificiais (RNA)

Utilizou-se a rede neural artificial (RNA) a fim de possibilitar a separação dos tratamentos referência (TRB e TRV) e das amostras de leite de búfala com adição de leite de vaca em grupos, tais como, amostras com baixo nível de adição de leite de vaca (entre T10 a T20), médio nível de adição de leite de vaca (T30 a T40) e alto nível de adição de leite de vaca (T50 a T90).

4.5.4.1. Desenvolvimento e arquitetura da RNA

As RNA consistiram em uma estrutura complexa, composta de neurônios artificiais de entrada, da camada oculta e de saída capazes de realizar cálculos para o processamento de dados, sem o conhecimento prévio da relação existente entre eles. Cada neurônio na camada de entrada recebeu um sinal e enviou um sinal de saída para geração da informação, cuja força das ligações estava diretamente relacionada aos pesos aplicados em cada neurônio durante a etapa de aprendizagem.

Foi escolhida a metodologia supervisionada de redes neurais, denominada perceptrons de multicamadas ou *multilayers perceptrons* (MLPs), as quais são amplamente utilizadas dentre as classes de redes neurais supervisionadas. Foram utilizadas conexões “feed-forward” com parâmetros variando de -1 a 1 e o algoritmo de treinamento utilizado foi o *Resilient Propagation*, uma versão aperfeiçoada do algoritmo *Backpropagation*, cujo objetivo é tornar o processo de convergência mais eficiente (SILVA et al., 2010).

Para ativar um neurônio (transferir um sinal), algumas funções de ativação são usadas, como funções tangentes sigmóides ou hiperbólicas. A última é preferida quando se estudam sistemas biológicos ou mais complexos (RAI et al., 2005). A função de ativação utilizada foi do tipo de função hiperbólica (Equação 8):

$$\varphi = \tanh(\nu_k) \quad (8) \text{ Eq.}$$

Onde: ν_k é o potencial de ativação do neurônio; k^{th} (soma dos sinais de entrada ponderados pelos pesos sinápticos e o viés).

Com a finalidade de obter a melhor configuração da rede neural para a classificação das amostras de leite foi realizado um rastreamento do número de iterações, das camadas ocultas, e do valor do erro quadrático médio (RMSE - *Root Mean Square Errors*) nas etapas de treinamento e validação.

As redes foram formadas por 15 neurônios de entrada (correspondente as 15 variáveis/ 15 picos) e duas camadas ocultas, suficientes para aproximar qualquer função matemática. A escolha do número de neurônios em cada camada oculta foi baseada na minimização do RMSE associado à capacidade preditiva da rede e na maximização da taxa de classificação que é o mesmo que taxa de acertos dos grupos estudados, quanto maior a taxa de classificação, melhor a capacidade da rede em diferenciar e separar os grupos (RAI et al., 2005).

Nas duas camadas ocultas foram testados de 0 a 50 neurônios, totalizando 10 configurações com um número diferente de neurônios em cada camada oculta, e foram utilizadas 200 iterações até a escolha da melhor. Todas as redes testadas tinham os neurônios de uma camada totalmente conectados com os neurônios de uma camada vizinha (Tabela 2).

Tabela 2: Arquitetura das redes neurais artificiais testadas.

Arquitetura Utilizada	Parâmetros
Algoritmo	<i>Resilient Propagation</i>
Aprendizado	Supervisionado
Função de Ativação	Tangente Hiperbólica (tanh)
n° de variáveis de entrada	15
n° de camadas ocultas	2
n° de neurônios da 1ª camada oculta	10 a 50
n° de neurônios da 2ª camada oculta	0 a 50
n° de neurônios de saída	5
Iterações	200

Foi utilizado o software Java Neural Network Simulator, JavaNNS, versão 1.1 (Wilhelm-Schickard-Institute (WSI), Tübingen, Alemanha), para realização dos procedimentos matemáticos e estatísticos necessários (FISCHER et al., 2001).

4.5.4.2. Treinamento e Validação

Os dados foram aleatorizados e divididos em dois conjuntos de dados. O primeiro contendo 80% dos dados utilizado para a etapa de treinamento e o segundo contendo 20% para a etapa de validação. A aleatorização e distribuição dos dados ocorreram em cada grupo, constando 36 amostras para treinamento e 9 para validação de modo que o número de dados fosse igual para todos os grupos, totalizando 225 amostras para este experimento. Foram utilizadas 15 camadas de entrada (15 variáveis/picos). A saída da RNA foi representada por vetores multidimensionais contendo 5 neurônios, correspondente aos grupos de TRB, TRV, tratamentos com baixo nível de adição de leite de vaca ao leite de búfala (entre T10 a T20), médio nível de adição de leite de vaca (T30 a T40) e alto nível de adição de leite de vaca (T50 a T90), representados pelos vetores iguais ao número de grupos, sendo o conjunto de

dados criado em formato numérico, atribuindo 1 quando o grupo foi corretamente classificado e 0 caso contrário. Sendo representados da seguinte maneira:

(1,0,0,0,0) = correspondente às amostras TRB;

(0,1,0,0,0) = correspondente às amostras TRV;

(0,0,1,0,0) = baixo nível de adição de leite de vaca ao leite de búfala

(T10 a T20);

(0,0,0,1,0) = nível médio de adição de leite de vaca ao leite de búfala

(T30 a T40);

(0,0,0,0,1) = alto nível de adição de leite de vaca ao leite de búfala

(T50 a T90).

Os dados de entrada e saída foram normalizados por meio das Equações 9 e 10.

$$x_{i,norm} = \frac{x_i - \bar{x}_i}{R_{i,m\acute{a}x}} \quad (9) \text{ Eq.}$$

$$x_{i,norm} = \frac{x_i - x_{i,min}}{x_{i,m\acute{a}x} - x_{i,min}} \quad (10) \text{ Eq.}$$

Onde: $x_{i,norm}$ é o dado normalizado; x_i é o dado experimental; \bar{x}_i é a média dos dados; $R_{i,m\acute{a}x}$ é dado por $R_{i,m\acute{a}x} = \max[x_{i,m\acute{a}x} - x_i, x_i - x_{i,min}]$, $x_{i,min}$ e $x_{i,m\acute{a}x}$ são os valores máximos e mínimos de x_i , respectivamente.

As performances das diferentes configurações das redes neurais foram comparadas utilizando a taxa de classificação e o RMSE, expresso pela Equação 11, onde n é o número de pares de dados e x_d e x_p são os dados experimentais (também chamados valores desejados) e dados preditos, respectivamente.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_d - x_p)^2} \quad (11) \text{ Eq.}$$

4.5.4.3. Classificação e Predição

Os dados foram classificados de acordo com a taxa de acertos de cada grupo, a qual se refere a uma classificação correta. Caso os valores experimentais fossem identificados de maneira correta pela rede, a mesma correspondia a uma boa taxa de classificação, sendo escolhida, portanto, a rede que apresentou melhor taxa de classificação e menor erro (RMSE).

Para verificar a capacidade de generalização da rede criada, utilizou-se esse modelo para prever o percentual de leite de vaca adicionado ao leite de búfala utilizando-se os dados da etapa de validação do processo de generalização. Os valores preditos pelo modelo e os obtidos experimentalmente foram utilizados para obtenção do coeficiente de correlação e por meio do *Statistical Analysis System (SAS)® Studio* foi observada a sua significância pela teste F ($p < 0,0001$)

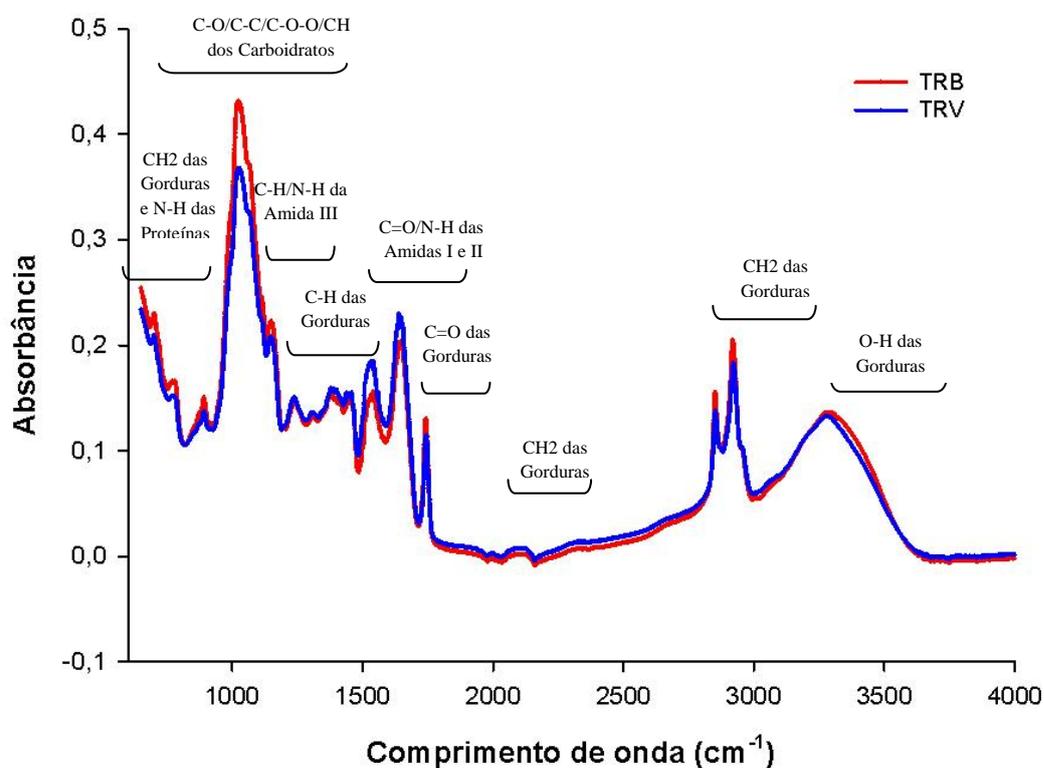
5. RESULTADOS E DISCUSSÃO

5.1. Espectros das amostras referência e das amostras de leite de búfala com adição de leite de vaca

A espectroscopia no Infravermelho FTIR-ATR foi empregada no intuito de realizar comparações entre os perfis espectroscópicos do leite de búfala (TRB), vaca (TRV) e amostras com inclusão de diferentes percentuais de leite bovino ao bubalino.

Foram observadas as características dos espectros das amostras TRB e TRV (Figura 3) e a sobreposição das amostras referência e das amostras de leite de búfala com adição de leite de vaca (Figura 4). Foram identificados 15 picos de absorção importantes cujas absorbâncias máximas, que de acordo com Pappas et al. (2008), Santos et al. (2013), Botelho et al. (2015) são correspondentes às vibrações das ligações dos grupos funcionais presentes no leite, sendo que cada grupo absorve uma frequência característica e corresponde aos níveis de energia da molécula.

Figura 3: Espectros obtidos por espectrômetro FTIR-ATR das amostras do tratamento referência búfala (TRB) e tratamento referência vaca (TRV).



Apesar da semelhança entre os espectros foram notadas algumas diferenças entre as amostras de búfala e vaca evidenciadas pela variação de intensidade das bandas de uma espécie para outra, o que gerou, em grande parte, maiores intensidades de absorção para os picos referentes ao leite bubalino (Figura 3).

De acordo com Nicolaou (2010), o grau de absorção dos picos se correlaciona com a quantidade do componente presente em cada tipo de leite. Considerando que o leite de búfala da região sudoeste da Bahia apresenta maior teor dos seus principais constituintes (4,26-5,8% de gordura, 3,05-4,0% de proteína e 4,27% de lactose) em relação ao leite de vaca (3,8% de gordura, 2,92% de proteínas e 4,14% de lactose) (VENTURINI et al., 2007; CALDEIRA et al., 2010; PIGNATA et al., 2014), essas diferenças podem ter repercutido na absorção dos seus respectivos grupos funcionais.

Comparando os espectros, nota-se que a maioria dos picos gerados se refere às vibrações das ligações de gordura, como os picos da região de comprimento de onda 3278 cm^{-1} , $2919\text{-}2922\text{ cm}^{-1}$, $1742\text{-}1746\text{ cm}^{-1}$ e $2851\text{-}2855\text{ cm}^{-1}$ devido à presença do O-H, CH₂ e C-O, respectivamente, os quais se apresentaram maiores para as amostras da espécie bubalina.

Em menor proporção aparecem os picos relacionados com proteínas e carboidratos e esta diferença de intensidade se aplicou, principalmente, à região do pico $1147\text{-}1149\text{ cm}^{-1}$, referente ao carboidrato que apresentou maior absorção de energia em relação aos demais.

