



UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I - CAMPINA GRANDE
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

WEDJA MARCELINO DA SILVA

**DETERMINAÇÃO SIMULTÂNEA DE ADULTERAÇÃO EM LEITE DE
CABRA POR ADIÇÃO DE LEITE DE VACA E TEOR DE LIPÍDIOS USANDO
NIR PORTÁTIL E ALGORITMOS PLS**

CAMPINA GRANDE

2023

WEDJA MARCELINO DA SILVA

DETERMINAÇÃO SIMULTÂNEA DE ADULTERAÇÃO EM LEITE DE CABRA POR ADIÇÃO DE LEITE DE VACA E TEOR DE LIPÍDIOS USANDO NIR PORTÁTIL E ALGORITMOS PLS

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Universidade Estadual da Paraíba, como requisito para a obtenção do título de Mestre em Ciências Farmacêuticas.

Área de concentração:

Desenvolvimento e controle da qualidade de produtos farmacêuticos.

Orientador: Prof. Dr. José Germano Vêras Neto.

Coorientadora: Dra. Elaine Virginia dos Santos Pereira

CAMPINA GRANDE

2023

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

S586d Silva, Wedja Marcelino da.

Determinação simultânea de adulteração em leite de cabra por adição de leite de vaca e teor de lipídios usando NIR portátil e algoritmos PLS [manuscrito] / Wedja Marcelino da Silva. - 2023.

51 p.

Digitado.

Dissertação (Mestrado em Ciências Farmacêuticas) - Universidade Estadual da Paraíba, Centro de Ciências Biológicas e da Saúde, 2023.

"Orientação : Prof. Dr. José Germano Vêras Neto , Departamento de Química - CCT. "

"Coorientação: Profa. Dra. Elainy Virginia dos Santos Pereira , UFPB - Universidade Federal da Paraíba"

1. Mínimos quadrados parciais. 2. Regressão multivariada.
3. Leite caprino. I. Título

21. ed. CDD 637.17

WEDJA MARCELINO DA SILVA

DETERMINAÇÃO SIMULTÂNEA DE ADULTERAÇÃO EM LEITE DE CABRA POR ADIÇÃO DE LEITE DE VACA E TEOR DE LIPÍDIOS USANDO NIR PORTÁTIL E ALGORITMOS PLS

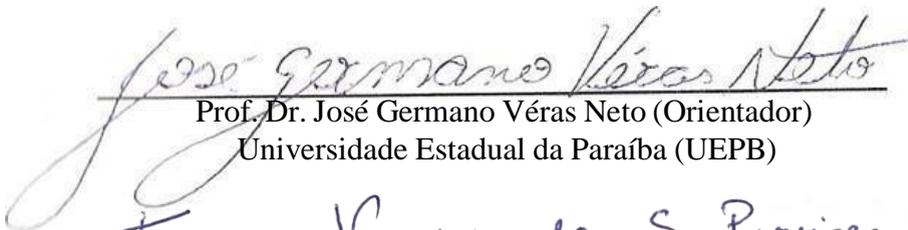
Dissertação apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Universidade Estadual da Paraíba, como requisito para obtenção do título de Mestre em Ciências Farmacêuticas

Área de concentração:

Desenvolvimento e controle de qualidade de produtos farmacêuticos.

APROVADA EM: 15 / 02 / 2023

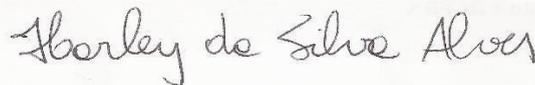
BANCA EXAMINADORA



Prof. Dr. José Germano Vêras Neto (Orientador)
Universidade Estadual da Paraíba (UEPB)



Dra. Elaine Virginia dos Santos Pereira (Coorientadora)
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Harley da Silva Alves
Universidade Estadual da Paraíba – UEPB (Examinador interno)



Dr. Valber Elias de Almeida (Examinador externo)

A Deus que me guiou, iluminou e me
proporcionou estabilidade emocional,
DEDICO.

AGRADECIMENTOS

Agradeço primeiramente a Deus por sua infinita bondade e por me conceder sabedoria, força e coragem para dar continuidade a minha caminhada.

A minha família, em especial a meu pai Cicero Matias da Silva que não teve oportunidade de estudar, mas é PhD na vida e com seu infinito amor e simplicidade sempre ensinou grandes lições. Que honra e orgulho de tê-lo na minha vida, eu te amo meu “painho”.

A Fabiano Geraldo dos Santos por toda paciência, carinho, cuidado, encorajamento e principalmente por sempre estar ao meu lado nos momentos difíceis. “Um amigo fiel é uma poderosa proteção: quem o achou descobriu um tesouro” Eclesiastes 6,14.

A Matheus Kayke, um anjo de luz que Deus me proporcionou a honra de conhecer aqui na terra. Minha eterna gratidão por todo apoio, encorajamento e por sempre acreditar em mim.

A Eullália Gonçalo das Neves e Silva, quantas coisas dividimos minha amiga, quantos risos e choros compartilhados, mas o que mais importou/importava era/é que tínhamos/temos uma a outra para nos apoiar. Gratidão por tudo, saiba que és muito especial para mim.

A Aline Lacerda Lino da Silva pelo apoio e incentivo, além sempre disponibilizar o seu tempo para ler meu trabalho. Minha eterna gratidão, saiba que pode contar comigo também.

A Karllos Matheus Campos pela amizade, preocupação e apoio sempre demonstrado tão carinhosamente.

A Danúbio Leonardo Bernardino de Oliveira pela preocupação, apoio e principalmente pelos conselhos. Gratidão, és um ser humano incrível.

A Simone Santos que em épocas difíceis, sempre me orientou tão pacientemente e docemente, resgatando sempre meus potenciais.

A Lêda Maria Marinho Santos de Figueiredo pelo acolhimento e pelas conversas descontraídas que tornavam meus dias mais leves. Gratidão por tudo.

A Maria Eduarda da Silva Sales por sempre me ajudar no trabalho para que eu pudesse assistir aula. Grata pela amizade, parceria e companheirismo, você é maravilhosa.

A meu orientador prof. Dr. José Germano Vêras Neto, um homem íntegro e

sábio. Ao senhor minha gratidão pela paciência em transmitir o seu conhecimento e na execução deste trabalho. Saiba que “Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes” (Marthin Luther King).

A Dr^a. Elaine Virginia dos Santos Pereira pela atenção, carinho, preocupação, apoio e em especial por partilhar parte do seu banco de dados.

Ao prof. Dr. Felipe Hugo Alencar Fernandes pela magnitude que transborda em seu ser e irradia a todos. Gratidão por todo o conhecimento transmitido.

A Valber e a Harley que tão gentilmente aceitaram o convite para compor a minha banca. Grata pela honra dada em compartilhar seus conhecimentos, só irá agregar no meu trabalho.

A Thiago Araujo, secretário do programa de pós graduação, por sempre ser tão gentil e sempre disposto a ajudar.

A Edvânia Emanuelle Pinheiro e Milena Nogueira que sempre estiveram me auxiliando em vários momentos do mestrado.

A todos os meus amigos que aqui não foram citados, mas que ocupam um lugarzinho especial no meu coração.

“Que a Vossa bondade permita aos espíritos consoladores derramarem por toda a parte a paz, a esperança e a fé.”

Prece de Cáritas

RESUMO

A procura pelo leite de cabra vem crescendo nos últimos anos por ser um alimento nutricionalmente rico. Para dispor de todos os benefícios ofertados e melhor utilização dos nutrientes, tais como os lipídios que desempenham funções fisiológicas do leite caprino no organismo do consumidor, é necessário que o fornecedor garanta a conservação das suas características originais para que o nível de qualidade do produto seja adequado. Assim, o presente trabalho tem como objetivo comparar o desempenho de modelos quimiométricos de regressão PLS utilizado para determinação do percentual da mistura de leite de vaca em leite de cabra e da concentração de lipídios utilizando espectrometria NIR portátil. Os pré-processamentos utilizados nos dados espectrais foram SNV e MSC para correção do espalhamento, baseline e derivação de Savitzky Golay para correção de linha de base. Para a construção do modelo de calibração foi utilizado o método de quadrados mínimos parciais (PLS). Com base nos resultados após a aplicação dos pré-processamentos os melhores modelos obtidos para quantificação do percentual de mistura foram os modelos utilizando derivação Savitzky Golay de 2º derivada, janela de 15 pontos e polinômio 2º ordem e Savitzky Golay de 2º derivada, janela 23 e polinômio 2º ordem. As figuras de mérito avaliadas foram: RMSEC (10,17 e 12,19%), RMSECV (14,14 e 12,19%), RMSEP (10,66 e 10,04%), r^2_{cal} (0,89 e 0,84), r^2_{val} (0,79 e 0,78) e r^2_{pred} (0,72 e 0,75) com as 15 e 8 variáveis latentes, respectivamente. Para concentração de lipídeos, o SNV apresentou os melhores resultados obtendo os seguintes valores para RMSEC (0,28 g/100g), RMSECV (0,33 g/100g), RMSEP (0,27 g/100g), r^2_{cal} (0,46), r^2_{val} (0,33) e r^2_{pred} (0,27) usando 7 variáveis latentes. É possível observar valores das figuras de mérito adequadas para a proporção da mistura de leite bovino em leite caprino. Em termos do modelo para determinação de lipídios não foi possível obter resultados adequados para ser usado em aplicações reais.

Palavras-chave: Mínimos quadrados parciais; Leite de cabra; Regressão multivariada.

ABSTRACT

The demand for goat's milk has been growing in recent years because it is a nutritionally rich food. In order to enjoy all the benefits offered and better use of nutrients, such as the lipids that play a regulatory role in goat milk in the consumer's body, the supplier must ensure that its original characteristics are preserved so that the product's quality level is maintained adequate. Thus, the present work aims to compare the performance of chemometric PLS regression models used to experiment with the percentage of cow's milk mixture in goat's milk and the lipid concentration using portable NIR spectrometry. The pre-processing used in the spectral data were SNV and MSC for scattering correction, baseline and Savitzky Golay derivation for baseline correction. For the construction of the calibration model, the partial least squares method (PLS) was used. Based on the results after applying the pre-processing, the best models obtained for quantifying the percentage of mixture were the models using Savitzky Golay derivation with 2nd derivative, 15-point window and 2nd order polynomial and Savitzky Golay with 2nd derivative, window 23 and 2nd order polynomial. The figures of merit evaluated were: RMSEC (10.17 and 12.19%), RMSECV (14.14 and 12.19%), RMSEP (10.66 and 10.04%), r^2_{cal} (0.89 and 0.84), r^2_{val} (0.79 and 0.78) and r^2_{pred} (0.72 and 0.75) with the 15 and 8 latent variables, respectively. For lipid concentration, SNV showed the best results, obtaining the following values for RMSEC (0.28 g/100g), RMSECV (0.33 g/100g), RMSEP (0.27 g/100g), r^2_{cal} (0.46), r^2_{val} (0.33) and r^2_{pred} (0.27) using 7 latent variables. It is possible to observe values of figures of merit suitable for the proportion of bovine milk in goat milk. In terms of the model for lipid naivety, it was not possible to obtain adequate results to be used in real applications.