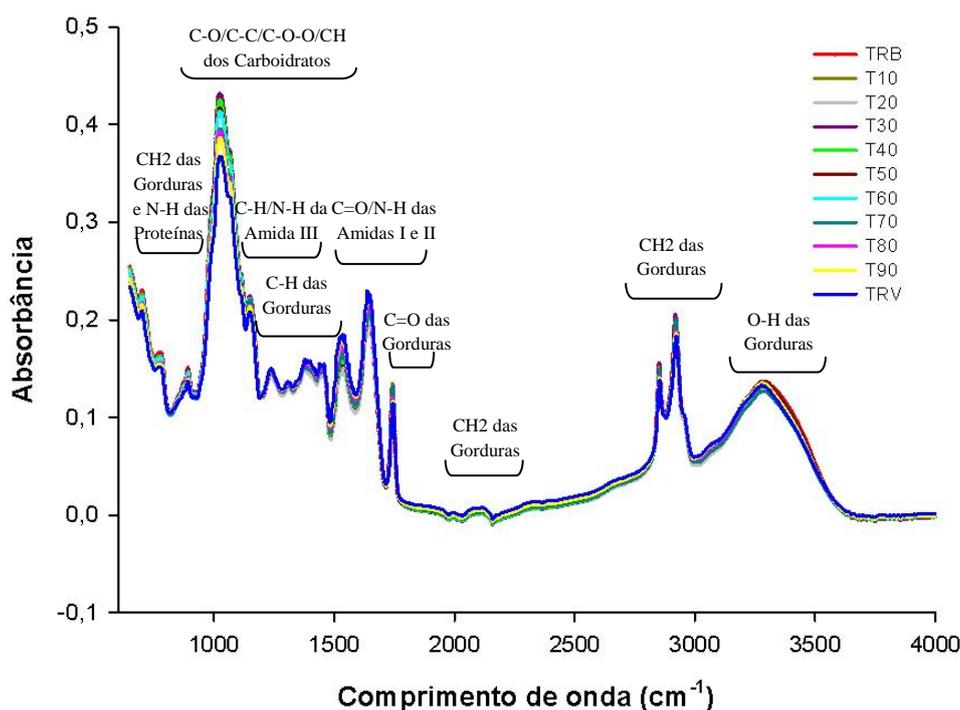
Nicolaou et al. (2010) verificaram por meio do gráfico de sobreposição das bandas derivadas dos espectros dos leites caprino, ovino e bovino as diferenças existentes em suas composições. Observou-se o aparecimento de bandas mais intensas no comprimento de onda de 2927 cm^{-1} para o leite de ovelha devido à presença do CH₂ e C-O, relacionados aos ácidos graxos, justificado pelo maior teor de gordura do leite ovino em relação ao caprino e bovino e, possivelmente também pela diferença no perfil de ácidos graxos.

Os picos referentes às vibrações das ligações das proteínas na região entre $699\text{-}703\text{ cm}^{-1}$, $1240\text{-}125\text{ cm}^{-1}$ também se apresentaram maiores para as amostras TRB, assim como os demais picos referentes aos carboidratos ($1019\text{-}1133\text{ cm}^{-1}$ e $890\text{-}893\text{ cm}^{-1}$) e às gorduras ($768\text{-}780\text{ cm}^{-1}$), confirmando o que a literatura aborda relação entre o grau de absorção dos picos e o teor dos componentes de cada tipo de leite.

Pappas et al. (2008) avaliaram o comportamento dos espectros de amostras de leite de cabra e ovelha desengordurados. Observou-se que, dos onze picos gerados no intervalo entre 1840-950 cm^{-1} de comprimento de onda, pequenas diferenças foram notadas entre as amostras, das quais o leite de cabra apresentou maior intensidade do pico 1745 cm^{-1} em relação ao leite de ovelha atribuída à presença dos carboidratos. Como as amostras foram previamente desengorduradas, possivelmente não foi detectada nessa região a vibração das moléculas de gordura e sim dos carboidratos.

O comportamento dos espectros entre as amostras referência e as adulteradas apresentaram ligeiras diferenças e variações de absorbância pouco perceptíveis tanto no menor (T10) quanto nos maiores níveis de adulteração (T90), sendo evidente a grande aproximação entre os espectros (Figura 4).

Figura 4: Espectros obtidos por espectrômetro FTIR-ATR das amostras do tratamento referência búfala (TRB), tratamento referência vaca (TRV), 10% de leite de vaca em leite de búfala (T10), 20% de leite de vaca em leite de búfala (T20), 30% de leite de vaca em leite de búfala (T30), 40% de leite de vaca em leite de búfala (T40), 50% de leite de vaca em leite de búfala (T50); 60% de leite de vaca em leite de búfala (T60), 70% de leite de vaca em leite de búfala (T70), 80% de leite de vaca em leite de búfala (T80), 90% de leite de vaca em leite de búfala (T90).



À medida que se aumentaram os níveis de adição de leite de vaca ao leite de búfala, as amostras se aproximaram do espectro referente ao TRV sendo que, as amostras em seu menor nível de adição tenderam aos espectros do TRB, os quais apresentaram, em certas regiões, maior absorção de energia.

O pico observado na região 1147-1149 cm^{-1} evidenciou as diferenças existentes entre as amostras com adição de leite de vaca, uma vez que os espectros se justapõem de forma gradativa em função do aumento do nível de adição. O espectro referente ao T10 apresentou comportamento parecido ao espectro do TRB enquanto que o T90, com maior teor de leite de vaca adicionado ao leite de búfala, se aproximou do espectro do TRV. Foi observado que a partir do espectro T50 e T60 os espectros passaram a se aproximar do TRV, ou seja, com o aumento do nível de adição as amostras de leite de búfala adicionadas de leite de vaca passaram a se comportar como o TRV. As demais amostras adicionadas de leite de vaca se apresentaram entre os espectros dos tratamentos referência, acompanhando o desenho espectral de ambos (TRB e TRV).

Dentre os principais picos de absorção e seus respectivos compostos orgânicos foram identificadas as ligações das gorduras em picos entre 3278-1746 cm^{-1} e 1450-1399 cm^{-1} advindos de ligações O-H, CH_2 e C-O, respectivamente. Os grupos funcionais C=O, N-H e C-N originaram picos referentes às amidas I (1645-1251 cm^{-1}), II (1540-1439 cm^{-1}) e III (1240-1251 cm^{-1}). Os carboidratos geraram bandas envolvendo ligações C-O/C-C/C-O-O/CH entre 1147-893 cm^{-1} de absorbância (ZALESKA et al., 2002; CHOUDHURRY et al., 2005; BARBOSA, 2007; PAPPAS et al., 2008; LEI, 2010; NICOLAOU et al., 2010; SUBRAMANIAN et al., 2011; SANTOS et al., 2013a; SANTOS et al., 2013b; JAISWAL et al., 2014).

Jaiswal et al. (2014) detectaram a adulteração de leite de búfala com extrato de soja por meio do FTIR com diferenças claras entre o leite e o extrato de soja em todos os espectros sendo mais visível nas regiões de 1680-1058 cm^{-1} , faixa onde as proteínas de soja absorveram a energia eletromagnética com maior intensidade com marcantes variações de absorção em função do nível de adulteração em diferentes regiões espectrais.

Carvalho et al. (2015) realizaram a detecção de adulteração de leite em pó com adição de soro de leite por meio do infravermelho médio cujas adulterações variaram de 1% a 20%, com 11 níveis de adulteração. Observando os espectros obtidos, os lipídeos foram detectados na faixa de comprimento de onda entre 2924 cm^{-1} e 2853 cm^{-1} e as

vibrações de energia das amidas I e II foram observadas na faixa entre 1611 cm^{-1} e 1539 cm^{-1} , respectivamente, vibrações de energia também observadas por Karoui et al. (2006) (1666-1609 cm^{-1} e 1544 cm^{-1}), resultado semelhante ao observado no presente estudo.

5.2. Análise Multivariada dos dados

Após a coleta dos valores de absorvância e padronização dos dados ($\mu=0$, $S=1$) estes foram submetidos à Análise de Componentes Principais (ACP), Análise de Agrupamento, Análise Discriminante (AD) e Redes Neurais Artificiais (RNA) a fim de investigar a diferença entre os tipos de leite e a diferença entre as amostras referência (TRB e TRV) e as amostras de leite de búfala adicionadas de leite de vaca. Realizou-se a comparação do perfil das amostras TRB e TRV e o comportamento das amostras adicionadas de leite de vaca quando comparadas às amostras de referência.

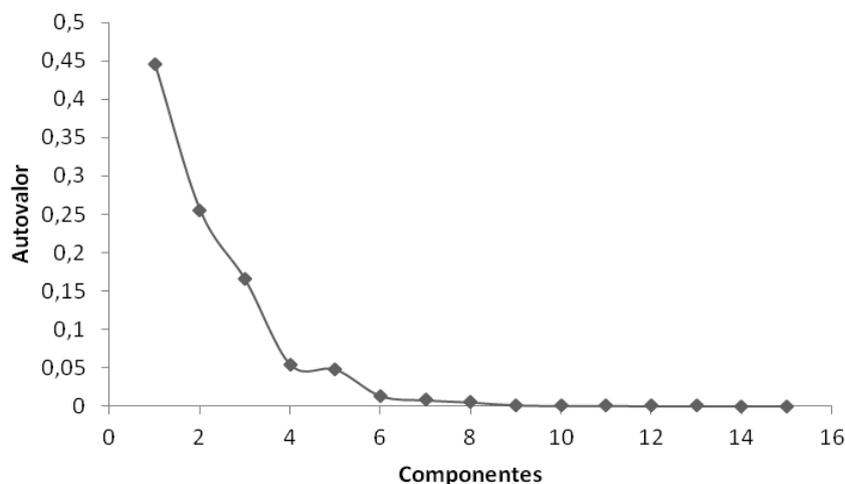
5.2.1. Análise de Componentes Principais (ACP)

A técnica de Análise de Componentes Principais (ACP) teve como propósito reduzir a dimensionalidade dos dados, preservando, simultaneamente, a sua variância e substituindo o conjunto de variáveis correlacionadas em variáveis não correlacionadas, os componentes principais (CP).

O comportamento das amostras TRB e TRV foi avaliado por meio da comparação entre elas. A proporção da variância total com relação à ACP aplicada teve como objetivo verificar e/ou explicar a estrutura dos dados. Observou-se por meio dos critérios de Kaiser, que elege o número de componentes principais a partir de autovalores >1 , do critério dos fatores interpretáveis que considera a soma acumulada dos autovalores $> 70\%$ e da observação do *Scree Plot* a necessidade de 3 componentes principais para a explicação da variabilidade dos dados (Figura 5).

A estabilização da curva ocorreu a partir do quarto autovalor. Dessa forma, os autovalores anteriores ao ponto de estabilização 1, 2 e 3 foram os escolhidos para representar a variância dos dados também segundo este critério (OGILVIE et al., 2008).

Figura 5: *Scree Plot* dos autovalores associados aos componentes principais.

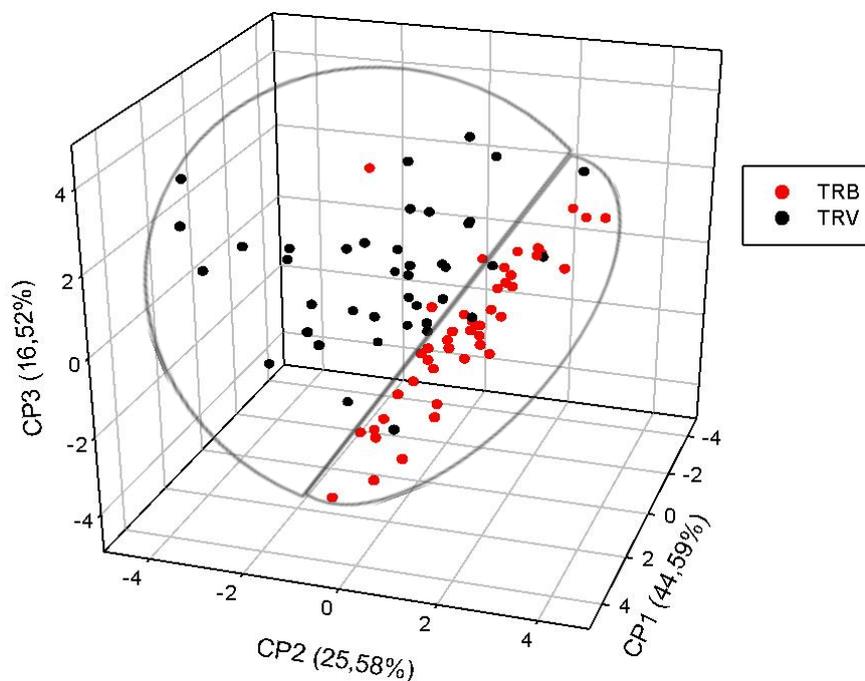


Os três componentes juntos representaram 86,70% da variância dos dados. Com base nos coeficientes de correlação de Pearson e na probabilidade de correlação (0,65-0,7), o primeiro componente principal (CP1- 44,59%) apresentou correlação com as variáveis correspondentes referentes às gorduras (3278 cm^{-1} , 1458 cm^{-1} , 1376 cm^{-1} , 768 cm^{-1}), às proteínas (1647 cm^{-1} , 1241 cm^{-1} e 700 cm^{-1}) e carboidrato (1023 cm^{-1}). O CP2 foi capaz de explicar 25,58% da variabilidade dos dados e se apresentou altamente correlacionável com a maioria das variáveis relacionadas com a gordura (2919 cm^{-1} , 2851 cm^{-1} , 1742 cm^{-1} , 1148 cm^{-1} , 1541 cm^{-1}), enquanto que o CP3 (16,52%) apresentou correlação significativa apenas com o pico 1742 cm^{-1} referente à gordura, e as demais variáveis não obtiveram correlação com o CP3.