Keywords: Partial Least Squares; Goat milk; Multivariate regression.

LISTA DE ABREVIATURAS E SIGLAS

MSC	Correção Multiplicativa de Espalhamento (<i>Multiplicative Scatter Corretion</i>)
NIR	Infravermelho próximo (<i>Near Infrared</i>)
NIRS	Espectroscopia no Infravermelho Próximo (<i>Near Infrared Spectroscopy</i>)
NUPEA	Núcleo de Pesquisa e Extensão em Alimentos
PC	Componentes Principais (<i>Principal Component</i>)
PCA	Análise de Componentes Principais (<i>principal component analysis</i>)
PLS	Mínimos Quadrados Parciais (<i>Partial Least Squares</i>)
PLSR	Regrssão por Mínimos Quadrados Parciais (<i>Partial Least Squares Regression</i>)
PTFE	Politetrafluoretuleno
r^2	Coefficiente de determinação
REP	Erro relativo de predição (<i>Relative Error of Predictions</i>)
RMSEC	Raiz quadrada do erro médio quadrático de calibração (<i>Root Mean Square Error of Calibration</i>)
RMSECV	Raiz quadrada do erro médio quadrático de validação cruzada (<i>Root Mean Square Error of Cross Validation</i>)
RMSEP	Raiz quadrada do erro médio quadrático de predição (<i>Root Mean Square Error of Prediction</i>)
SG	Savitsky Golay
SNV	Varição Normal Padrão (<i>Standard Normal Variate</i>)
SPXy	Partição do conjunto de amostras baseado nas distâncias conjuntas (<i>Sample set Partitioning based on joint X-y distances</i>)
TACO	Tabela Brasileira de Composição de Alimentos

SUMÁRIO

1	INTRODUÇÃO	12
2.1	Objetivo Geral	14
2.2	Objetivos Específicos	14
3	REFERENCIAL TEÓRICO.....	15
3.1	Leite	15
3.2	Leite de cabra	16
3.3	Controle de qualidade.....	17
3.4	Adulteração de leite de cabra por adição de leite de vaca.....	17
3.5	Espectrometria no infravermelho próximo (NIR)	18
3.6	Quimiometria.....	19
3.7	Técnicas quimiométricas	21
3.7.1	<i>Pré - processamentos.....</i>	21
3.7.2	<i>Análise de Componentes Principais (PCA).....</i>	21
3.7.3	<i>Calibração multivariada.....</i>	22
3.7.4	<i>Mínimos quadrados parciais (PLS).....</i>	23
4	MATERIAIS E MÉTODOS.....	25
4.1	Amostras	25
4.2	Preparo das amostras	26
4.3	Determinação de lipídios	26
4.4	Análises das amostras	26
4.4.1	<i>NIR portátil.....</i>	26
4.4.2	<i>Banco de dados.....</i>	27
4.4.3	<i>Métodos quimiométricos</i>	27
4.4.4	<i>Seleção das amostras.....</i>	27
4.4.5	<i>Pré-processamentos.....</i>	27

4.4.6	<i>Construção do modelo</i>	28
5	RESULTADOS E DISCUSSÃO	30
5.1	Avaliação das propriedades de interesse	30
5.2	Análise espectral	31
5.3	Análise exploratória das propriedades de interesse	32
5.4	Modelos PLS para misturas de leite de vaca em leite de cabra	36
5.4.1	<i>Modelos PLS para dados brutos nas misturas de leite de vaca em leite de cabra</i>	36
5.4.2	<i>Modelo PLS para dados pré-processados nas misturas do leite de vaca em leite de cabra</i>	37
5.5	Modelos PLS para determinação de lipídios	40
5.5.1	<i>Modelo PLS para dados brutos na determinação de lipídios</i>	40
5.5.2	<i>Construção dos modelos de PLS para lipídios com espectros pré- processadas</i> 41	
6	CONCLUSÃO	43
	REFERÊNCIAS	44

1 INTRODUÇÃO

O leite é um alimento bastante consumido pela população mundial por ser constituído de nutrientes básicos e essenciais que desempenham um papel vital no desenvolvimento dos seres humanos ao longo da vida (ADINEPOUR et al., 2022; AKINYEMI et al., 2022; LEMOS et al., 2022).

O leite de cabra é o segundo alimento lácteo mais consumido, tendo um valor nutricional significativo, além de possuir componentes bioativos que apresentam efeitos benéficos à saúde humana (HAN et al., 2022; WANG; WANG; GUO, 2020). Este produto se destaca por possuir uma quantidade menor de gordura, sendo considerado hipoalergênico, em relação ao leite de vaca. Além disso, é recomendado para crianças, idosos e pessoas em recuperação de doenças ou distúrbios digestivos (MARQUES et al., 2021). Vale destacar que a composição do leite caprino pode variar de acordo com fatores genéticos, fisiológicos, ambientais e alimentícios, como de todo e qualquer produto de origem animal (VERRUCK et al., 2019).

Devido a ser rico em vitaminas, minerais e aminoácidos essenciais (MARQUES et al., 2021) a procura por leite de cabra vem crescendo nos últimos anos. Com isso, fica mais propenso a falsificações no mercado (LIU et al 2020), já que muitos comerciantes, com o intuito de elevar os seus lucros, utilizam substâncias para aumentar o volume ou conferir alguma outra propriedade do leite (água, leite proveniente de outro animal e até produtos danosos à saúde como soda cáustica e peróxido de hidrogênio) (GOLINELLI et al., 2014; AGRIMONTI, PIRONDINI, MARMIROLI; MARMIROLI, 2015; DI PINTO et al., 2017).

Junto à preocupação com problemas de falsificação, outro fator que precisa ser observado é a garantia da manutenção de todas as propriedades benéficas à saúde de quem consome leite de cabra, como o de lipídios que desempenham funções fisiológicas importantes no organismo do consumidor. Dessa forma, enfatiza-se a necessidade de ferramentas que sejam capazes de auxiliar na identificação dos constituintes inerentes à matéria analisada e detectar substâncias que não estão naturalmente presentes nas amostras.

Sendo assim, técnicas analíticas capazes de identificar e quantificar compostos moleculares que sejam rápidas, de baixo custo e não-destrutivas estão sendo cada vez mais utilizadas, a exemplo de métodos como a espectrometria no infravermelho próximo (NIRS, do inglês *Near Infrared Spectroscopy*). Esta técnica é capaz de fornecer

informações para a caracterização de matérias-primas e quantificação de compostos moleculares (GRASSI; ALAMBRESE, 2018). No entanto, a obtenção de espectros por si só não fornece outras informações além da presença de determinados grupos funcionais. E com o intuito de extrair o máximo de informações químicas sobre o comportamento dos dados em estudo, o emprego de ferramentas quimiométricas têm sido imprescindíveis para complementar o uso de dados provenientes da espectroscopia NIR.

A quimiometria utiliza ferramentas que compreendem algoritmos estatísticos multivariados, sendo os mais difundidos a regressão por mínimos quadrados parciais (PLS, do inglês *Partial Least Square*) e regressão por componentes principais (PCR, do inglês *Principal Component Regression*) (ROHMAN; MAN, 2010; TEYE et al., 2019).

Desse modo, este trabalho tem como objetivo aplicar metodologias alternativas para determinação de lipídios em leite de cabra e misturas de leite bovino em leite de caprino a partir de NIR portátil e calibração multivariada.

2 OBJETIVOS

2.1 Objetivo Geral

Comparar o desempenho de modelos quimiométricos de regressão PLS utilizado para determinação da concentração de lipídios e percentual da mistura de leite de vaca em leite de cabra utilizando espectrometria NIR portátil.

2.2 Objetivos Específicos

- a) Empregar a análise de componentes principais (PCA) para observar o comportamento das amostras em estudo;
- b) Utilizar a regressão pelo método dos quadrados mínimos parciais – PLS na construção dos modelos quimiométricos;
- c) Realizar um estudo comparativo entre os modelos construídos e os dados brutos.

3 REFERENCIAL TEÓRICO

3.1 Leite

O leite dos diferentes mamíferos é um alimento nutricionalmente rico e consumido globalmente. É constituído principalmente por água, lipídios, proteínas, lactose, vitaminas e minerais que podem variar de quantidade dependendo da espécie do animal e do período de lactação. Além disso, fatores genéticos, fisiológicos e nutricionais, bem como as condições ambientais desempenham uma grande influência nessas diferenças (PIETRZAK-FIEĆKO; KAMELSKA-SADOWSKA, 2020; SUH, 2022).

O leite pode ser proveniente da vaca, cabra, ovelha, camela, búfala e outras, em que desempenha papel fundamental na vida dos seres humanos, principalmente em crianças e idosos por ser uma rica fonte de suplementação alimentar (CHEN et al., 2021). Este líquido branco, de odor e sabor característico, tem a gordura como um dos principais componentes por representar um importante indicador de qualidade (GU et al., 2021).

A proteína do leite contém aminoácidos essenciais que são indispensáveis e benéficos à saúde humana, além de estarem ligadas as características organolépticas do mesmo. Já os lipídios representam cerca de 3 a 5% da composição do leite e são responsáveis por atribuir características sensoriais ao alimento e seus derivados (CLARKE et al., 2019; LIU et al., 2020; ZHANG et al., 2022).

O leite é comercializado na forma líquida, em pó, em apresentações integrais, desnatados, pasteurizados e ultraprocessados. É a base para a origem de diversos produtos tais como: queijos, kefir, coalhada, iogurte, manteigas, sorvetes, leite condensado, creme de leite, bebidas fermentadas e não fermentadas. O leite e seus derivados proporcionam diversos benefícios a seus consumidores, desempenhando papéis importantes no crescimento, desenvolvimento e imunidade dos humanos (GU, et al., 2021; LIU; CAI; ZHANG, 2021).

A conservação das propriedades originais do leite cru é um fator primordial para garantir a segurança e a higiene do produto e qualidade dos seus derivados. Um suprimento adequado de leite de qualidade é a chave para o desenvolvimento da indústria leiteira moderna (ZHIQIANG, 2021).

3.2 Leite de cabra

As cabras fazem parte do grupo de pequenos ruminantes e é uma das espécies de animais domesticados mais antigos da história da humanidade. Sua produção vem aumentando em todo o mundo devido características peculiares, tais como: capacidade de atingir a maturidade precoce, fácil adaptação a diferentes climas e alimentação e sua lucratividade (NAMONJE-KAPEMBWA; CHIWAWA; SITKO, 2022; SHARMA et al., 2020).

A comercialização de cabras é, principalmente, realizada informalmente sendo dominada por pequenos agricultores. A cabra é um animal atrativo por fornecer carne, pele e, principalmente, leite e vários produtos lácteos. Embora as cabras ainda contribuam cerca de 2% da produção de leite total do mundo, estudos sugerem que haverá um aumento em seu consumo até 2030 (FOROUTAN et al., 2019; NAYIK et al., 2021; NAMONJE-KAPEMBWA; CHIWAWA; SITKO, 2022 ; PULINA et al., 2018). E isso é justificável diante dos diversos benefícios que o leite caprino oferece aos seres humanos.

O leite de cabra é um produto que pode ser utilizado para fabricar uma grande variedade de produtos. No entanto, produtos de alta qualidade e valor agregado só podem ser produzidos a partir de leite de cabra de excelente qualidade tendo o potencial para ser transformado em um produto que satisfaça as expectativas dos consumidores, em termos de atributos nutricionais, sanitários e sensoriais (DELGADO; STOCK; SIQUEIRA, 2020).

É um alimento rico em nutrientes básicos e essenciais para o homem, dentre os quais se destacam os lipídios que são constituídos de pequenas e médias partículas, possuem maior número de triacilgliceróis insaturados, menor peso molecular e glicoproteínas que corroboram para baixa alergenicidade, fácil digestão e absorção e maior valor biológico que o leite de vaca (KOSTIC et al., 2021; LIU; CAI; ZHANG, 2021; OLIVEIRA et al., 2021; ZHANG; ZHU; JIA, 2022; ZHAO; WANG; MAO, 2022).

Os lipídios no leite de cabra compreendem uma mistura complexa de várias espécies lipídicas, com dominância quantitativa de triacilgliceróis, seguidos por fosfolipídios, colesterol, diacilgliceróis, ácidos graxos livres, entre outras partículas constituintes. Os triacilgliceróis são a estrutura básica dos glóbulos de gordura do leite de cabra e a principal substância de energia para o desenvolvimento físico natural humano (YUAN et al., 2021; ZHANG; JIA; SHI, 2022).

O leite de cabra apresenta vários benefícios a saúde, tornando-o um substituto

potencial para a produção de fórmulas infantis. Além disso, possui constituintes capazes de desempenhar proteção contra infecções, tendo em vista que hidrolisados da caseína melhoram a resistência à insulina, antioxidantes, entre outras (CAKIR; TUNALI-AKBAY, 2021; GONG, H et al., 2020; KOSTIC et al., 2021; OLIVEIRA et al., 2021).

3.3 Controle de qualidade

Para garantir um produto de qualidade se faz necessário ter o controle processual para cada atividade a ser executada. A adoção de manuais de procedimentos operacionais padrão são necessários para a realização de qualquer tipo de atividade a ser desenvolvida. Na área de alimentos, as boas práticas de fabricação são indispensáveis para obtenção de um produto final de qualidade, seguro e que atenda as exigências estabelecidas no mercado, bem como supra os requisitos determinados pela legislação vigente (VANZELLA; SANTOS, 2020).

Além disso, manter um controle de qualidade rigoroso acarreta diversos benefícios aos seus adeptos, tais como: redução dos custos, já que evita desperdícios, aprimoramento dos métodos de produção corroborando, dessa maneira, com a cultura de boas práticas dentro da equipe de trabalho. Segundo Vanzella e Santos (2020) a qualidade, nas últimas décadas, tornou-se um dos temas mais trabalhados pelas organizações, principalmente para comprovarem a qualidade dos seus produtos, visando manterem-se competitivas no mercado nacional e internacional.

Com isso, a busca por métodos simples, baratos e que sejam eficazes na execução da leitura de matrizes complexas estão vez mais procurados pelas grandes empresas com o intuito de otimizar a produção, mas sem perder a agilidade da mesma.

3.4 Adulteração de leite de cabra por adição de leite de vaca

As fraudes em leite tem sido relatadas em todo o mundo, em que tem afetado a qualidade e a segurança desse alimento para a população. Os adulterantes usuais incluem conservantes para aumentar a vida útil do produto, adição de substâncias para aumentar o volume, reconstituintes da densidade e outros que tem a função de mascarar a acidez da fermentação microbiana (ALMEIDA, 2013; FARIA et al., 2022).

Para Cohen (2018) a adulteração alimentar é dividida em categorias e o leite pertence a adulteração econômica, na qual um ingrediente mais barato é adicionado em um produto

de maior valor. Logo, a adulteração do leite de cabra misturando-o com proporções variadas de leite de vaca no leite de cabra é comumente aplicada, uma vez que o valor comercial do leite de cabra e seus respectivos derivados é maior do que o de vaca (RODRIGUES et al., 2012).

Segundo Egito et al. (2006) alguns fatores podem corroborar para este tipo de fraude, entre eles está os eventos ocasionados pela sazonalidade que acabam afetando a disponibilidade do leite de algumas espécies, como o de cabra, tornando-se um incentivo a este tipo de prática.

Para a identificações das adulterações provocadas pela adição de leite bovino no leite caprino podem ser identificadas por divergências nos parâmetros estabelecidos pelo Ministério da Agricultura e Pecuária e Abastecimento (MAPA). De acordo com a ISO 9622 / IDF 141 (2013) teores de gordura e proteínas totais, lactose anidra e sólidos não gordurosos e totais podem ser analisados por medidas de absorção da radiação no infravermelho médio e uso de sensores de condutividade. Ressaltando que não são medidas tão precisas e podem promover degradação da amostra de leite durante a análise (MAPA, 2018).

3.5 Espectrometria no infravermelho próximo (NIR)

A espectrometria é um conjunto de técnicas analíticas que aplica a interação entre radiação eletromagnética em comprimentos de ondas específicos com a matéria para determinar a concentração de substâncias, sendo estes divididos por regiões no espectro (FRANÇA, 2019; SKOOG; HOLLER; CROUCH, 2017).

A região do espectro eletromagnético onde fica a região de radiação infravermelha está dentro da faixa de frequência que varia de $3 \cdot 10^{11}$ até $4 \cdot 10^{14}$ Hz (OLIVEIRA et al., 2015), sendo dividida em 3 diferentes regiões: infravermelho próximo (Near Infrared - NIR), infravermelho médio (Middle Infrared - MID) e infravermelho distante (Far Infrared – FIR).

O infravermelho próximo é uma região do espectro eletromagnético que fica próxima da região visível e seu comprimento de onda vai de 750 a 2500 nm. Para que ocorra a geração de dados a molécula passa para um estado de energia excitado quando absorve radiação infravermelha, como consequência suas vibrações são alteradas, contudo a frequência permanece inalterada (MARIANI et al., 2015; SALA, 1996).

Para que as moléculas absorvam essa radiação, elas precisam apresentar variação no

momento dipolo ($\Delta\mu \neq 0$), assim quando a frequência de vibração da molécula coincide com a da radiação ocorre a absorção de energia. Assim, suas análises são baseadas nas transições das vibrações das ligações C-H, N-H, S-H e O-H (MARIANI et al., 2015; SALA, 1996).

As vantagens inerentes a essa metodologia são a capacidade de previsão rápida, ambientalmente amigável, não destrutiva e não invasiva (PRIETO et al., 2017).

3.6 Quimiometria

A quimiometria surgiu na década de 1970, onde seu nome foi apresentado pela primeira vez em um artigo de Svante Wold. Mas, apenas na década de 1980 este termo se difundiu, devido a uma série de eventos, em especial, uma reunião sediada em Conzeza, Itália, que proporcionou visibilidade ao tema (BRERETON et al., 2017).

As primeiras aplicações da quimiometria ocorreram durante este período, tendo como principal área a química analítica quantitativa. (BRERETON et al., 2017). No Brasil essa ciência começou a ser utilizada a pouco mais de 30 anos e, apesar de ter como seu maior aliado os cursos de química, vem se destacando em outras áreas como no setor de controle de qualidade na indústria farmacêutica e na indústria de alimentos por se tratar de uma técnica altamente interdisciplinar (BRERETON et al., 2018; NETO; SCARMINIO; BRUNS, 2006).

A quimiometria possui diversos métodos que podem ser aplicados com diferentes finalidades, tais como:

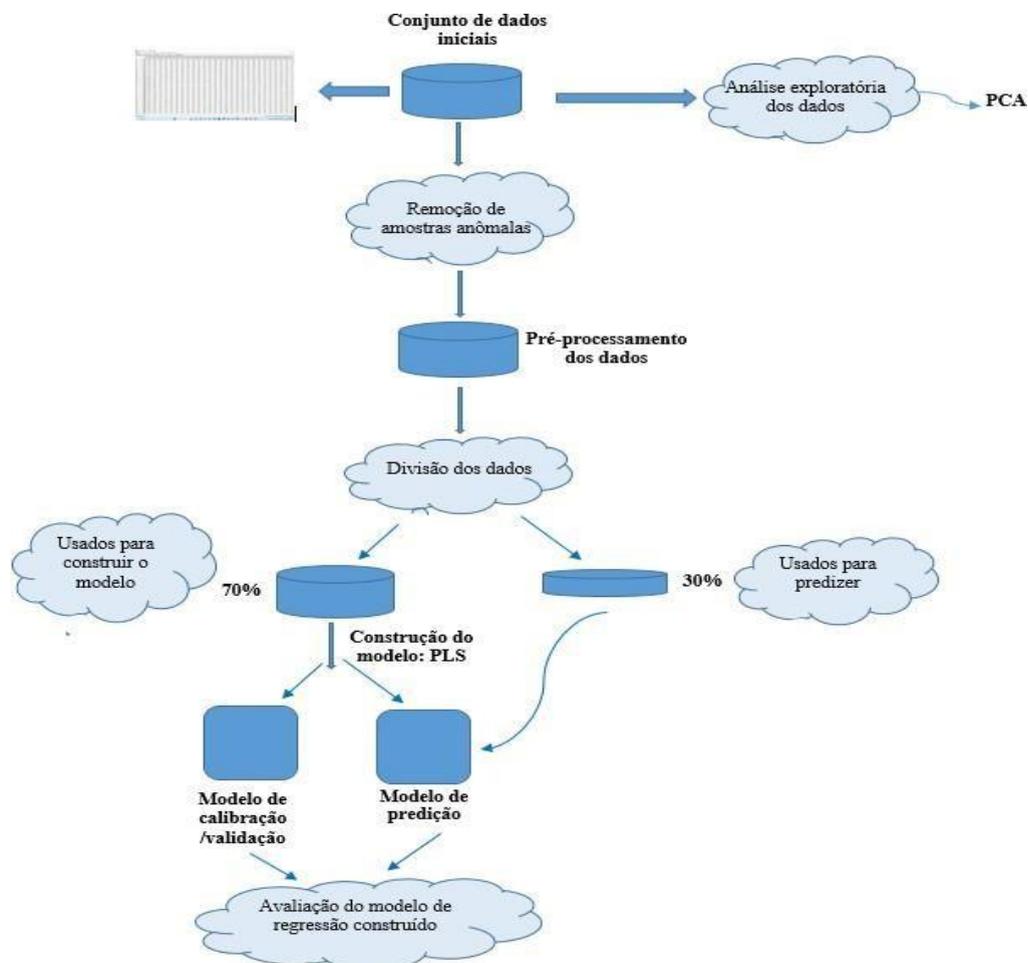
- Análise exploratória de dados;
- Calibração multivariada;
- Planejamento e otimização experimental;
- Reconhecimentos de padrões.