A partir da aquisição dos escores dos CP foi realizada a análise por meio do gráfico de dispersão das amostras TRB e TRV no qual foi possível observar a formação de um padrão de agrupamento por tipo de leite com o aparecimento de dois grupos homogêneos com pouca dispersão entre eles (Figura 6).

As amostras TRB se apresentaram de forma mais concentrada enquanto que as TRV apresentaram forma dispersa. No entanto, não houve separação total das amostras, com algumas pertencentes ao grupo TRV se comportando como TRB. O contrário também ocorreu, mas com uma dispersão bem menor, o que provavelmente pode ter ocorrido em função da composição semelhante entre as amostras, fato que dificultou a completa separação das mesmas.

Figura 6: Gráfico de dispersão das amostras do tratamento referência búfala (TRB) e do tratamento referência vaca (TRV) em relação aos componentes principais CP1, CP2 e CP3.



A disposição gráfica dos escores com relação à área que compreende CP2 e CP3 proporciona uma melhor visualização dos grupos formados (Figura 6). O CP2 e CP3, que possuem correlação com a gordura, são capazes de descrever melhor o conjunto de dados e, possivelmente, devido a estes componentes que acontece a separação das espécies. Como o CP1 apresentou correlação com praticamente todas as variáveis, a separação possivelmente aconteceu com base nos componentes principais CP2 e CP3 que são mais específicos e apresentaram correlação com variáveis referentes a apenas um constituinte do leite que é a gordura.

Devido o CP2 apresentar correlação com grande parte das variáveis correspondentes à gordura, graficamente nota-se que o mesmo demonstrou influência na separação dos grupos das espécies. Após a água, a gordura é o componente majoritário do leite de búfalas com teores acima de 5,5%, sendo esta uma característica particular da espécie, independente das condições experimentais. Os elevados percentuais de gordura e proteína do leite de búfala lhe conferem alto valor nutritivo o que distingue

decisivamente a qualidade físico-química do leite bubalino em relação ao leite bovino (FILHO et al., 2014; PIGNATA et al., 2014).

Para verificar a influência dos níveis de adição de leite de vaca ao leite de búfala na separação dos grupos foi realizada a Análise de Componentes Principais com os tratamentos referência TRB, TRV e os níveis de adição T10 ao T90. Nessa etapa, também foram selecionados três CP utilizando os mesmos critérios dos fatores interpretáveis, critérios de Kaiser e o *Scree Plot*. Juntos, os três componentes principais foram capazes de explicar entre 78-86,45% da variabilidade dos dados originais. Com base na aquisição dos escores dos componentes principais foram plotados gráficos em um espaço tridimensional com a comparação entre as amostras de leite de búfala adicionadas de leite de vaca e os tratamentos referência em função do aumento no nível de adição pré-estabelecidos (Figura 7).

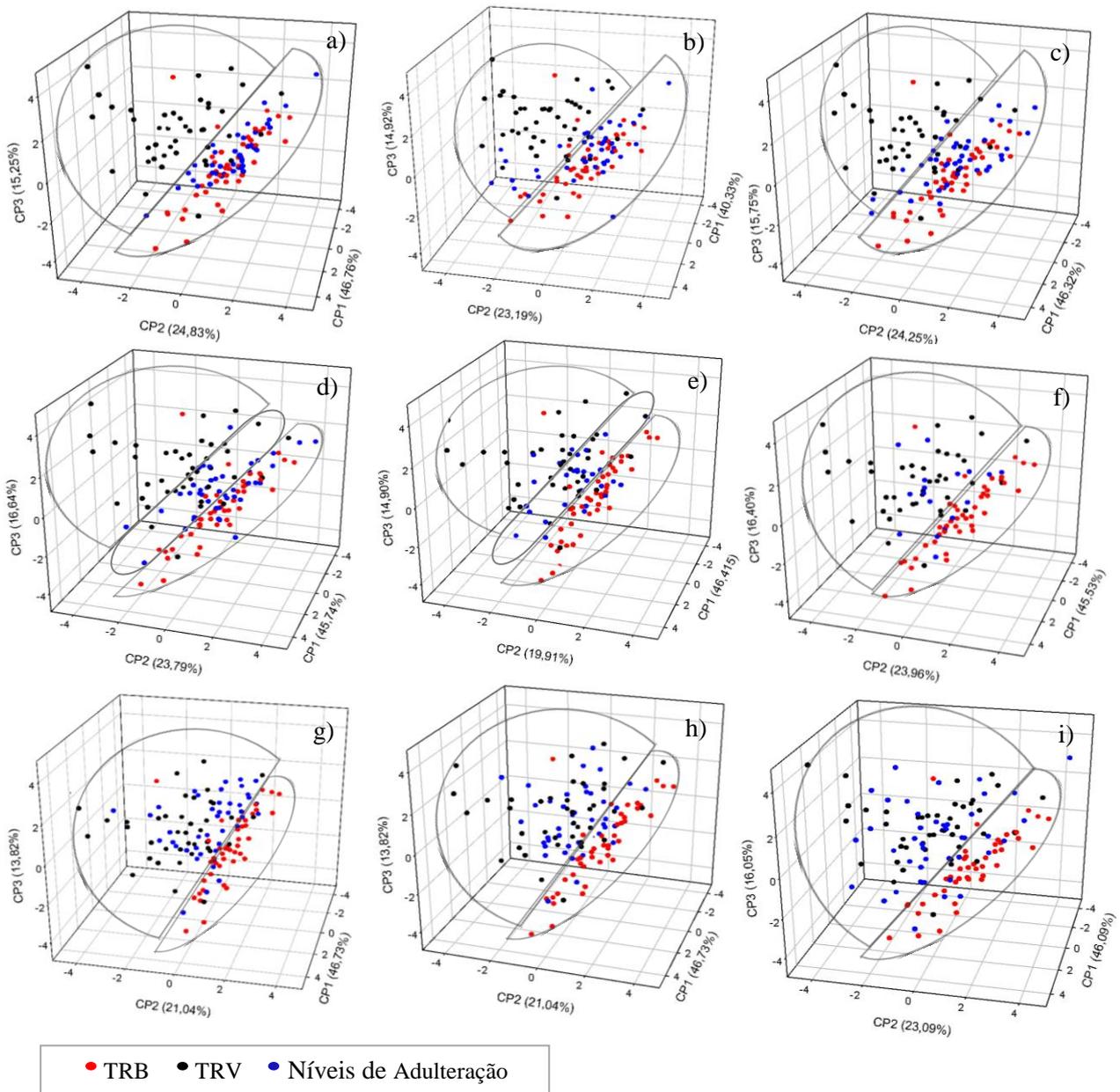
O CP1 obteve boa correlação com a maioria das variáveis referentes aos carboidratos (1023 cm^{-1} , 892 cm^{-1}), proteínas (1647, 1541, 1241, 700 cm^{-1}) e gorduras (3278 cm^{-1} , 1458 cm^{-1} , 1376 cm^{-1} , 768 cm^{-1}), o CP2 apresentou correlação com a gordura (2919 cm^{-1} , 2851 cm^{-1} , 2081 cm^{-1}), enquanto que o CP3 apresentou correlação com algumas variáveis de gordura (2919 cm^{-1} , 2851 cm^{-1}).

As amostras adicionadas de 10%, 20% e 30% de leite de vaca ao leite de búfala apresentaram comportamento semelhante ao TRB, enquanto que o TRV se comportou de forma dispersiva, evidenciando a dificuldade de separação por esta técnica, inclusive nos menores níveis de adição (Figuras 7a, 7b e 7c).

Foi observada a formação de um terceiro grupo localizado entre TRB e TRV referentes a 40% e 50% de adição de leite bovino ao leite bubalino (Figura 7d e 7e). Entretanto, não houve completa separação, pois ainda algumas amostras pertencentes ao grupo T40 e T50 se comportaram como TRB e TRV.

Amostras com 60%, 70%, 80% e 90% de adição de leite bovino ao leite bubalino apresentaram comportamento semelhante ao TRV (Figuras 7f, 7g, 7h e 7i). A partir da análise utilizada observou-se a falta de sensibilidade para a detecção da adição de leite de vaca ao leite de búfala por meio do FTIR-ATR, levando em consideração que as ações fraudulentas com diferentes espécies normalmente acontecem utilizando baixos teores do leite de maior disponibilidade e menor preço para fraudar o leite, possivelmente, mais nobre.

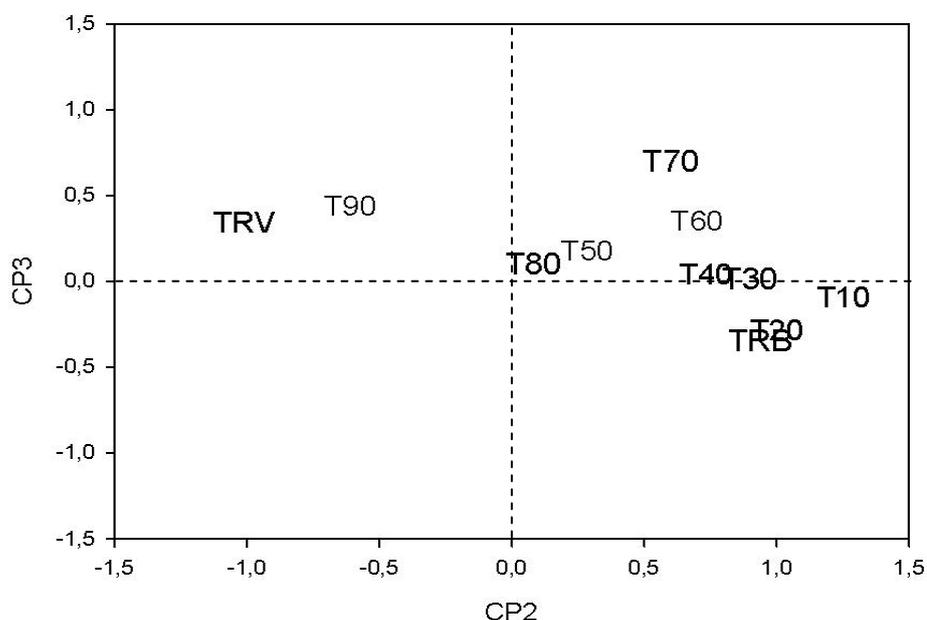
Figura 7: Gráficos de dispersão das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV) e das amostras de leite de búfala adicionadas de leite de vaca (T10 ao T90) em relação aos componentes principais CP1, CP2 e CP3. (a) TRB, TRV e 10% de leite de vaca ao leite de búfala (T10); (b) TRB, TRV e 20% leite de vaca ao leite de búfala (T20); (c) TRB, TRV e 30% leite de vaca ao leite de búfala (T30); (d) TRB, TRV e 40% leite de vaca ao leite de búfala (T40); (e) TRB, TRV e 50% de leite de vaca ao leite de búfala (T50); (f) TRB, TRV e 60% leite de vaca ao leite de búfala (T60); (g) TRB, TRV e 70% de leite de vaca ao leite de búfala (T70); (h) TRB, TRV e 80% de leite de vaca ao leite de búfala (T80); (i) TRB, TRV e 90% de leite de vaca ao leite de búfala (T90).



Por meio da análise da média dos escores de todos os tratamentos estudados em função do CP2 e CP3 foi possível confirmar que estes componentes de fato exerceram influência na separação dos dados em função dos níveis de adição (Figura 8). O CP2 foi capaz de separar o TRV e o T90 das demais amostras estudadas, enquanto que o CP3 foi determinante no processo de separação do grupo das amostras TRB, T10, T20, apresentadas no quadrante inferior direito da figura, do grupo formado por T30, T40, T50, T60, T70 e T80.

Percebe-se a separação das amostras TRB e TRV em quadrantes diferentes e a divisão de um grupo com menores níveis de adição (T10, T20) que se aproximou ao tratamento referência de búfala (TRB), as amostras do T30 e T40 se localizaram entre os quadrantes inferior e superior podendo ser considerados níveis intermediários de adição e o grupo de maiores níveis de adição de leite de vaca ao leite de búfala que corresponderam às amostras T40, T50, T60, T70 e T80, enquanto que o T90 apresentou características muito similares ao tratamento referência de vaca (TRV).