A escolha do método quimiométrico a ser empregado depende de quais informações se deseja extrair do conjunto de dados em estudo. Neste trabalho serão utilizados a análise exploratória dos dados e a calibração multivariada. A Figura 1 mostra de forma esquemática o caminho percorrido para calibração multivariada.

A quimiometria tem a capacidade de otimizar o número de experimentos, estimar possíveis interações entre variáveis, além de selecionar condições experimentais ótimas. Por isso, este conjunto de técnicas denominado quimiometria vem apresentando um crescente número de publicações tendo em vista a sua diversidade de formas para resolver

limitações analíticas, além de fornecer informações através da medição do sinal (TARAPOULOUZI; ORTONE; CINTI, 2022).

Figura 1 - Esquema para construção de calibração multivariada.



Fonte: Elaborada pela autora, 2023.

Medir o sinal significa registrar a magnitude da resposta de um instrumento em função da sua variável determinada pelo domínio da medida. Por exemplo, em espectrometria as informações obtidas pelos espectros são organizadas em uma matriz de dados em que a posição de um pico é determinada pelo comprimento de onda, denominada de variável (FERREIRA, 2015).

A espectrometria associada a quimiometria (análise estatística multivariada) possui várias vantagens, entre elas estão: a interpretação de conjuntos de dados complexos em um tempo de análise consideravelmente reduzido e quantificação simultânea, de forma não destrutiva de vários analitos a partir de uma única medida espectral (DOS SANTOS et al., 2022).

Esta combinação está se tornando cada vez mais investigada nos domínios acadêmico e industrial para atender às crescentes demandas por técnicas de análise rápidas, precisas, econômicas e fáceis de usar que podem ser aplicadas no local quando refere-se a NIR portátil, bem como desenvolvidas em sistemas de monitoramento, otimização e controle de processos (DOS SANTOS et al., 2022).

3.7 Técnicas quimiométricas

3.7.1 Pré - processamentos

Com o objetivo de reduzir e/ou remover variações indesejáveis, a etapa de pré-processamento é fundamental e até mesmo indispensável em alguns casos para obtenção do êxito na realização da análise multivariada. Nesta etapa busca-se o aprimoramento das condições da matriz de dados para que as informações sejam melhor exploradas (RINNAN; VANDENBERG; ENGELSEN, 2009).

Para a matriz dos dados em estudo pode-se utilizar correções de linha de base do espectro, aplicando o método de baseline, derivação ou alisamento pela técnica Savitzky-Golay, que também é capaz de remover ruídos. Para correção de sinal multiplicativo e espalhamento do espectro, pode-se aplicar Variação Normal Padrão (do inglês, *Standard Normal Variate* - SNV) e Correção Multiplicativa de Espalhamento (do inglês, *Multiplicative Scatter Corretion* - MSC).

É indispensável que sejam realizadas de forma cuidadosa e criteriosamente a escolha adequada do pré-processamento, levando em consideração que é um fator essencial para o sucesso de qualquer análise multivariada de dados (FERREIRA, 2015).

3.7.2 Análise de Componentes Principais (PCA)

A análise de componentes principais é uma investigação exploratória, usada para reduzir a dimensionalidade dos dados originais e fornecer uma avaliação mais fácil da relação entre amostras semelhantes (HARROU et al., 2015; MOUFID et al., 2022). A PCA é uma técnica não supervisionada, linear e frequentemente usada em sistemas multisensoriais (HARROU et al., 2015; MOUFID et al., 2022).

Segundo Akulich et al. (2022), a PCA pode ser exemplificada como um espaço que foi transformado, podendo ser imaginado como uma forma de elipsoide de dimensão P,

em que o maior eixo da elipsoide simboliza a direção de variância máxima dentro dos dados para $P < M$, em que o M representa o número das dimensões originais. Dado N número de amostras e uma matriz $N \times M$ de dados X , e $X_i \in \mathbb{R}^M$, para $i = 1, \dots, N$ e $j = 1, \dots, M$, os dados devem estar dispostos de forma padronizada para serem centrados em torno de zero subtraindo o valor médio de cada variável dos pontos de dados individuais em uma determinada dimensão.

A PCA é um ferramenta de análise exploratória que permite evidenciar a existência ou não de amostras anômalas, demonstra as relações entre as variáveis medidas e revela as relações ou agrupamentos entre amostras (LYRA et al., 2010).

As componentes principais (PC) são não correlacionadas e ortogonais entre si e a informação contida em uma delas não está presente em outra. Matematicamente, sua aplicação implica em decompor a matriz de dados (X) em duas outras matrizes: a de *scores* ou escores (T) e a de *loadings* ou pesos (P'), como mostra a equação a seguir (FERREIRA et al., 1999; WOLD et al., 1987).

$$X = TP^T + E \quad \text{Equação 1}$$

Observa-se na equação que a matriz de escores T corresponde às coordenadas das amostras em relação aos eixos das componentes principais (PCs) e possui dimensões $I \times K$. Já a matriz de loadings P , de dimensão $K \times J$ representa a relação entre o eixo da PC e o eixo da variável original, sendo calculado pelo cosseno do ângulo entre esses eixos. Por fim, a variabilidade dos dados não explicada pelo modelo é expressa por uma matriz de resíduos E , que apresenta a mesma dimensão da matriz X , $I \times J$.c (GARCIA, 2021).

3.7.3 Calibração multivariada

O conjunto de calibração multivariada é utilizada com a finalidade de desenvolver um modelo para quantificar propriedades de interesse, em que os mesmos podem ser obtidos pelos sinais instrumentais. Ou seja, o objetivo da calibração multivariada é construir um modelo capaz de estabelecer uma relação dos sinais analíticos obtidos da amostra com uma propriedade específica. O conjunto de calibração vem sendo associada a medições espectrométricas (DIAZ et al., 2021).

A calibração multivariada é empregada para medir, indiretamente, composições químicas de produtos e substâncias de interesse e tem sido uma ferramenta analítica

atrativa na área de controle de qualidade por apresentar bons resultados de forma eficaz e econômica (DIAZ et al., 2021).

Dentre estes métodos, os de regressão mais utilizados está o PLS, devido principalmente à sua ampla gama de aplicações e à sua facilidade no manuseio. Este método executa a calibração a partir da construção de um modelo de regressão que correlaciona as informações de todo espectro com a propriedade a ser determinada (SHAO et al., 2010; SOUSA et al., 2019).

3.7.4 *Mínimos quadrados parciais (PLS)*

O método a regressão PLS é uma técnica muito versátil, pois é usado em diferentes cenários, como em ciência de alimentos, biodiesel e diesel, inovação em alianças cooperativas, aplicações forenses e medicina (MEHMOOD et al, 2012). A regressão por mínimos quadrados parciais é baseado na decomposição de todas as variáveis independentes e dependentes, X e Y, respectivamente, em um produto com duas matrizes menores, mais uma matriz de erro, como descritas nas Equações (2) e (3) (BRERETON, 2003).

$$X = TP^t + E \quad \text{Equação 2}$$

$$Y = UQ^t + F \quad \text{Equação 3}$$

Em que T e U são as matrizes de escores das matrizes X e Y, P e Q são as matrizes de pesos (*loadings*) de X e Y, e E e F são os resíduos de X e Y, respectivamente.

Há duas variantes do método PLS, denominadas como PLS1 e PLS2. Na primeira, cada coluna de y é modelada individualmente, isto é, para cada constituinte de interesse, é calculado um conjunto separado de scores e variáveis latentes. Enquanto que no PLS2, calcula-se um único conjunto de scores e variáveis latentes para todas as colunas da matriz Y, que é comum a todas as propriedades de interesse que estão sendo modeladas (BRERETON, 2018; FERREIRA, 2015).

Como o método PLS maximiza a relação entre a variável dependente e os escores, pode-se afirmar que as variáveis latentes representam as direções que melhor discriminam as classes e que definem a máxima separação entre elas (FERREIRA, 2015). Para isso, modelo PLS ajusta uma função linear dos scores da matriz X (dados instrumentais) aos

scores do vetor y (propriedade de interesse) de modo a minimizar os desvios numa soma de variáveis latentes (VLs) (OLIVEIRA, 2016).

O cálculo das variáveis latentes são feitos a partir de combinações lineares das variáveis originais para predição da propriedade de interesse. Cada variável latente relaciona a matriz X à propriedade de interesse y . Ela é obtida maximizando a covariância entre os escores das matrizes X e Y (T e U) (FERREIRA, 2015).

O número de variáveis latentes utilizadas na construção de um modelo é de extrema importância na obtenção dos resultados. Um modelo contendo um número de VLs menor que o ideal resultará em subajuste, pois não levará em consideração toda informação contida no conjunto de dado e, por outro lado, quanto maior o número de VLs, maior será o erro, resultando em um sobreajuste (*overfitting*). Assim, pode-se dizer que o número ótimo de VLs é aquele no qual a diminuição do erro é compensada pelo aumento de erro de superavaliação (OLIVEIRA, 2016).

4 MATERIAIS E MÉTODOS

4.1 Amostras

As amostras de leite de caprino e bovino foram adquiridas mediante parcerias com pequenos produtores de leite de cabra e vaca pertencentes a municípios da Paraíba. As amostras obtidas foram provenientes de animais de diferentes períodos de lactação e idades, e não houve controle da alimentação ingerida pelas mesmas (PEREIRA et al., 2021). A Tabela 1 apresenta mais detalhes sobre os lotes das amostras.

Tabela 1 - Descrição das amostras de leite de cabra e leite de vaca.

Lote	Tipo de criação	Raça	Localização geográfica	Armazenamento
Caprino 1	Semi-extensivo	Moxotó	Taperoá	Resfriamento
Caprino 2	Semi-extensivo	Moxotó	Taperoá	Resfriamento
Caprino 3	Semi-extensivo	Murciana	Taperoá	Resfriamento
Caprino 4	Semi-extensivo	Alpina Americana	Taperoá	Resfriamento
Caprino 5	Semi-extensivo	Parda	Taperoá	Resfriamento
Caprino 6	Intensivo	Anglo Nubiana	Campina Grande	Resfriamento
Bovino 1	Intensivo	Sidi	Taperoá	Resfriamento
Bovino 2	Intensivo	Guzerá	Taperoá	Resfriamento
Bovino 3	Intensivo	Sidi/Guzerá	Taperoá	Resfriamento
Bovino 4	Extensivo	Girolando	Sobrado	Resfriamento

Fonte: PEREIRA et al., 2021.

As amostras foram provenientes de 6 lotes de leite caprino sendo 5 advindas do município de Taperoá-PB e 1 de Campina Grande-PB. Já o leite de vaca foram adquiridos 4 lotes em que 3 foram adquiridos em Taperoá-PB e 1 em Sobrado-PB (PEREIRA et al., 2021).

As cabras de Taperoá-PB eram criadas em sistemas semi-extensivos e pertenciam as raças: Moxotó, Murciana, Alpina Americana e Parda. As cabras de Campina Grande-PB pertenciam a um sistema de produção intensivo e eram da raça Anglo Nubiana. As vacas de Taperoá-PB eram mantidas em regimes semi-extensivo e intensivo e pertenciam as raças Sindi, Guzerá e Sidi/Guzerá. O lote adquirido em Sobrado-PB foi proveniente de uma vaca da raça Girolando e vivia em regime de criação extensivo. A aquisição de todas as amostras foi realizada no mesmo período (outubro de 2019) (PEREIRA et al., 2021).

As misturas dos leites foram obtidas por meio da adição individual e aleatória dos diferentes lotes de leite de cabra e vaca. Estas misturas ocorreram nos seguintes níveis: 5, 10, 15, 20, 30, 40, 50, 60, e 75 g/100 g. Um total de 10 g para cada mistura foi preparado em duplicatas autênticas. A homogeneização foi realizada em tubos Falcon de 50 mL por inversões sucessivas (PEREIRA et al., 2021).

4.2 Preparo das amostras

As amostras foram coletadas de acordo com o acompanhamento da ordenha, e quando adquiridas seguiram devidamente acondicionadas sendo transportadas com o auxílio de caixa isotérmica sob refrigeração para o Núcleo de Pesquisa e Extensão em Alimentos (NUPEA) da Universidade Estadual da Paraíba (UEPB) de Campina Grande-PB (PEREIRA et al., 2021).

Após a chegada ao laboratório, as amostras foram imediatamente pasteurizadas a uma temperatura de 65°C por um período de 30 minutos e resfriadas a 4°C com o intuito de eliminar os microorganismos presentes e conseqüentemente aumentar conservação das mesmas (PEREIRA, et al., 2021).

4.3 Determinação de lipídios

Para determinação do teor de gordura das amostras foi empregado o método Gerber (2000.18), de acordo com o recomendado pela AOAC *Official Methods*.

4.4 Análises das amostras

4.4.1 NIR portátil

As amostras foram dispostas em placas de Petri de poliestireno de 60 mm × 15 mm e seus espectros NIR adquiridos em triplicata na faixa de 900 a 1650 nm usando um instrumento NIR portátil, modelo DLP NIRscan Nano Evaluation Module (Texas Instruments, Dallas, TX). O registro dos espectros de absorvância foi realizados usando o software integrado, em resolução digital de 210 e integrando 32 varreduras com transformada de Hadamard. Todas as medições foram realizadas na sala temperatura (23 ± 1°C). Um espectro médio de cada amostra e suas triplicatas foi calculado e então

utilizado para a construção dos modelos de regressão. Para o branco foi utilizado uma pastilha de politetrafluoretuleno (PTFE) (PEREIRA, et al., 2021).

4.4.2 Banco de dados

Foi disponibilizado para este estudo um banco de dados de leite caprino contendo 181 amostras e 212 variáveis para os dados brutos. Em que foram selecionadas as variáveis lipídios para posterior análise.

4.4.3 Métodos quimiométricos

Para o tratamento dos dados obtidos pelo NIR, foram utilizadas diferentes ferramentas quimiométricas como a análise de componentes principais (PCA) e regressão de mínimos quadrados parciais (PLS). As modelagens desses métodos foram realizadas utilizando The Unscrambler[®] versão 9.7 e Matlab[®] versão R2016a. A PCA foi aplicada para verificar a possível presença de *outliers* o qual nenhum foi detectado. PLS foi utilizada para construir uma regressão linear dos espectros obtidos.

4.4.4 Seleção das amostras

As amostras dos conjuntos de calibração (136 amostras) e predição (45 amostras) foram selecionadas através do algoritmo SPXy a partir da interface Data Hand Gui, em Matlab[®] versão R2016a (GALVÃO et al., 2005).

4.4.5 Pré-processamentos

Realizou-se pré-tratamentos espectrais no conjunto de dados. Inicialmente, avaliou-se o espectro dos dados brutos além dos seguintes pré-processamentos:

- Dados brutos;
- Variação Normal Padrão (do inglês *Standard Normal Variate* - SNV);
- Correção Multiplicativa de Espalhamento (do inglês, *Multiplicative Scatter Correction* – MSC);
- Baseline: correção linear e offset da linha de base;

- Savitzky-Golay com janela de 15 pontos, polinômio de 1º grau e derivação de primeira ordem (SG J15P1D1);
- Savitzky-Golay com janela de 15 pontos, polinômio de 2º grau e derivação de primeira ordem (SG J15P2D1);
- Savitzky-Golay com janela de 15 pontos, polinômio de 2º grau e derivação de segunda ordem (SG J15P2D2);
- Savitzky-Golay com janela de 19 pontos, polinômio de 1º grau e derivação de primeira ordem (SG J19P1D1);
- Savitzky-Golay com janela de 19 pontos, polinômio de 2º grau e derivação de primeira ordem (SG J19P2D1);
- Savitzky-Golay com janela de 19 pontos, polinômio de 2º grau, e derivação de segunda ordem (SG J19P2D2);
- Savitzky-Golay com janela de 23 pontos, polinômio de 1º grau, e derivação de primeira ordem (SG J23P1D1);
- Savitzky-Golay com janela de 23 pontos, polinômio de 2º grau, e derivação de primeira ordem (SG J23P2D1);
- Savitzky-Golay com janela de 23 pontos, polinômio de 2º grau, e derivação de segunda ordem (SG J19P2D2);

Para realização dos pré-processamentos foi utilizado software The Unscrambler® versão 9.7.

4.4.6 Construção do modelo

Após a obtenção do banco de dados, espectros NIR, e parâmetros de interesse, os dados brutos (*full spectra*) e os dados pré-processados foram construídos os conjuntos de calibração e predição através de algoritmo de seleção de amostras SPXy, com posterior seleção de variáveis e ao final a construção dos modelos de regressão PLS.

Os modelos de PLS foram avaliados por meio do coeficiente de determinação (R^2), pelo erro quadrático médio de calibração (do inglês *Root Mean Square Error of Calibration* - RMSEC) e raiz quadrada do erro médio quadrático de validação cruzada (do inglês *Root Mean Square Error of Cross Validation* - RMSECV). Também foram projetadas a capacidade preditiva, em que foi avaliado através do coeficiente de determinação (R^2), o erro quadrático médio de predição (do inglês *Root Mean Square*

Error of Prediction - RMSEP) e variáveis latentes.

5 RESULTADOS E DISCUSSÃO

5.1 Avaliação das propriedades de interesse

Foi realizada, inicialmente, uma análise das medidas de tendência central e de dispersão das propriedades de interesse a fim de observar de maneira preliminar, o comportamento do conjunto de dados obtidos por meio dos experimentos físico-químicos, indicados na Tabela 3.

Tabela 2 – Medidas de tendência central e de dispersão das propriedades de interesse das amostras estudadas.

Propriedade de interesse	Medidas	Valores
Percentual da mistura do leite de vaca em leite de cabra (v/v%)	Mínimo	0
	Máximo	100
	Média	45,0079
	Desvio Padrão	28,7424
Lipídios (g/100g)	Mínimo	3,5700
	Máximo	5,5500
	Média	4,7650
	Desvio Padrão	0,3643

Fonte: Elaborada pela autora, 2023.

No parâmetro de misturas de leite de vaca em leite de cabra, valores compreendidos entre 0 a 100% foram considerados para o este estudo em decorrência a evolução da adição do leite de vaca ao leite de cabra. Em que 0% significa leite puro caprino e 100% leite puro de vaca.

Sabe-se que vários fatores podem contribuir para variação do conteúdo lipídico das misturas estudadas e até mesmo de amostras analisadas de maneira isoladas, tais como raça, sazonalidade, alimentação, período de lactação e a região de criação (LI et al., 2017; LIU et al., 2020). Assim, os valores obtidos de lipídios que compreendeu entre 3,5700 a 5,5500 gramas foram considerados para o estudo.

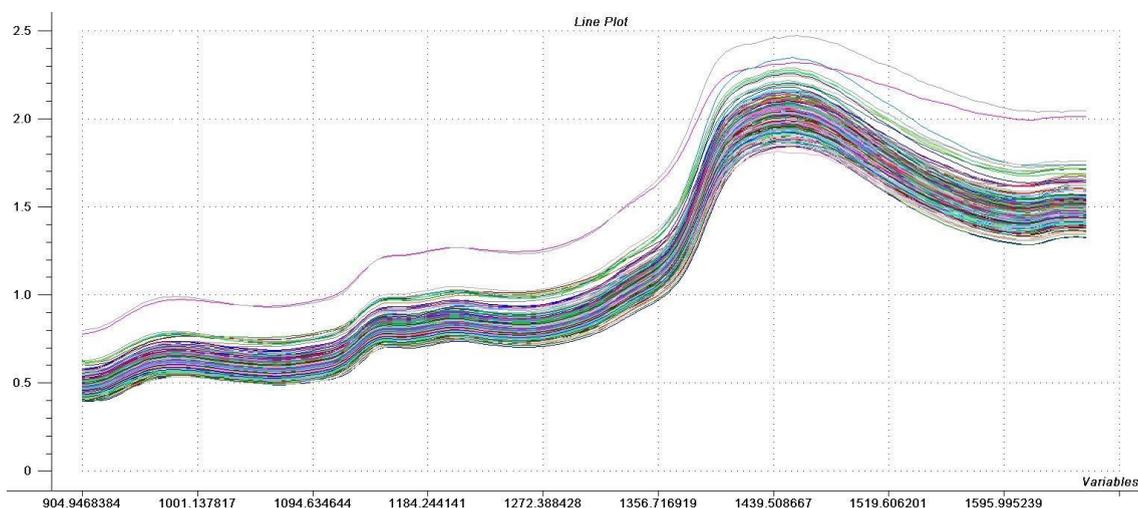
A média obtida para a propriedade de interesse lipídios observa-se valores dentro do esperado. Em que estipulasse uma média 3,80 a 4,99 g para lipídios (FERNANDES et al., 2008; TACO, 2011). Vale ressaltar que este valor de lipídios está sendo considerado

para os leites classificados como integrais. Em relação ao desvio padrão, apresentou número decimal para lípidios, sendo considerado um valor baixo para esta propriedade de interesse.