Figura 8: Gráfico de médias dos escores das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV), T10 (10% de leite de vaca em leite de búfala); T20 (20% leite de vaca em leite de búfala); T30 (30% leite de vaca em leite de búfala); T40 (40% leite de vaca em leite de búfala); T50 (50% de leite de vaca em leite de búfala); T60 (60% leite de vaca em leite de búfala); T70 (70% de leite de vaca em leite de búfala); T80 (80% de leite de vaca em leite de búfala); T90 (90% de leite de vaca em leite de búfala) em relação ao CP2 e CP3.



Santos et al. (2013a) e Santos et al. (2013b) empregaram a Análise de Componentes Principais usando espectros obtidos por meio da espectroscopia no infravermelho de amostras controle de leite bovino e adulteradas com peróxido de hidrogênio, leite sintético (emulsão de óleos vegetais, detergentes e uréia), soro de leite e uréia. Na apresentação espacial dos dados os autores indicaram a separação de grupos bem definidos do leite bovino e das amostras adulteradas permitindo a sua discriminação total em função dos tipos de adulterantes, os quais possuem componentes químicos diferentes da composição original do leite, o que pode facilitar a separação. No caso de espécies com a mesma composição qualitativa, como é o caso do presente trabalho, a análise fica mais complexa e difícil.

Jaiswal et al. (2014) ao utilizarem a ACP demonstraram o agrupamento claro de amostras de leite de búfala em relação aos níveis de adulteração com extrato de soja. Mesmo com o aparecimento dos grupos com menores níveis de adulteração (2%, 4%, 5% de extrato de soja em leite de búfala) próximos uns dos outros, houve uma discriminação precisa já que apresentaram grandes diferenças em suas composições e propriedades bioquímicas.

5.2.2. Análise de Agrupamento (AA) ou *Cluster Analysis*

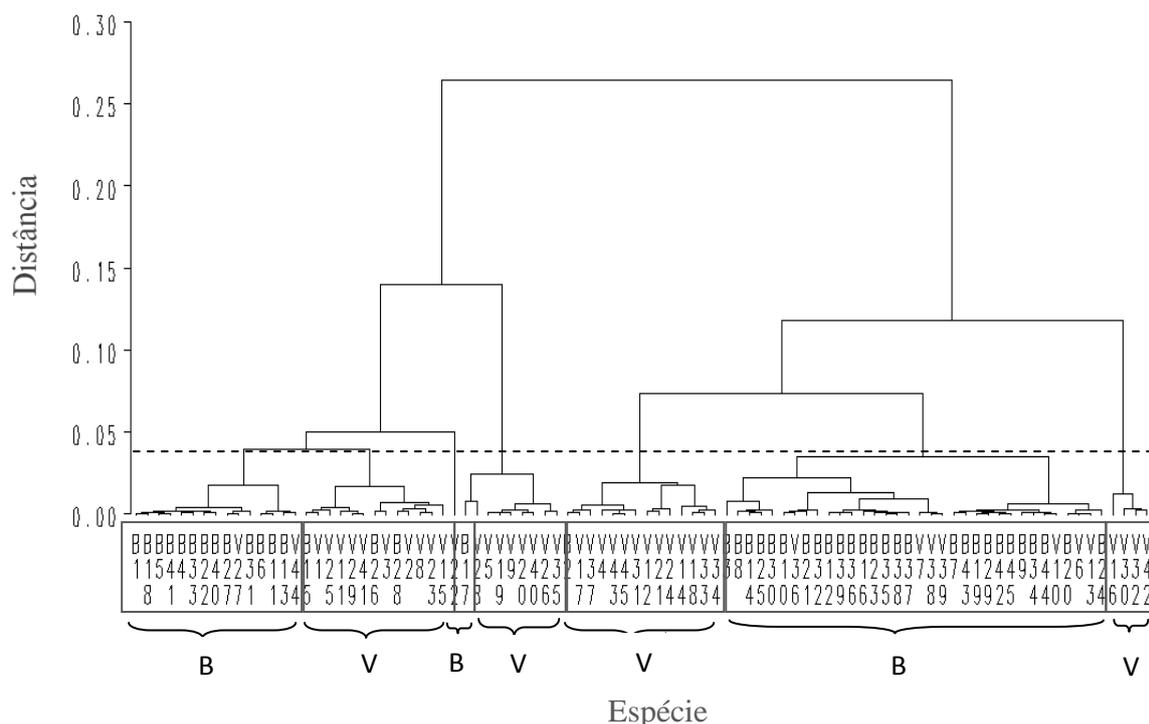
Por meio da Análise de Agrupamento pôde-se verificar a discriminação das amostras TRB e TRV e os grupos formados foram evidenciados pela plotagem do dendograma, gráfico de agrupamento (Figura 9). Os algoritmos de agrupamento não apresentaram solução exata para a determinação do número ideal de grupos, sendo uma decisão subjetiva, realizada de acordo com o conhecimento prévio sobre a estrutura dos dados e o número de grupos desejados, por meio da análise visual.

O eixo vertical do gráfico corresponde ao nível de similaridade entre as amostras, onde quanto maior a altura de cada grupo maior o erro que distancia um grupo do outro, já o eixo horizontal representa as amostras estudadas. O conjunto de dados foi agrupado conforme o grau de semelhança entre eles apresentado.

O agrupamento foi baseado no corte referente à distância entre 0 e 0,05, correspondendo à formação de sete grupos. Dos sete grupos formados, foram obtidos 4 grupos referentes ao TRV (V) e 3 ao TRB (B), verificando-se ainda a presença de algumas amostras tanto do TRB quanto do TRV em ambos os grupos formados, o que,

possivelmente ocorreu devido às semelhanças entre os picos de absorvância obtidos pelo FTIR-ATR dos leites das duas espécies estudadas. Apesar do método ter realizado o agrupamento, este fato dificultou a separação de 100% das amostras.

Figura 9. Dendograma das amostras do tratamento referência búfala (TRB – B), do tratamento referência vaca (TRV – V).



Não foi possível identificar a formação de grupos por meio da análise de agrupamento das amostras de leite de búfala adicionadas de leite de vaca (T10 ao T90) devido à impossibilidade de visualização dos mesmos. Muitas amostras foram confundidas com amostras semelhantes impossibilitando a identificação correta dos grupos, o que não é interessante para o estudo, demonstrando que o método não foi capaz de agrupar as amostras de modo satisfatório para a possível identificação de fraude.

Pappas et al. (2008) descreveram em seu trabalho o desenvolvimento de um método para a identificação e diferenciação de leites de cabra e de ovelha por meio da análise em FTIR por reflexão difusa. Oitenta e sete espectros de ambas as amostras foram obtidos e os valores de absorções utilizados para a diferenciação das amostras de leite. Foi utilizado o método de Análise de Agrupamento Hierárquico para a separação

entre ambos os grupos e para a mensuração das distâncias entre os grupos foram utilizadas a Correlação de Pearson, já que se tratava de dois grupos de amostras. Os autores indicaram por meio do dendograma uma diferenciação entre o conjunto de amostras de cabra e ovelha, resultado semelhante ao do presente estudo.

5.2.3. Análise Discriminante (AD)

A Análise Discriminante (AD) foi utilizada para separar e classificar as amostras de leite estudadas em cinco classes: TRB, TRV, amostras com baixo nível de adição de leite de vaca ao leite de búfala (entre T10 a T20), médio nível de adição de leite de vaca ao leite de búfala (T30 a T40) e alto nível de adição de leite de vaca ao leite de búfala (T50 a T90) por meio da determinação da taxa de classificação, que corresponde ao somatório dos acertos de cada grupo.

A primeira etapa, consistiu do treinamento, parte exploratória da análise para verificar características capazes de serem utilizadas para alocar as amostras nos cinco grupos previamente definidos por meio do uso da função linear discriminante. Cinco funções discriminantes foram obtidas como combinações lineares das variáveis utilizadas no estudo (Equações 12, 13, 14, 15 e 16) correspondentes a cada grupo estudo, as quais maximizaram a razão entre as classes e minimizaram a dispersão dentro de cada grupo (FISHER, 1936).

$$Y_B = -163.94 - 12.12 * \gamma_{3278} + 1001 * \gamma_{2919} - 278.91 * \gamma_{2851} + 13.86 * \gamma_{2081} + 488.40 * \gamma_{1742} - 1415 * \gamma_{1647} + 1536 * \gamma_{1541} + 17.75 * \gamma_{1458} - 337.39 * \gamma_{1376} - 615.70 * \gamma_{1241} + 508.73 * \gamma_{1148} + 190.27 * \gamma_{1023} + 784.54 * \gamma_{892} - 12.37 * \gamma_{768} - 328.44 * \gamma_{700}$$

(12) Eq.

$$Y_V = -181.45 - 8.32 * \gamma_{3278} + 1177 * \gamma_{2919} - 364.01 * \gamma_{2851} + 26.37 * \gamma_{2081} + 606.25 * \gamma_{1742} - 1512 * \gamma_{1647} + 1611 * \gamma_{1541} - 32.75 * \gamma_{1458} - 339.50 * \gamma_{1376} - 555.85 * \gamma_{1241} + 472.86 * \gamma_{1148} + 194.50 * \gamma_{1023} + 825.74 * \gamma_{892} - 11.97 * \gamma_{768} - 351.79 * \gamma_{700}$$

(13) Eq.

$$Y_C = -164.31 - 18.88 * \gamma_{3278} + 1015 * \gamma_{2919} - 288.99 * \gamma_{2851} + 13.35 * \gamma_{2081} + 481.86 * \gamma_{1742} - 1461 * \gamma_{1647} + 1553 * \gamma_{1541} - 31.97 * \gamma_{1458} - 329.13 * \gamma_{1376} - 632.77 * \gamma_{1241} + 523.49 * \gamma_{1148} + 196.41 * \gamma_{1023} + 785.70 * \gamma_{892} - 5.64 * \gamma_{768} - 335.37 * \gamma_{700}$$

(14) Eq.

$$\begin{aligned}
Y_D = & -170.95 - 9.63 * \gamma_{3278} + 949.94 * \gamma_{2919} - 262.07 * \gamma_{2851} + 8.90 * \gamma_{2081} + 510.13 * \\
& \gamma_{1742} - 1494 * \gamma_{1647} + 1592 * \gamma_{1541} - 4.47 * \gamma_{1458} - 361.79 * \gamma_{1376} - 585.07 * \gamma_{1241} + \\
& 494.14 * \gamma_{1148} + 204.77 * \gamma_{1023} + 793.35 * \gamma_{892} - 12.35 * \gamma_{768} - 346.47 * \gamma_{700}
\end{aligned}$$

(15) Eq.

$$\begin{aligned}
Y_E = & -180.23 - 6.63 * \gamma_{3278} + 1178 * \gamma_{2919} - 340.28 * \gamma_{2851} + 12.27 * \gamma_{2081} + 540.38 * \\
& \gamma_{1742} - 1519 * \gamma_{1647} + 1638 * \gamma_{1541} - 6.09 * \gamma_{1458} - 357.05 * \gamma_{1376} - 617.10 * \gamma_{1241} + \\
& 513.26 * \gamma_{1148} + 195.47 * \gamma_{1023} + 846.60 * \gamma_{892} - 12.37 * \gamma_{768} - 357.95 * \gamma_{700}
\end{aligned}$$

(16) Eq.

Onde: Y_B : Função linear discriminante do grupo TRB; Y_V : Função linear discriminante do grupo TRV; Y_C : Função linear discriminante do grupo com baixo nível de adição de leite de vaca ao leite de búfala (T10 a T20); Y_D : Função linear discriminante do grupo com médio nível de adição de leite de vaca ao leite de búfala (T30 a T40) e Y_E : Função linear discriminante do grupo com alto nível de adição de leite de vaca ao leite de búfala (T50 a T90).

Na etapa de validação, utilizada para avaliar o desempenho da classificação, os dados foram aplicados para que a partir das equações obtidas na fase anterior os novos dados fossem classificados em diferentes grupos.

Na etapa de treinamento, parte exploratória da análise, as amostras TRB e TRV apresentaram o percentual de acerto de 52,0% e 72,0%, respectivamente, e na etapa de validação 44,44% e 33,33%, respectivamente evidenciando que o método foi capaz de diferenciar melhor os tratamentos referência na etapa de treinamento quando comparado com a etapa de validação. Entretanto, em ambas as etapas foi obtida uma baixa capacidade de discriminação entre as amostras de leite de vaca e leite de búfala (Tabela 3).

Para a classificação dos grupos contendo amostras com baixo nível de adição (T10 e T20), médio nível (T30 a T40) e alto nível (T50 a T90) foram apresentados 61,11%, 58,33%, 52,78% de acertos na etapa de treinamento, respectivamente, e na fase de validação os três grupos obtiveram taxas de classificação iguais, 44,44%, as quais evidenciaram a dificuldade para a separação destes grupos.