5.2 Análise espectral

As misturas analisadas foram obtidas com a adição de leite de vaca em leite de cabra em diferentes proporções, variando entre 5 e 75%. Os espectros foram adquiridos utilizando infravermelho próximo no equipamento portátil, na faixa espectral de 904 a 1650 nm. A Figura 2 apresenta o comportamento dos dados brutos (sem pré-processamento) referente as 181 amostras estudadas.

Figura 2 - Comportamento dos dados brutos no espectro do infravermelho próximo (NIR).



Fonte: Elaborada pela autora, 2023.

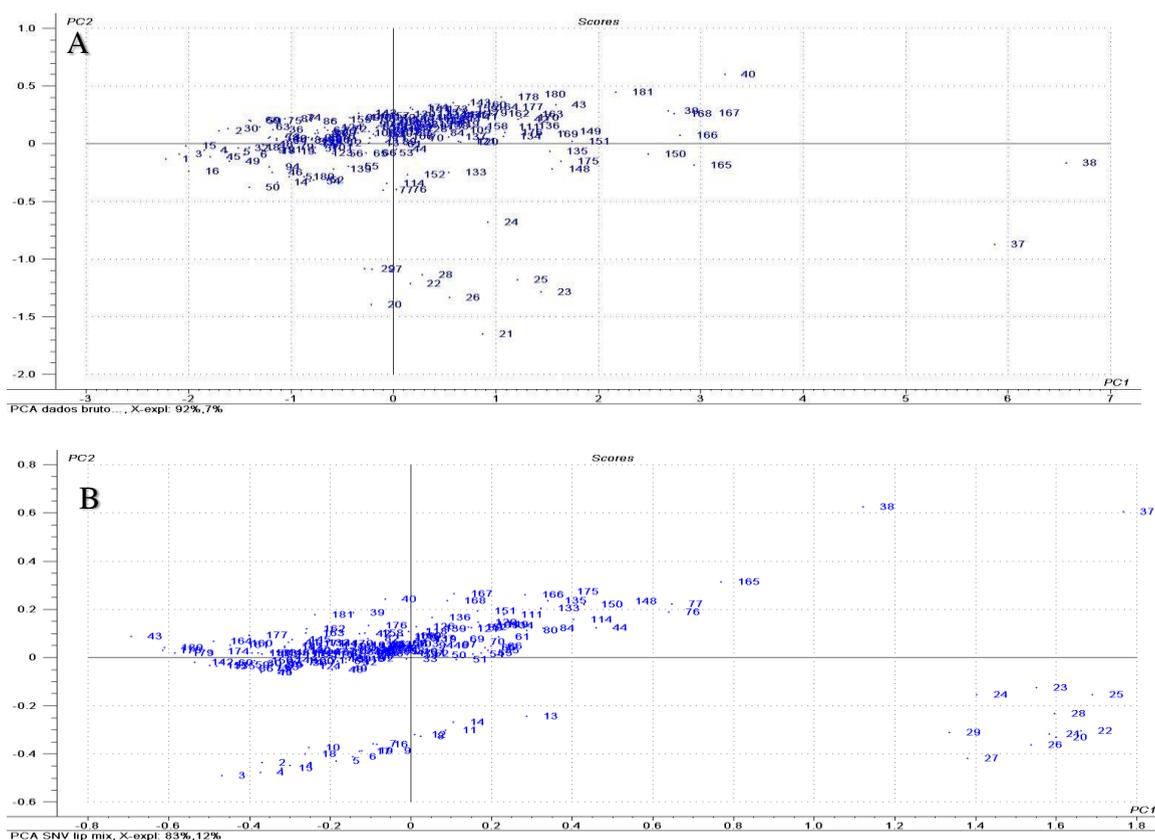
Analisando o perfil químico das amostras de percentual de misturas de leite de vaca em leite de cabra observa-se uma banda proeminente por volta da região de 1450 nm caracterizando o primeiro sobretom do estiramento vibracional da ligação O-H correspondente a água. Nota-se a presença de outras bandas menores tais como na região aproximada de 1630 nm em que há combinação harmônica de estiramento de N-H livres e amidas secundárias indicando a presença de proteínas nas amostras. Além disso podemos observar outra banda de menor intensidade de sinal na região espectral de 1100 e 1200 nm correspondendo ao segundo sobretom e estiramento da ligação -CH e designando a presença de gordura que é a propriedade de interesse do estudo.

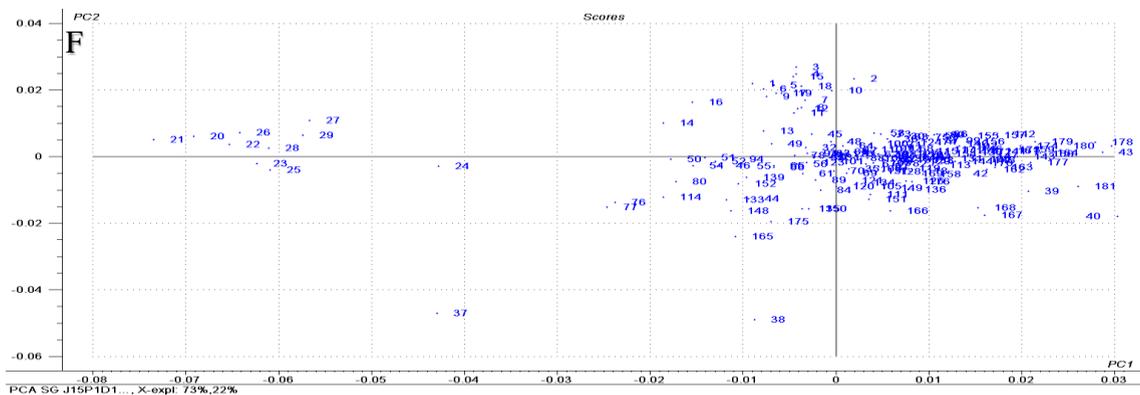
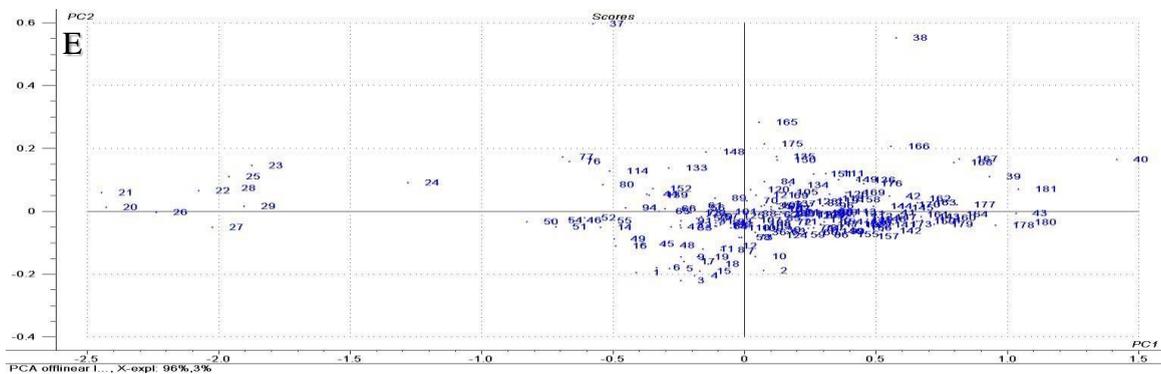
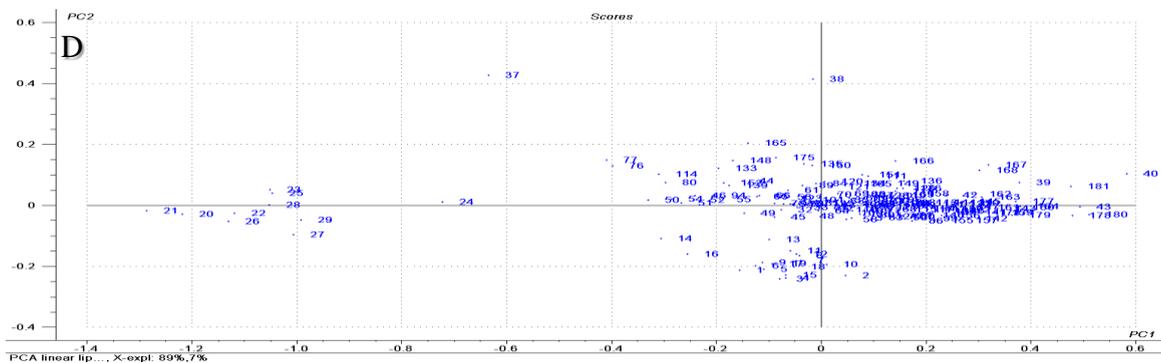
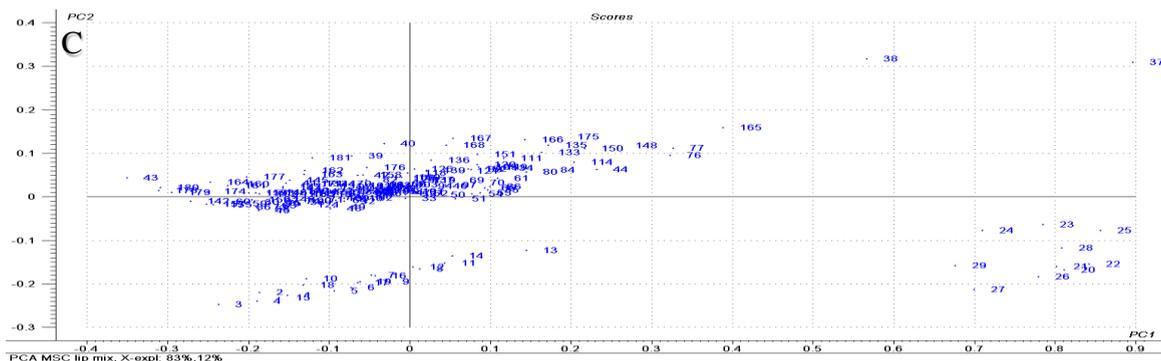
Explorando o comportamento dos espectros de forma analítica, observa-se que não há presença de ruídos, permitindo que toda a faixa espectral escolhida fosse considerada para o estudo em relação a linha de base, nota-se um deslocamento, provavelmente decorrente da variabilidade da composição e características físico-química das amostras, além de conter espalhamento dos espectros.