Para os grupos de amostras de leite de búfala adicionadas de leite de vaca, o ideal é que estas sejam classificadas como elas mesmas e que sejam obtidas baixas taxas de erro. Neste trabalho o grupo contendo amostras com baixo nível de adição, 33,33% das amostras, foram confundidas com o TRB na etapa de validação e o mesmo número de amostras foi confundido com amostras pertencentes ao grupo T30 e T40, valores considerados altos e não satisfatórios para o processo de detecção de adulterações.

Tabela 3: Número de observações e o percentual de classificação por grupo no treinamento (T)/validação (V) das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV) e amostras com baixo nível de adição de leite de vaca ao leite de búfala (T10 a T20), médio nível de adição de leite de vaca ao leite de búfala (T30 a T40) e alto nível de adição de leite de vaca ao leite de búfala (T50 a T90).

Grupos	TRB (T/V)	TRV (T/V)	Baixo Nível (T10-T20) (T/V)	Nível Médio (T30-T40) (T/V)	Alto Nível (T50-T90) (T/V)	Total (T/V)
TRB	19/4	3/1	11/3	3/1	0/0	36/9
	52,00/44,44	8,33/11,11	30,56/33,33	8,33/11,11	0,00/0,00	100,00
TRV	1/0	26/3	0/0	2/2	7/4	36/9
	2,78/0,00	72,00/33,33	0,00/0,00	5,56/22,22	19,44/44,44	100,00
Baixo Nível (T10-T20)	10/3	0/0	22/4	3/2	1/0	36/9
	27,78/33,33	0,00/0,00	61,11/44,44	8,33/22,22	2,78/0,00	100,00
Nível Médio (T30-T40)	5/0	0/0	6/3	21/4	4/2	36/9
	13,89/0,00	0/0,00	16,67/33,33	58,33/44,44	11,11/22,22	100,00
Nível Alto (T50-T90)	3/0	6/2	3/2	5/1	19/4	36/9
	8,33/0,00	16,67/22,22	8,33/22,22	13,89/11,11	52,78/44,44	100,00

Pappas et al. (2008) identificaram e diferenciaram amostras de leite de cabra e ovelha desengorduradas utilizando a AD por meio de um gráfico de dispersão das distâncias obtidas na análise discriminante entre as amostras, sendo observada uma discriminação satisfatória nas amostras de leite. Enquanto que no presente estudo foram analisadas as taxas de classificação de cada grupo para a realização da identificação e separação das mesmas.

A análise discriminante não foi capaz de identificar e diferenciar os grupos estudados. Isto provavelmente ocorreu não pelo fato da má qualidade dos dados, mas sim devido às semelhanças entre as amostras.

5.2.4. Redes Neurais Artificiais (RNA)

A seleção da rede neural de melhor arquitetura foi realizada utilizando como critério a escolha do número de neurônios das camadas ocultas que apresentaram de forma conjunta um dos menores valores de RMSE na etapa de validação e taxa máxima de classificação dos grupos estudados e total.

5.2.4.1. Classificação

Obteve-se a melhor configuração da RNA (Tabela 4) com quinze neurônios na camada de entrada, cinquenta neurônios na primeira camada oculta, vinte neurônios na segunda camada oculta e 5 neurônios na camada de saída apresentando um RMSE de 0,23, taxa de classificação total de 95,55% e 100% de classificação correta para as amostras dos grupos TRB, TRV, com nível médio de adição de leite de vaca ao leite de búfala (T30 a T40) e com alto nível de adição (T50 a T90). Outras configurações testadas também mostraram baixo RMSE e alta taxa de classificação, porém foi escolhida a configuração mais simples de neurônios, uma vez que o seu processamento requer menor esforço computacional.

Tabela 4: Configuração das Redes Neurais Artificiais testadas, Erro Quadrático Médio (RMSE), e taxas de classificação (%) do tratamento referência de búfala (TRB), do tratamento referência de vaca (TRV), do grupo de amostras com baixo nível de adição de leite de vaca ao leite de búfala (T10 a T20), com médio nível de adição de leite de vaca ao leite de búfala (T30 a T40), com alto nível de adição de leite de vaca ao leite de búfala (T50 a T90) e taxa de classificação total (T).

NCE	1 ^a CO	2 ^a CO	NCS	RMSE	Taxa de Classificação (%)					
					TRB	TRV	Baixo	Médio	Alto	T
15	50	50	5	0,32	100	100	55,55	88,88	100	88,88
15	50	40	5	0,17	88,88	100	100	88,88	100	95,55
15	50	30	5	0,20	100	100	77,77	88,88	100	94,33
15	50	20	5	0,23	100	100	77,77	100	100	95,55
15	50	10	5	0,30	88,88	88,88	88,88	88,88	88,88	91,11
15	50	0	5	0,47	88,88	100	55,55	77,77	77,77	80,00
15	40	0	5	0,50	55,55	88,88	77,77	77,77	100	80,00
15	30	0	5	0,54	66,66	77,77	77,77	77,77	77,77	75,55
15	20	0	5	0,57	88,88	100	44,44	88,88	77,77	80,00
15	10	0	5	0,57	77,77	100	33,33	77,77	66,66	71,1

NCE=Neurônios da camada de entrada/ CO = Camada oculta/ NCS=Neurônios da camada de saída/
RMSE= *Root Mean Square Errors* ou Erro quadrático médio.

No presente estudo, a RNA foi capaz de discriminar corretamente as amostras em seus respectivos grupos de acordo com suas características e ainda com boas taxas de classificação, fato que não ocorreu completamente nas outras análises estatísticas multivariadas realizadas anteriormente como a ACP, AA e AD, que não foram capazes de separar tão bem as amostras e/ou as classificaram como sendo pertencente a outros semelhantes.

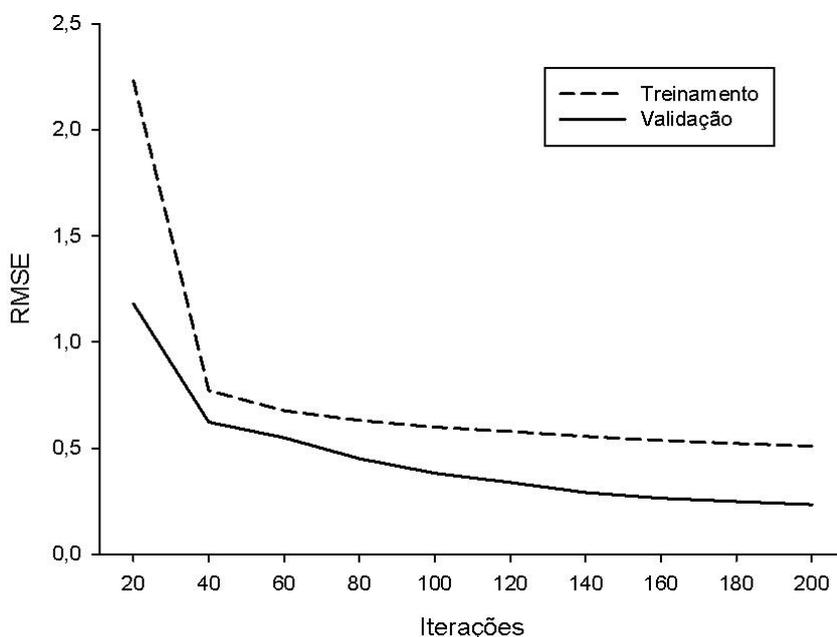
De acordo com Cybenko (1988), uma camada oculta é suficiente para aproximar qualquer função contínua, enquanto duas camadas escondidas são suficientes para aproximar qualquer função matemática. Pode ser observado que, em geral, quando apenas uma camada escondida é utilizada, há um valor de RMSE mais elevado, como o que ocorreu nas redes com 20 e 10 neurônios na camada oculta para o grupo de baixo nível de adição, apresentando $RMSE=0,57$ e taxas de classificação baixas (44,44% e 33,33%, respectivamente).

Nota-se que o grupo contendo os tratamentos T10 e T20 (baixo nível de adição) apresentou menores taxas de classificação para a grande maioria das arquiteturas testadas em comparação com os demais grupos, sendo possível prever o quão é complexa a separação de amostras adulteradas, principalmente com baixos teores de leite de vaca em leite de búfala.

Foi monitorado o RMSE em função do número de iterações em 200 ciclos, para a etapa de treinamento e de validação da rede que apresentou melhor arquitetura (50 e 20 neurônios na primeira e segunda camada oculta, respectivamente) (Figura 10).

O erro do treinamento normalmente se apresenta maior que o de validação na fase inicial e ambos diminuem ligeiramente durante as primeiras iterações. Observou-se que a partir de 40 iterações o erro tende a reduzir numa faixa menor e se torna constante, sendo capaz de discriminar corretamente as amostras em seus verdadeiros grupos. No presente trabalho, embora tenham sido utilizadas 200 iterações só foram necessárias 40 iterações para o processo de aprendizagem.

Figura 10. Gráfico do erro quadrático médio (RMSE) em função do número de iterações nas etapas de treinamento e validação da rede neural de melhor arquitetura.



Os valores de RMSE e de taxa de classificação obtidos no presente trabalho estão de acordo com outros estudos. Valente et al. (2014) empregaram as Redes Neurais Artificiais para classificar amostras de leite adulteradas quanto à adição de soro de queijo com diferentes concentrações de soro (0, 1, 5, 10 e 20%). Dos 167 ensaios realizados ocorreu erro de classificação total de 4,8%. Para o leite que não foi adulterado (0%) não ocorreu erro de classificação e para amostras adulteradas houve classificação errônea, mas numa porcentagem considerada baixa, sendo que, a melhor arquitetura testada apresentou 95% de acertos na classificação.

Barile et al. (2006) encontraram valores de RMSE inferiores a 0,08 e taxa de classificação de 93% para a etapa de validação para uma rede com configuração 9-16-2 (9 neurônios de entrada e 16 na primeira camada oculta e 2 saídas) usados para classificar um tipo de queijo italiano com base em propriedades físico-químicas.

Nazário et al. (2009) projetaram uma rede neural para detectar adulteração de amostras de leite por adição de água na faixa de 1 a 60% por meio de ultra-som. Para a detecção da adulteração foram obtidos resultados com boa capacidade de classificação, sendo que as configurações contendo 40 e 52 neurônios obtiveram o maior potencial de classificação, obtendo 95,7% e 96,3% e erro de 0,43 e 0,37, respectivamente.

Balestrieri et al (2001), realizaram na Itália a classificação de amostras de leite de cabra, ovelha e búfalo por meio da RNA de acordo com a composição de cada espécie determinada pelo método oficial de análise química do país. Foram utilizados 21 neurônios de entrada, 4 neurônios de saída e 1 camada oculta com 10 neurônios tanto para treinamento quanto para validação e apresentou 96% de classificação correta e erro de 0,199.

5.2.4.2. Predição

Os parâmetros da RNA utilizados para a predição demonstraram que a configuração da Rede Neural que apresentou melhor estrutura para a predição dos teores de leite de vaca utilizados para a adição ao leite de búfala incluiu 2 camadas ocultas, com 50 e 20 neurônios, respectivamente, a qual apresentaram boa capacidade de generalização, sendo os valores RMSE e coeficiente de correlação considerados satisfatórios, de 0,23 e 0,971 (Tabela 5).

O número ótimo de camadas ocultas e de seus neurônios foram selecionados utilizando o critério de tentativa e erro até encontrar a rede com menor valor de RMSE e maior coeficiente de correlação.

Tabela 5: Configurações das redes neurais testadas para a predição e os coeficientes preditivos.

NCE	1ª CO	2ª CO	NCS	RMSE	Correlação (R ²)
15	50	50	5	0,32	0,911
15	50	40	5	0,17	0,903
15	50	30	5	0,20	0,920
15	50	20	5	0,23	0,971
15	50	10	5	0,30	0,803
15	50	0	5	0,47	0,701
15	40	0	5	0,50	0,702
15	30	0	5	0,54	0,660
15	20	0	5	0,57	0,754
15	10	0	5	0,57	0,708

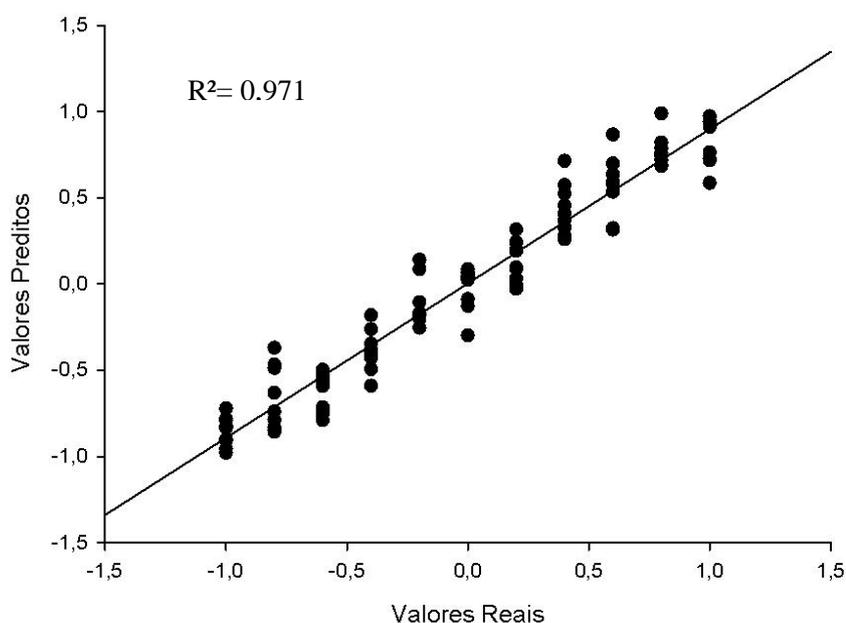
NCE=Neurônios da camada de entrada/ CO = Camada oculta/NCS=Neurônios da camada de saída
 RMSE= *Root Mean Square Errors* ou Erro quadrático médio.

A capacidade preditiva da Rede Neural Artificial gerada a partir do conjunto de dados também foi observada por meio da comparação entre os valores experimentais e/ou reais referentes aos níveis de adição e os valores preditos pela rede que obteve

menor RMSE e maior correlação (Figura 11).

Observa-se com o gráfico que as RNA demonstraram capacidade preditiva satisfatória, apresentando os dados de forma ascendente e concentrados em torno da linha de tendência, apresentando valores experimentais e preditos altamente correlacionados ($R^2 = 0,971$), sendo que esta correlação foi significativa ($p < 0,001$) segundo o teste F.

Figura 11. Valores reais versus valores preditos pela rede das amostras do tratamento referência búfala (TRB), do tratamento referência vaca (TRV), e das amostras T10 (10% de leite de vaca em leite de búfala); T20 (20% leite de vaca em leite de búfala); T30 (30% leite de vaca em leite de búfala); T40 (40% leite de vaca em leite de búfala); T50 (50% de leite de vaca em leite de búfala); T60 (60% leite de vaca em leite de búfala); T70 (70% de leite de vaca em leite de búfala); T80 (80% de leite de vaca em leite de búfala) e T90 (90% de leite de vaca em leite de búfala).



Goyal et al. (2012) utilizaram as RNA para prever o tempo de vida útil do queijo processado correlacionando nitrogênio solúvel, pH, contagem de leveduras e de formas esporuladas com os escores de análise sensorial realizada. Os resultados revelaram uma alta correlação entre os dados experimentais e os valores preditos ($R^2 = 0,999$), e baixa

taxa de erro (RMSE = 0,0009), demonstrando que o modelo de RNA desenvolvido foi capaz de prever os valores estudados com excelente desempenho.

Em um estudo desenvolvido na Índia foi realizada a predição da quantidade de leite produzido por vacas leiteiras utilizando as redes neurais artificiais. A arquitetura da rede selecionada possuía 4 neurônios de entrada, 2 camadas ocultas e um neurônio de saída (4-2-1), a qual foi capaz de prever o total de leite com erro mínimo e uma correlação máxima, apresentando coeficiente próximo de 1 (0,986) (CHATURVEDI et al., 2013)

Assim, a RNA desenvolvida no presente trabalho pode ser utilizada para estimar os níveis de adição de outras amostras de leite de búfala adicionado de leite de vaca que não foram utilizadas neste estudo, demonstrando que a RNA gerada apresenta uma boa capacidade de generalização.

6. CONCLUSÕES

A detecção da adição do leite de vaca ao leite de búfala utilizando o FTIR-ATR neste trabalho só foi possível por meio da aplicação das RNA, entretanto, as análises quimiométricas ACP, AA e RNA foram capazes de identificar e diferenciar as amostras TRB e TRV.

A Análise de Componentes Principais (ACP) possibilitou a diferenciação dos tratamentos TRB e TRV e a detecção da adição apenas a partir de 40% de leite de vaca ao leite de búfala, ressaltando a importância das variáveis relacionadas com a gordura para a separação e diferenciação das amostras.

As amostras TRB e TRV foram agrupadas por meio da Análise de Agrupamento, porém não apresentaram a formação de grupos de maneira satisfatória das amostras de leite de búfala com adição de leite de vaca.

A análise discriminante (AD) apresentou baixas taxas de classificação na fase de validação tanto para a identificação e diferenciação das amostras TRB e TRV quanto para a separação dos grupos de amostras de leite de búfala adicionadas de leite de vaca, não sendo satisfatório para detecção de inclusão de leite de espécies produtoras diferentes.

As RNA se apresentaram mais eficazes na classificação das amostras em relação às demais análises multivariadas. Foram obtidas taxas de classificação total de 95,55% e de 100% para as amostras TRB, TRV e a maioria dos grupos de amostras de leite de vaca com adição de leite de búfala além da boa capacidade de predição, proporcionando correlações e erros satisfatórios.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ABER, S.; AMANI-GHADIM, A. R.; MIRZAJANI, V. Removal of Cr (VI) from polluted solutions by electrocoagulation: Modeling of experimental results using artificial neural network. **Journal of Hazardous Materials**, v. 171, n. 1-3, p. 484-490, 2009.

ALMEIDA, A. L. Avaliação do método de determinação da enzima xantino oxidase para identificação da adulteração de leite de cabra por adição de leite de vaca. **Revista do Instituto de Laticínios Cândido Tostes**, v. 56, n. 321, p. 55-62, 2001.

AMARAL, F. R.; ESCRIVÃO, S. C. Aspectos relacionados à búfala leiteira. **Revista Brasileira de Reprodução Animal**, v. 29, n. 2, p.111-117, 2005.

ANDRIGHETTO, C. Cadeia produtiva do leite de búfala. In: Simpósio da Cadeia Produtiva da Bubalinocultura, 2011, Botucatu. **Anais...** Botucatu, 2011.

ANDUEZA, D.; AGABRIEL, C.; CONSTANT, I.; LUCAS, A.; MARTIN, B. Using visible or near infrared spectroscopy (NIRS) on cheese to authenticate cow feeding regimes. **Food Chemistry**, v. 141, p. 209-214, 2013.

ARAÚJO, T. P. M.; RANGEL, A. H. N.; SOARES, A. D.; LIMA, T. C. C.; JUNIOR, D. M. L.; NOVAES, L. P. Influência das estações do ano sobre a composição do leite de búfalas mantido em tanque de resfriamento. **Revista Agropecuária Científica no Semi-Árido**, v.07, n. 01, p. 01-05, 2011.

BALESTRIERI, F.; DAMIANI, F.; MARINI, D. Artificial neural networks to classify some dairy products. **Journal of Commodity Science**, v. 40, n.1, p. 17-31, 2001.

BARBOSA, Luiz Cláudio de Almeida. **Espectroscopia no infravermelho na caracterização de compostos orgânicos**. Viçosa: Ed. UFV, 2007.

BARILE, D.; COÏSSON, J. D.; ARLORIO, M.; RINALDI, M. Identification of production area of Ossolano Italian cheese with chemometric complex approach. **Food Control**, v. 17, p. 197–206, 2006.

BAS, D.; BOYACI, İ. H. Modeling and optimization II: Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. **Journal of Food Engineering**, v. 78, n. 3, p. 846–854. 2007.

BORGOGNONE, M. G.; BUSSI, J.; HOUGH, G. Principal component analysis in sensory analysis: covariance or correlation matrix? **Food Quality and Preference**, v.12, n. 5-7, p. 323–326, 2001.

BORIN, A.; FERRÃO, M. F.; MELLO, C.; MARETTO, D. A.; POPPI, R. J. Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered Milk. **Analytica Chimica Acta**, v. 579, n. 1, p. 25-32, 2006.

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**, v. 181, p. 31–37, 2015.

BRASIL. Decreto nº 30.691, de 29 de março de 1952. Regulamento de Inspeção Industrial e Sanitária de Produtos de Origem Animal. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 07 jul. 2007, Seção 1.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Dados de rebanho bovino e bubalino no Brasil -2015, Brasília, DF, Dezembro de 2015.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa nº 62 de 2011. Regulamento Técnico de Produção, Identidade e Qualidade de Leite. **Diário Oficial da União**, Brasília, DF, 29 de dezembro de 2011. 24 p.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. RIISPOA. Regulamento de Inspeção Industrial e Sanitária de Produtos de Origem Animal. **Diário Oficial da União**, Brasília, DF, 1997. 154 p.

CALDEIRA, L. A.; FERRÃO, S. P. B.; FERNANDES, S. A. A.; MAGNAVITA, A. P. A.; SANTOS, T. D. R. Desenvolvimento de bebida láctea sabor morango utilizando diferentes níveis de iogurte e soro lácteo obtidos com leite de búfala. **Ciência Rural**, v.40, n.10, p.2193-2198, 2010.

CAPUANO, E.; RADEMAKER, J.; VAN DEN BIJGAART, H.; VAN RUTH, S. M. Verification of fresh grass feeding, pasture grazing and organic farming by FTIR spectroscopy analysis of bovine milk. **Food Research International**, v. 60, n. 1, p. 59-65, 2014.

CARVALHO, B. M. A.; CARVALHO, L. M.; COIMBRA, J. S. R.; MININ, L. A.; BARCELLOS, E. S.; JUNIOR, W. F. S.; DETMANN, E.; CARVALHO, G. G. P. Rapid detection of whey in milk powder samples by spectrophotometric and multivariate calibration. **Food Chemistry**, v. 174, p. 1-7, 2015.

CASSOLI, L. D.; SARTORI, B.; ZAMPAR, A.; MACHADO, P. F. An assessment of Fourier transform infrared spectroscopy to identify adulterated raw milk in Brazil. **International Journal of Dairy Technology**, v. 64, p. 480-485, 2011.

CHATURVEDI, S.; YADAV, R. L. GUPTA, A. K.; SHARMA, A. K. Life time milk amount prediction in dairy cows using artificial neural networks. **International Journal of Recent Research and Review**, v. 5, p 1-6, 2013.

CHOUDHURRY, A.; MUKHERJEE, M.; ADHIKARI, B. Thermal stability and degradation of the post-use reclaim milk pouches during multiple extrusion cycles. **Thermochimica Acta**, n. 430, p. 87-94, 2005.

CYBENKO, G. **Continuous valued neural networks with two hidden layers are sufficient**. Techinal Report, Department of Computer Science, Tufts University, 1988

COATES, J. P. Interpretation of Infrared Spectra, A Practical Approach. **Encyclopedia of Analytical Chemistry**, Ed. R.A. Meyers, J. Wiley & Sons, Ltd., Chichester, UK, p. 10815-10837, 2000.

COATES, J. Vibrational Spectroscopy: Instrumentation for infrared and Raman spectroscopy. **Applied Spectroscopy Reviews**, v. 33, p. 267-425, 1998.

COROIAN, A.; TRIF, M. CORIAN, C.O.; MIRESAN, V.; RĂDUCU, C.; DĂRĂBAN, S. Qualitative evaluation of buffalo cheese using FTIR spectroscopy. **International Journal of the Bioflux Society**, v. 4, n.2, p.66-70, 2012.

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 Asiagod'Allevo cheese. **Journal of Agricultural and Food Chemistry**, v. 57, p. 11449-11454, 2009.

- CZERWENKA, C.; MÜLLER, L.; LINDNER, W. Detection of the adulteration of water buffalo milk and mozzarella with cow's milk by liquid chromatography-mass spectrometry analysis of β -lactoglobulin variants. **Food Chemistry**, v. 122, p. 901-908, 2010.
- DALMASSO, A.; CIVERA, T.; L.A. NEVE, F.; BOTTERO, M. T. Simultaneous detection of cow and buffalo milk in mozzarella cheese by Real-Time PCR assay. **Food Chemistry**, v. 124, n. 1, p. 362-366, 2011.
- DANESHVAR, N.; KHATAEE, A. R.; DJAFARZADEH, N. The use of artificial neural networks (ANN) for modeling of decolorization of textile dye solution containing C.I. Basic Yellow 28 by electrocoagulation process. **Journal of Hazardous Materials**, v. 137, n. 3, p. 1788-1795, 2006.
- DA SILVA, N. R.; PADOVANI, C. R. Utilização de componentes principais em experimentação agrônômica. **Energia na Agricultura**, v.21, n. 4, p. 98-113, 2006.
- DAS, K. NENADIC, Z. Approximate information discriminant analysis: A computationally simple heteroscedastic feature extraction technique. **Pattern Recognition**, v. 41, n. 5, p. 1548-1557, 2008.
- DENG, Y.; ZHOU, Q.; SUN, S. Analysis and discrimination of infant powdered milk via FTIR spectroscopy. **Spectroscopy and Spectral Analysis**, v. 26, n. 4, p. 636-639, 2006.
- DIAS, S. S. **Avaliação da qualidade e autenticidade de derivados de leite de búfala encontrados no varejo do Rio de Janeiro**. 59f, 2009. Dissertação (Mestre em Ciência de Alimentos) – Instituto de Tecnologia, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2009.
- DIAS, S. S.; LOBATO, V.; VERRUMA-BERNARDI, M.R. Metodologias para identificar adulteração em queijos produzidos com leite de diferentes espécies de animais. **Revista do Instituto Adolfo Lutz**, São Paulo, v.68, n.3, 2009.
- DOMINGO, E.; TIRELLI, A. A.; NUNES, C. A.; GUERREIRO, M. C.; PINTO, S. M. Melamine detection in milk using vibrational spectroscopy and chemometrics analysis: a review. **Food Research International**, v. 60, n. 131-139, 2014.