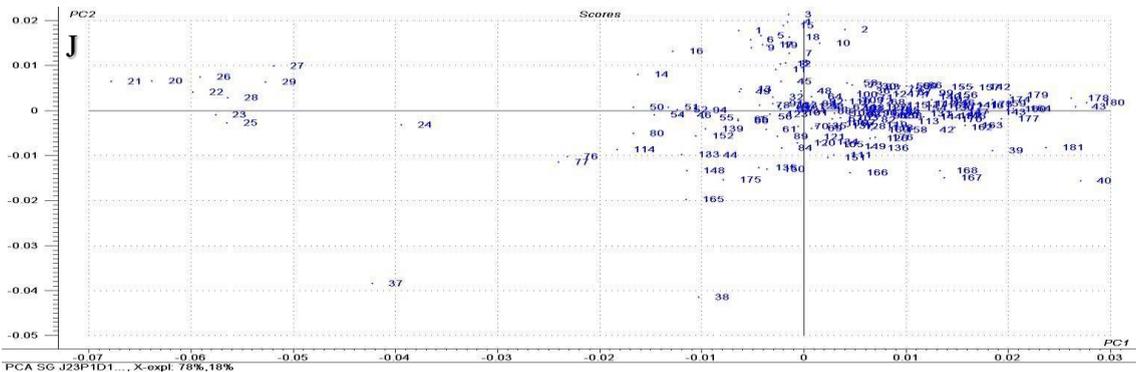
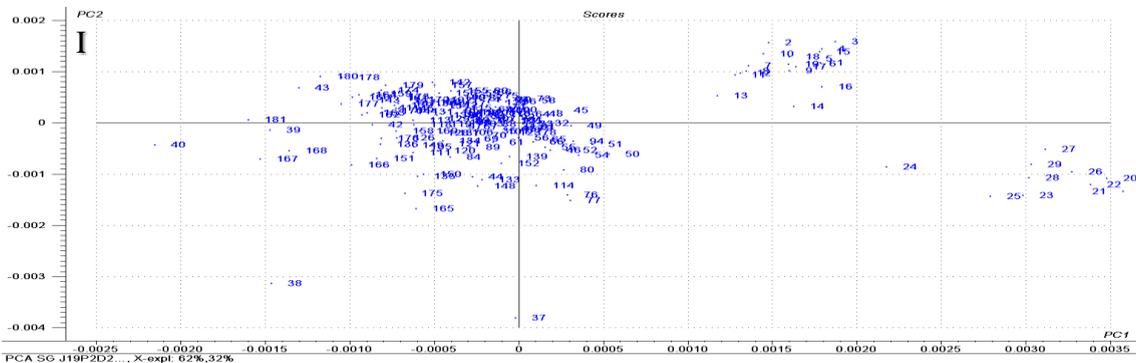
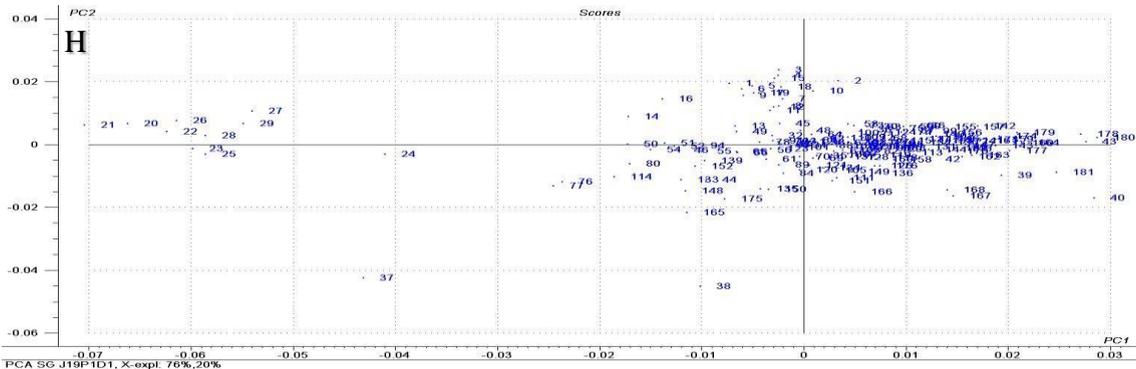
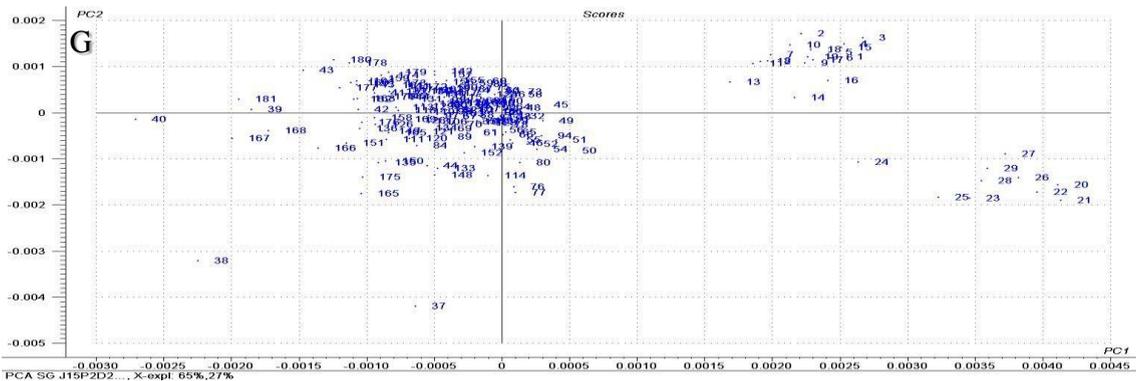
5.3. Análise exploratória das propriedades de interesse

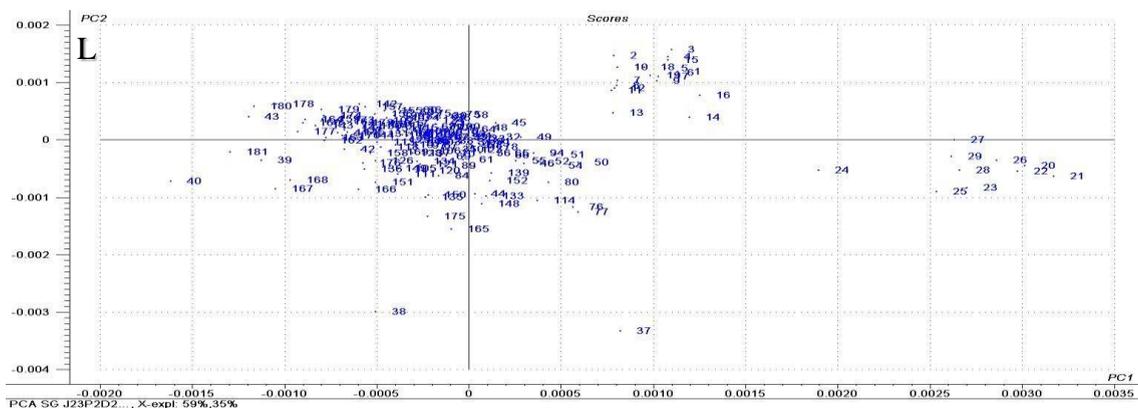
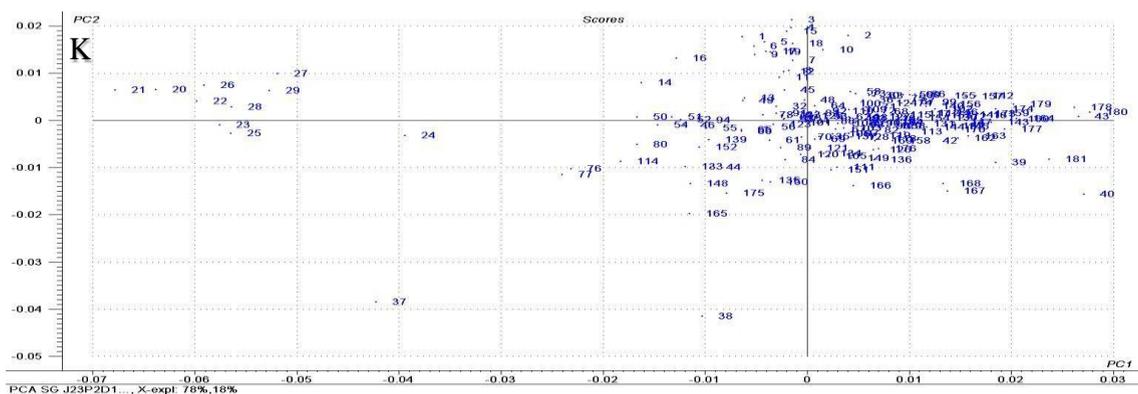
Com a finalidade de realizar uma análise exploratória e observar o comportamento dos dados, foi aplicada a técnica de reconhecimento de padrão não supervisionado, Análise de Componentes Principais (do inglês *Principal Component Analysis* - PCA), O método foi utilizado para verificar as tendências de todas as amostras de leite em relação às faixas de concentrações e possíveis semelhanças e/ou diferenças. Este procedimento foi aplicado para tanto para os dados brutos como os pré-processados. A Figura 3 apresenta o comportamento das amostras de leite dos dados brutos e pré-processados no gráfico de scores em PC1 por PC2.

Figura 3 - Gráficos de escores dos dados brutos e pré-processados das amostras estudadas.









Legenda: A - Gráfico referente aos dados brutos, B - gráfico pré-processado com SNV, C - MSC, D - baseline offlinear, E - baseline offset, F - SG J15P1D1, G - SG J15P2D2, H - SG J19P1D1, I - SG J19P2D2, J - SG J23P1D1, K - SG J23P2D1, L - SG J23P2D2.

Fonte: Elaborada pela autora, 2023.

Foram aplicados pré-processamentos de acordo com as características apresentadas pelo comportamento dos dados brutos, visando obter melhores resultados ao final da realização dos mesmos. Logo, para correção do efeito multiplicativo dos espectros foi escolhido o SNV e MSC. Para correção de linha de base foram selecionados baseline offlinear e offset e derivações de Savitzky-Golay janelas 15, 19 e 23, polinômios 1 e 2 e 1ª e 2ª derivada.