ELMOLLA, E. S.; CHAUDHURI, M.; ELTOUKHY, M. M. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. **Journal of Hazardous Materials**, v. 179, n. 1-3, p. 127-134, 2010

ENNE, G.; ELEZ, D.; FONDRINI, F.; BONIZZU, I.; FELIGINI, M.; ALEANDRI, R. High-performance liquid chromatography of governing liquid to detect illegal bovine milk's addition in water buffalo Mozzarella: Comparison with results from raw milk and cheese matrix. **Journal of Chromatography A**, v. 1094, p. 169-174, 2005.

EGITO, A. S.; ROSINHA, G. M. S.; LAGUNA, L. E.; MICLO, L.; GIRARDET, J. M.; GAILLARD, J. L. Método eletroforético rápido para detecção da adulteração do leite caprino com leite bovino. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 58, n. 5, p. 932-939, 2006.

ETZION, Y.; LINKER, R.; COGAN, U.; SHMULEVICH, I. Determination of protein concentration in raw milk by mid-infrared Fourier transform infrared/attenuated total reflectance spectroscopy. **Journal of Dairy Science**, v. 87, n. 9, p. 2779–2788, 2004.

FARIA, C. P.; BENEDET, H. D.; GUERROUE, J. L. L. Parâmetros de produção de leite de búfala fermentado por *Lactobacillus casei*. **Pesquisa Agropecuária Brasileira**, v.41, n.3, p.511-516, 2006.

FARIA, M. H.; TONHATI, H.; CERÓN M. M.; DUARTE, J. M. C. Características físico-químicas do leite de búfalas ao longo da lactação. **Revista do Instituto de Laticínios Cândido Tostes**, v. 57, n. 324, p. 3-7, 2002.

FAYE, B.; KONUSPAYEVA, G.; MESSAD, S.; LOISEAU, G. Discriminant milk components of Bactrian camel (*Camelus bactrianus*), dromedary (*Camelus dromedarius*) and hybrids. **Dairy Science and Technology**, v. 88, n. 6, p. 607–617, 2008.

FERNANDES, S. A. A.; MATTOS, W. R.; MATARAZZO, S. V.; ROSETO, C. V.; MACHADO, P. H. Componentes do leite de bubalinos ao longo da lactação no Estado de São Paulo. **Revista do Instituto de Laticínios Cândido Tostes**, v. 60, n. 346-347, p. 71-78, 2005.

FERRÃO, M. F.; CARVALHO, C. W.; MULLER, E. I. Determinação simultânea dos teores de cinza e proteína em farinha de trigo empregando NIR-PLS e DRIFT-PLS. **Ciência e Tecnologia de Alimentos**, v. 24, p. 333-340, 2004.

FERRÃO, M. F.; MELLO, C.; BORIN, A.; MARRETO, D. A.; POPPI, R. J. LS-SMV: Uma ferramenta quimiométrica para regressão multivariada. Comparação de modelos de regressão LS-SVM e PLS na quantificação de adulterantes em leite em pó empregando NIR. **Química Nova**, v. 30, n. 4, p. 852-859, 2007.

FERREIRA, D. F. **Estatística multivariada**. Lavras: Editora UFLA, 2011. 675p.

FIORIN, D. V.; MARTINS, F. R.; SCHUCH, N. J.; PEREIRA, E. B. Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares. **Revista Brasileira de Ensino de Física**, n.33, n.1, p. 01-20, 2011.

FILHO, M. H. B. C; JUNIOR, D. M. L; RANGEL, A. H. N.; SILVA, F. J. S.; NOVAES, L. P.; JÚNIOR, G. B. G; SILVA, M. J. M. S.; MORENO, G. M. B Sazonalidade e variação na qualidade do leite de búfalas no rio grande do norte. **Acta Veterinaria Brasilica**, v.8, n.3, p.201-208, 2014.

FISCHER, I.; HENNECKE, F.; BANNES, C.; ZELL, A. **Java Neural Network Simulator – JavaNNS. Version 1.1**. Wilhelm-Schickard-Institute for Computer Science - Department of Computer Architecture, University of Tübingen, Alemanha, 2001.

FISHER, R. A. Has Mendel's work been rediscovered? **Annals of Science**, v. 1, p. 115–137, 1936.

FONTENELE, M. A. **Caracterização físico-química, avaliação sensorial, proteômica e bioquímica do queijo coalho do Jaguaribe - CE visando o processo de indicação geográfica - IG**. 2013. 146p. Tese (Doutorado) - Universidade Estadual do Ceará, Ceará.

FOOD AND AGRICULTURE ORGANIZATION – FAO. **FAOSTAT: Agriculture**. Acesso em: Agosto de 2014.

GOYAL, S. Artificial neural networks (ANNs) in food science and a review. **International Journal of Scientific World**, v. 1, n. 2, p.19-28, 2013.

GÓMEZ-RUIZ, J. A.; MIRALLES, B.; AGÜERA, P.; AMIGO, L. Quantitative determination of α s2- and α s1-casein in goat's milk with different genotypes by capillary electrophoresis. **Journal of Chromatography A**, v. 1054, p. 279-284, 2004.

GUERREIRO, J. S.; BARROS, M.; FERNANDES, P.; PIRES, P.; BARDSLEY, R. Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. **Food Chemistry**, v. 136, n. 3-4, p. 1526-1532, 2013.

HAIR JR.; J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. **Multivariate Data Analysis**. 7ed. Pearson Prentice Hall, 2010. 816 p.

HELPER, G. A.; FERRÃO, M. F.; FERREIRA, C. V. Aplicação de métodos de análise multivariada no controle qualitativo de essências alimentícias empregando espectroscopia no infravermelho médio. **Ciência e Tecnologia de Alimentos**, v. 26, n. 4, p. 779-786, 2006.

HOLLER, F. J.; SKOOG, D. A.; CROUCH, S. R. **Princípios da análise instrumental**. 6ª ed. Porto alegre, Bookman Companhia Editora LTDA, 2009. 1056 p.

IBGE. Instituto Brasileiro de Geografia e Estatística. **SIDRA**. Disponível em: <<http://www.sidra.ibge.gov.br/>> Acesso em: Agosto de 2015.

INON, F. A.; GARRIGUES, S.; GUARDIA, M. Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. **Analytica Chimica Acta**, v. 513, n. 2, p. 401-412, 2004.

JAISWAL, P.; JHA, S. N.; BORAH, A.; GAUTAM, A.; GREWAL, M. K.; JINDAL, G. Detection and quantification of soymilk in cow-buffalo milk using Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR-FTIR). **Food Chemistry**, v. 168, n.1, p. 41-47, 2015.

JIN, J.; AN, J. Robust discriminant analysis and its application to identify protein coding regions of rice genes. **Biosciences Matemática**, v. 232, n. 2, p.96-100, 2011.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. New Jersey: Prentice-Hall, 1999.816p.

KAMAL, M.; KAROUI, R. Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: A review. **Trends in Food Science & Technology**, v. 46, p. 27-48, 2015.

KAROUI, R.; DOWNEY, G.; BLECKER, C. Mid-infrared spectroscopy coupled with chemometrics: a tool for the analysis of Intact food systems and the exploration of their molecular structure-quality relationships and a review. **Chemical Reviews**, n. 110, p. 6144-6168, 2010.

KAROUI, R.; HAMMAMI, M.; ROUISSI, H.; BLECKER, C. Mid infrared and fluorescence spectroscopies coupled with factorial discriminant analysis technique to identify sheep milk from different feeding systems. **Food Chemistry**, v. 127, n. 2, p. 743-748, 2011.

KAROUI, R.; MOUAZEN, A. M.; DUFOUR, I.; PILLONEL, L.; PICQUE, D.; BOSSET, J. O. Mid infrared spectrometry: A tool for the determination of chemical parameters of Emmental cheeses produced during winter. **Le Lait Dairy Science and technology**, v. 86, p. 83–97, 2006.

KASEMSUMRAN, S.; THANAPASE, W.; KIATSSONTHON, A. Feasibility of nearinfrared spectroscopy to detect and to quantify adulterants in cow milk. **Analytical Sciences**, v. 23, n. 7, p. 907–910, 2007.

KHATAEE, A. R.; ZAREI, M.; POURHASSAN, M. Bioremediation of malachite green from contaminated water by three microalgae: Neural network modeling. **Clean-Soil, Air, Water**, v. 38, n. 1, p. 96-103, 2010.

LEI, Y.; QUN, Z.; ZHANG, Y-L.; CHEN, J-B.; SUN, S-Q.; NODA, I. Analysis of crystallized lactose in milk powder by Fourier-transform infrared spectroscopy combined with two-dimensional correlation infrared spectroscopy. **Journal of Molecular Structure**, v. 974, n. 1-3, p. 88–93. 2010.

LIN, M. A review of traditional and novel detection techniques for melamine and its analogues in foods and animal feed. **Frontiers of Chemical Engineering in China**, v. 3, p. 427-435, 2009.

LOHUMI, S. H.; LEE, S.; LEE, H.; CHO, B-K. A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. **Trends in Food Science & Technology**, v. 46, n.1, p. 85-98, 2015.

LOHUMI, S.; MO, C.; KANG, J-S.; HONG, S-J.; CHO, B-K. Nondestructive evaluation for the viability of watermelon (*Citrullus lanatus*) seeds using fourier

transform near infrared spectroscopy. **Journal of Biosystems Engineering**, v. 38, n. 4, 2013.

LOCCI, F.; GHIGLIETTI, R.; FRANCOLINO, S.; IEZZI, R.; OLIVIERO, V.; GAROFALO, A.; MUCCHETTI, G. Detection of the cow milk in cooked buffalo Mozzarella used as Pizza topping. **Food Chemistry**, v. 107, p. 1337-1341, 2008.

LÓPEZ-CALLEJA, I.; ALONSO, G.; FAJARDO, V.; RODRÍGUEZ, M. A.; HERNÁNDEZ, P. E.; GARCÍA, T.; MARTÍN, R. PCR detection of cow's milk in water buffalo milk and mozzarella cheese, **International Dairy Journal**, v. 15, n. 11, p. 1122-1129, 2005.

LUO, H.; ZHANG, J.; LIU, F.; QIN, Y. Quick detection of adulterant milk by infrared spectroscopy. **IDF World Summit Shanghai**, n.1, p.1-5, 2006.

LUYKX, D. M. A. M.; VAN RUTH, S. M. An overview of analytical methods for determining the geographical origin of food products. **Analytical Methods**, v. 107, p. 897-91, 2008.

LYRA, W. S.; DA SILVA, E. C.; DE ARAÚJO, M. C. U.; FRAGOSO, W. D.; VERAS, G. Classificação periódica: um exemplo didático para ensinar análise de componentes principais. **Química Nova**, v. 33, n. 7, p. 1594-1597, 2010.

MACEDO, M. P.; WECHSLER, F. S.; RAMOS, A. A.; AMARAL, J. B.; SOUZA, J. C.; RESENDE, F. D.; OLIVEIRA, J. V. Composição físico-química e produção do leite de búfalas da raça Mediterrâneo no Oeste do Estado de São Paulo. **Revista Brasileira de Zootecnia**, v. 30, n. 3, p.1, 2001.

MAROCO, J. **Análise Estatística com utilização do SPSS**, Lisboa: Sílabo, 2003. 824p.

MATTOS, B. C. Aspectos qualitativos do leite bubalino. **PUBVET**, v.1, n.9, 9 ed., 2007. Disponível em: <http://www.pubvet.com.br/artigos_det.asp?artigo=166> Acesso em: 15 abr. 2016.

MELLING, J. M.; THOMPSON, M. **Fiber-optic probes for midinfrared spectrometry. In Handbook of Vibrational Spectroscopy.** ed. John Wiley&Sons, Chichester, UK. Disponível em < http://www.remspec.com/pdfs/2703_o.pdf.> Acesso em: 13 de Maio de 2016.

MÉNARD, O.; AHMAD, S.; ROUSSEAU, F.; BRIARD-BRION, V.; GAUCHERON, F.; LOPEZ, C. Buffalo vs. cow milk fat globules: Size distribution, zeta-potential, compositions in total fatty acids and in polar lipids from the milk fat globule membrane. **Food Chemistry**, v. 120, n. 2, p. 544–551, 2010.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005. 297p.

MOORE, J. C.; SPINK, J.; LIPP, M. Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010. **Journal of Food Science**, v. 77, n.4, p. 118-126, 2012.

MORGANO, M. A.; FARIA, C. G.; FERRÃO, M. F. Determination of protein in raw coffee for NIR spectroscopy and regression PLS. **Ciência e Tecnologia de Alimentos**, v. 25, n. 1, p. 25-31, 2005.

NAZÁRIO, S. L. S.; BUIOCHI, F.; KITANO, C.; ISEPON, J. S.; ADAMOWSKI, J. C.; HIGUTI, R. T. Caracterização de leite bovino utilizando ultra-som e redes neurais artificiais. **Revista Controle & Automação**, v. 20, n. 4, p. 627-236, 2009.

NETO, J. M. M.; MOITA, G. C. Uma introdução à análise exploratória de dados multivariados. **Química Nova**, v. 21, n. 4, p. 467-469, 1998.

NICOLAOU, N.; XU, Y.; GOODACRE, R. Fourier transform infrared spectroscopy and multivariate analysis for the detection and quantification of different milk species. **Journal of Dairy Science**, v. 93, n. 12, p. 5651–5660, 2010

OGILVIE, D.; MITCHELL, R.; MUTRIE, N.; PETTICREW, M.; PLATT, S. Perceived characteristics of the environment associated with active travel: development and testing of a new scale. **International Journal of Behavioral Nutrition and Physical Activity**, v. 5, n. 32, p. 1-10, 2008.

OTTAVIAN, M.; FACCO, P.; BAROLO, M.; BERZAGHI, P.; SEGATO, S.; NOVELLI, E. Near-infrared spectroscopy to assist authentication and labeling of Asiago d'allevato cheese. **Journal of Food Engineering**, v.113, p. 289-298, 2012.

PACHECO, J.; CASADO, S.; PORRAS, S. Exact methods for variable selection in principal component analysis: Guide functions and pre-selection. **Computational Statistics and Data Analysis**, v. 57, n. 1, p. 95–111, 2013.

PAPPAS, C. S.; TARANTILIS, P. A.; MOSCHOPOULOU, E.; MOATSOU, G.; KANDARAKIS, I.; POLISSIOU, M. G. Identification and differentiation of goat and sheep milk based on diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) using cluster analysis. **Food Chemistry**, v. 106, n. 3, p. 1271–1277, 2008

PATRAS, A.; BRUNTON, N. P.; DOWNEY, G.; RAWSON, A.; WARRINER, K.; GERNIGON, G. Application of principal component and hierarchical cluster analysis to classify fruits and vegetables commonly consumed in Ireland based on in vitro antioxidant activity. **Journal of Food Composition and Analysis**, v. 24, n. 2, p. 250-256, 2011.

PAVIA, D. L.; LAMPMAN, G. M.; KRIZ, G. S. **Introdução à espectroscopia**. 4 ed. São Paulo: Cengage Learning, 2010. 716 p.

PEREIRA, E. B. B.; PEREIRA, M. B. Um critério para o descarte de variáveis na análise de componentes principais. **Revista Universidade Rural**, v. 23, n. 1-2, p. 1-7, 2004.

PESTANA, M. H.; GAGEIRO, J. N. **Análise de Dados para Ciências Sociais**. 2 ed. Lisboa: Sílabo, 2014, 1233 p.

PIGNATA, M. C.; FERNANDES, S. A. A.; FERRÃO, S. P. B.; FALEIRO, A. S.; CONCEIÇÃO, D. G.; Estudo comparativo da composição química, ácidos graxos e colesterol de leites de búfala e vaca. **Revista Caatinga**, v. 27, n. 4, p. 226-233, 2014.

RAI, P., MAJUMDAR, G.C., DASGUPTA, S., DE, S. Prediction of the viscosity of clarified fruit juice using artificial neural network: a combined effect of concentration and temperature. **Journal of Food Engineering**, v. 68, n.4, p. 527-533, 2005.

RAMOS, M.; JUAREZ, M. Sheep milk. In: ROGINSKI, H., FUQUAY, J.W., FOX, P.F. (Eds.), **Encyclopedia of Dairy Sciences**, Amsterdam: Academic Press, v. 4, p. 2539-2545, 2003.

RANGEL, A. H. N.; OLIVEIRA, J. P. F.; ARAÚJO, V. M.; BEZERRA, K. C.; MEDEIROS, H. R.; LIMA JÚNIOR, D. M.; ARAÚJO, C. G. F. Influência do estágio de lactação sobre a composição do leite de búfalas. **Acta Veterinaria Brasílica**, v. 5, n. 3, p. 306-310, 2011.

RAYNAL-LJUTOVAC, K.; LAGRIFFOUL, G.; PACCARD, P.; GUILLET, I.; CHILLIARD, Y. Composition of goat and sheep milk products: An update. **Small Ruminant Research**, v. 79, p. 57-72, 2008.

RECHE, A. N. **Análise espectroscópica no infravermelho do leite humano cru e processado/pasteurizado de banco de leite**, Dissertação (Mestrado em Bioengenharia) - Universidade do Vale do Paraíba (UniVap), São José dos Campos - SP, 2013.

REGAZZI, A. J. **Análise multivariada, notas de aula INF 766**, Departamento de Informática da Universidade Federal de Viçosa, v.2, 2000.

REID, L. M.; WOODCOCK, T.; O'DONNELL, C.; DOWNEY, G. Differentiation of Apple juice samples on the basis of heat treatment and variety using chemometric analysis of MIR and NIR. **Food Research International**, v. 38, p.1109-1115, 2005.

RICCI G. D.; DOMINGUES P. F. O leite de búfala. **Journal of Continuing Education in Animal Science of CRMV-SP**, v. 10, n. 1, p. 14–19, 2012.

RODRIGUES, N. P. A.; GIVISIEZ, P. E. N.; QUEIROGA, R. C. R. E.; AZEVEDO, P. S.; GEBREYES, W. A.; OLIVEIRA, C. J. B. Milk adulteration: Detection of bovine milk in bulk goat milk produced by smallholders in northeastern Brazil by a duplex PCR assay. **Journal of Dairy Science**, v. 95, n. 5, p. 2749-52, 2012.

RODRIGUEZ-NOGALES, J. M.; VASQUÉZ, F. Application of electrophoretic and chemometric analysis to predict the bovine, ovine and caprine milk percentages in Panela cheese, an unripened cheese. **Food Control**, v.18, n.5, p.570–576, 2007.

RONCADA, P.; PIRAS, C.; SOGGIU, A.; TURK, R.; URBANI, A.; BONIZZI, L. Farm animal milk proteomics. **Journal of Proteomics**, v. 75, n. 14, p. 4259-4274, 2012.

SACCO, D.; BRESCIA, M. A.; SGARAMELLA, S.; CASIELLO, G.; BUCCOLIERI, A.; OGRINC, N. Discrimination between Southern Italy and foreign milk samples using spectroscopic and analytical data. **Food Chemistry**, v.114, n.4, p.1559–1563, 2009.

SANTOS, P. M.; PEREIRA-FILHO, E. R.; RODRIGUEZ-SAONA, L. E. Rapid detection and quantification of milk adulteration using infrared microspectroscopy and chemometrics analysis. **Food Chemistry**, v. 138, n. 1, p. 19-24, 2013a.

SANTOS, P. M.; FILHO, E. R. P.; SOANA, L. E. R. Application of Hand-Held and Portable Infrared Spectrometers in Bovine Milk Analysis. **Journal of Agricultural and Food Chemistry**, n.61, p. 1205–1211, 2013b

SCHEEREN, L. W. et al. Agrupamento de unidades amostrais de *Araucaria angustifolia* (Bert.) O. Ktze. em função de variáveis do solo, da serapilheira e das acículas, na região de Canela, RS. **Ciência Florestal**, Santa Maria, v. 10, n. 2, p. 39 - 57, 2000.

SCHMILOVICH, Z.; SHMULEVICH, I.; MALTZ, E. Near infrared spectroscopy of milk in its heterogeneous state. **Computadores e Eletrônica em Agricultura**, v. 29, n. 3, p.195-207, 2000.

SENA, M. M.; POPPI, R. J.; FRIGHETTO, R. T. S.; VALARINI, P. J. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, v. 23, n. 4, p. 547-556, 2000.

SEIDEL, E. J.; JUNIOR, F. J. M.; ANSUJ, A. P. NOAL, M. R. C. Comparação entre o método Ward e o método K-médias no agrupamento de produtores de leite. **Ciência e Natura**, UFSM, v. 30, n. 1, p. 7-15, 2008.

SILVA, R. V. S. **Deteção da presença de leite bovino no leite caprino por imunocromatografia e ELISA**. Patos - PB: UFCG, 2010. 34 p. Dissertação - Mestrado em Medicina Veterinária.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. Redes Neurais Artificiais: para engenharia e ciências aplicadas. São Paulo: Editora Artliber Ltda, 2010, 399p.

SINDHU, J. S.; ARORA, S. **Buffalo Milk**. In: Encyclopaedia of Dairy Sciences, 2 ed. Academic Press, San Diego, 2011. p. 503–511

SINGH, K. P. et al. Artificial neural network modeling of the river water quality: A case study. **Ecological Modelling**, n. 220, v. 6, p. 888-895, 2009.

SKOOG, D. A.; HOLLER, F. J.; NIEMAN, T. A. **Princípios de análise instrumental**. 5 ed. Porto Alegre: Editora Bookman, 2002. 836 p.

SOLA-LANARRANGA, C.; NAVARRO-BLASCO, I. Chemometric analysis of minerals and trace elements in raw cow milk from the community of Navarra, Spain. **Food Chemistry**, v.112, n.1, p. 189–196, 2009.

SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial, parte 1. **Química Nova**, v. 35, n. 1, p. 223-229, 2012

SOUZA, S. S.; CRUZ, A. G.; WALTER, E. H. M.; FARIA, J. A. F.; CELEGHINI, R. M. S.; FERREIRA, M. M. C.; GRANATO, D.; SANT'ANA, A. S. Monitoring the authenticity of Brazilian UHT milk: A chemometric approach. **Food Chemistry**, v. 124, n. 2, p. 692–695, 2011

SUBRAMANIAN, A.; PRABHAKAR, V.; SAONA, L.R. **Infrared spectroscopy in dairy analysis**. In Encyclopedia of Dairy Sciences, New York, 2ed. 2011, 115–124p.

TEIXEIRA, L. V.; BASTIANETTO, E.; OLIVEIRA, D. A. A. Leite de Búfala na Indústria de Produtos Lácteos. **Revista Brasileira de Produção Animal**, v. 29, n. 2, p. 96-100, 2005.

THERMO NICOLET, **Introduction to Fourier Transform Infrared Spectrometry** 2001. Disponível em < <http://mmrc.caltech.edu/FTIR/FTIRintro.pdf>>. Acesso em: Maio de 2016

VALENTE, G. F. S.; GUIMARÃES, D. C.; GASPARDI, A. L. A.; OLIVEIRA, L. A. Aplicação de redes neurais artificiais como teste de detecção de fraude de leite por adição de soro de queijo. **Revista do Instituto de Laticínios Cândido Tostes**, v. 69, n. 6, p. 425-432, 2014.

VALENTI, B.; MARTIN, B.; ANDUEZA, D.; LEROUX, C.; LABONNE, C.; LAHALLE, F. Infrared spectroscopic methods for the discrimination of cows' milk according to the feeding system, cow breed and altitude of the dairy farm. **International Dairy Journal**, v. 32, n. 1, p. 26-32, 2013.

VELOSO, A. C.; TEIXEIRA, N.; FERREIRA, I. M. P. L. V. O.; FERREIRA, M. A.; Detecção de adulterações em produtos alimentares contendo leite e/ou proteínas lácteas. **Revista Química Nova**, v. 25, n. 4, p. 609-615, 2002.

VENTURINI, K. S.; SARCINELLI, M. F.; DA SILVA, L. C. **Características do Leite**. Boletim Técnico: Universidade Federal do Espírito Santo (UFES), 2007. 80 p.

WATKINS, P.; WIJESUNDERA, C. A preliminary study on the application of cluster analysis to the determinations of the geographical origin of cheddar cheese based on

semi-volatile composition. **Australian Journal of Dairy Technology**, v. 61, n. 3, p. 244–247, 2006.

WU, D.; FENG, S.; HE, Y. Infrared spectroscopy technique for the nondestructive measurement of fat content in milk powder. **Journal of Dairy Science**, v. 90, n. 8, p. 3613- 3619, 2007.

WU, D.; FENG, S.; HE, Y. Short-wave near-infrared spectroscopy of milk powder for brand identification and component analysis. **Journal of Dairy Science**, v. 91, n. 3, p. 939-949, 2008.

ZALESKA. H.; TOMASIK, P.; LIU, C. Y. Formation of carboxymethyl L cellulose-casein complexes by electrosynthesis. **Food Hydrocolloids**, v. 16, 215-224. 2002.

ZHOU, W.; THERDTHAI, N. **Artificial neural networks in food processing**. M.M. Farid (Ed.), **Mathematical modeling of food processing**. Boca Raton: CRC Press. p. 2010. 901– 920.