No desenvolvimento do estudo observou-se que os pré-processamentos de SG J15P1D1 e SG J15P2D1 apresentaram resultados iguais tanto para o comportamento das amostras no gráfico de PCA assim como nas métricas obtidas na construção do modelo e por isso só foi exposto neste trabalho a primeira opção. O mesmo ocorreu também com os pré-processamentos SG J19P1D1 e SG J19P2D1 em que foi aplicada a mesma metodologia citada anteriormente.

A PCA é uma análise que permite detectar amostras que apresentam um comportamento distinto das outras, pois a projeção de dados permite evidenciar as amostras atípicas e o gráfico de escores expressam as relações entre as amostras.

(FERREIRA, 2015). Logo, observa-se nos gráfico acima que não apresentam nenhuma amostra considerada anômala (*outliers*) nos dados brutos e nem em nenhum pré-processamento.

Os gráficos de escores dos dados brutos mostra que PC1 explica 92% e PC2 7% da variância dos dados. Observa-se também a similitudes dos gráfico dos pré-processamentos escolhidos para correção do efeito multiplicativo, tanto o SNV como MSC apresentaram a variância explicada para PC 1 de 83% e PC 2 de 12%. Para correção de linha de base os maiores e menores valores de PC1 e PC2 variaram entre 96% e 3% de varância explicada para o modelo baseline offlinear e 59% e 35% para o modelo de SG J23PD2.

5.4 Modelos PLS para misturas do leite de vaca em leite de cabra

5.4.1 Modelo PLS para dados brutos nas misturas do leite de vaca em leite de cabra

O desenvolvimento dos modelos de quantificação foi realizado utilizando o software unscrambler[®] por meio da aplicação do método PLS. Inicialmente foi desenvolvido o modelo utilizando apenas os dados brutos da amostra, com o intuito de observar os valores obtidos e realizar comparações aos resultados dos pré-processados. A Tabela 4 apresenta os resultados dos dados originais das amostras das misturas do leite de vaca em leite de cabra.

Tabela 4 - Resultado das figuras de mérito dos dados brutos das amostras das misturas de leite vaca em leite cabra.

MISTURAS	
RMSEC	12.84192
R²_{CALIBRAÇÃO}	0.823454
RMSEP	9.116916
R²_{PREDIÇÃO}	0.781366
RMSECV	13.39807
R²_{VALIDAÇÃO}	0.810647
VL	3

Fonte: Elaborada pela autora, 2023.

Através da análise dos dados brutos apresentados, nota-se resultados considerado altos para as figuras de mérito RMSEC, RMSEP e RMSECV. De acordo com Mochão (2015) são parâmetros importantes que contém informações valiosas do modelo. O RMSEC indica o ajuste do modelo aos dados de calibração, enquanto que o RMSEP é utilizado para testar o desempenho final do modelo. E RMSECV utiliza um conjunto de calibração para em simultâneo calibrar e validar o método. Logo, um bom modelo deverá ter baixos valores.

Sabe-se que o coeficiente de correlação (R^2) trata-se de uma medida de ajuste linear, que quando mais próximo de 1 melhores os resultados obtidos. De acordo com Elfadl, Reinbrechta e Claupein (2010), pode ser considerado resultados satisfatórios a partir de valores de R^2 superiores a 0,83, logo, o R^2 de calibração apresentou o mais próximo neste parâmetro. Quando refere-se a variáveis latentes, observa-se um valor baixo. Para Wold, Sjöström e Eriksson (2001), o número de variáveis latentes na construção de um modelo visa favorecer um menor erro de previsão.

De modo geral, o modelo construído para os dados brutos apresentou resultados consideráveis, tendo em vista que os valores obtidos das métricas RMSEC, RMSECV e RMSEP apresentaram próximas, assim como os coeficientes de correlação dos mesmos.

5.4.2 Modelo PLS para dados pré-processados nas misturas do leite de vaca em leite de cabra

Para obter melhores resultados na elaboração dos modelos, foram empregados pré-processamentos distintos os quais foram escolhidos a partir da observação do comportamento dos dados nas etapas anteriores deste trabalho. Assim, foram gerados 11 modelos de calibração para determinação de lipídeos nas amostras do percentual das misturas de leite de vaca em leite de cabra.

Foram considerados os melhores modelos de regressão PLS, aqueles que apresentarem o menor número de variáveis latentes, os maiores valores de $R^2_{\text{calibração}}$, $R^2_{\text{validação}}$, $R^2_{\text{predição}}$ e menores valores de RMSEC, RMSECV, RMSEP. A Tabela 5 apresenta os valores das figuras de mérito obtidas na construção dos modelos.

Analisando os resultados referentes aos percentuais de misturas, observa-se que os valores do RMSEC, RMSEP e RMSECV variaram entre 14.06037, 14.35195 e 15.40805 sendo os maiores valores obtidos e 10.17333, 10.04243 e 13.87954 os menores valores encontrados respectivamente. No entanto, de modo geral, os valores dessas três figuras

de mérito foram considerados altos.

Os valores do coeficiente de correlação, seguindo a ordem de maior ao menor valor, apresentaram os seguintes resultados para calibração (R^2) que variaram de 0.887480 a 0.823368. Já para predição o R^2 apresentou valores que compreenderam entre 0.755265 a 0.548850 e validação obteve resultados de 0.787935 a 0.740072. Segundo Martins (2014) O coeficiente de correlação não é uma medida resistente, logo podem sofrer variações por ser influenciado pelos dados em estudo, por isso deve ser interpretado com o devido cuidado. Para variáveis latentes, observa-se uma variação em que obteve como maior valor 15 e menor 6.

Analisando os modelos construídos com pré-processamento espectral e levando em consideração todas as figuras de mérito de forma simultânea, nota-se que o SG J15P2D2 e SG J23P2D2 obtiveram melhores resultados. No entanto, apesar do modelo de SG J15P2D2 apresentar maior RMSEC e RMSECV que o modelo de SG J23P2D2, este último exibe melhores figuras de mérito de predição, com maior $R^2_{\text{predição}}$ e menor RMSEP e número de variáveis latentes. Sabe-se que o objetivo da predição é testar se o conjunto de calibração consegue prever as concentrações das amostras de predição. Logo, quanto menor o valor do RMSEP, maior o percentual de predição dos modelos construídos.

Nota-se também que o modelo utilizando pré-processamento de SG J15P2D2 apresentou 15 variáveis latentes, sendo considerado o maior valor obtido neste parâmetro. Segundo Ferreira; Montanari e Gaudio (2002) quanto mais alto o número de variáveis latentes, maior a chance de ocorrer *overfitting*, que consiste na obtenção de valor elevado do coeficiente de correlação decorrente do número excessivo de variáveis incluídas no modelo e não de seu ajuste natural aos valores observados da atividade. Em resumo, *overfitting* pode ser considerado como ajuste forçado.

Comparando os modelos pré-processados ao modelo construído com os dados brutos, percebe-se que tanto utilizando SG J15P2D2 quanto SG J23P2D2 apresentam resultados menores para RMSEC e maiores $R^2_{\text{calibração}}$ do que os dados brutos. No entanto, quando trata-se do RMSEP e RMSECV observa-se valores menores para os dados brutos, assim como o $R^2_{\text{predição}}$ e $R^2_{\text{validação}}$. De modo geral nota-se que os modelos em que foram pré-processados apresentaram poucos ajustes em algumas figuras de mérito quando comparados aos modelos construídos dos dados brutos.

Tabela 5 - Resultados das figuras de mérito dos modelos referentes a determinação do percentual de misturas de leite de vaca em leite de cabra das amostras pré-processadas.

	RMSEC	R²_{CALIBRAÇÃO}	RMSEP	R²_{PREDIÇÃO}	RMSECV	R²_{VALIDAÇÃO}	VL
SNV	1.219.476	0.833859	1.132.534	0.755265	1.387.954	0.787935	7
MSC	1.268.925	0.820112	1.202.321	0.720882	1.528.921	0.742671	8
BASELINE OFFLINEAR	1.359.222	0.793410	1.235.228	0.699683	1.535.914	0.740072	6
BASELINE OFFSET	1.391.461	0.789772	1.273.910	0.622664	1.540.805	0.746000	7
SG J15P1D1	1.406.037	0.784003	1.435.195	0.548850	1.492.692	0.760125	5
SG J15P2D2	1.017.333	0.887480	1.065.521	0.723279	1.413.734	0.785894	15
SG J19P1D1	1.440.390	0.772747	1.427.006	0.558430	1.511.521	0.753414	5
SG J19P2D2	1.206.378	0.841757	1.156.918	0.673826	1.476.441	0.766450	9
SG J23P1D1	1.275.018	0.823368	1.163.596	0.675803	1.468.291	0.769192	8
SG J23P2D1	1.266.270	0.825861	1.163.596	0.675803	1.457.036	0.772894	8
SG J23P2D2	1.219.398	0.840769	1.004.243	0.753108	1.459.629	0.775192	8

Legenda: R² - coeficiente de correlação, RMSEC - raiz quadrada do erro médio quadrático de calibração, RMSEP - raiz quadrada do erro médio quadrático de previsão, RMSECV - raiz quadrada do erro médio quadrático de validação cruzada, VL - variáveis latentes, SG - Savitzky Golay, J - janela, P – polinômio, D – derivada.

Fonte: Elaborada pela autora, 2023.

De acordo com Breitzkreitz, Souza e Poppi (2014), é importante atentar-se ao ajuste do modelo empregado, pois este delega ao algoritmo a função estimar os parâmetros do modelo. Assim, é necessário que o modelo esteja devidamente ajustado para que ele represente adequadamente os pontos experimentais.

5.5 Modelos PLS para determinação de lipídios

5.5.1 Modelo PLS para dados brutos na determinação de lipídios

Para a construção dos modelos de propriedade de interesse concentração lipídios, foram utilizados os mesmos critérios e modelos descritos para as misturas de leite de cabra em leite de vaca (*item 5.4.1*). A Tabela 6 apresenta os valores obtidos para as concentrações de lipídios.

Com base nos resultados obtidos, vê-se que os modelos construídos a partir dos dados brutos não foram capazes de prever adequadamente a concentração de lipídeos para dados de infravermelho próximo. Logo, nota-se baixos valores dos coeficientes de correlação para os modelos. De acordo com Luz (2003), valores de R^2 baixos é decorrente principalmente a conjunto de dados que apresentam pouca variação nos valores reais usados na construção.

Tabela 6 - Resultado das figuras de mérito dos dados brutos das amostras de concentração de lipídios.

MISTURAS	
RMSEC	0.313122
$R^2_{\text{CALIBRAÇÃO}}$	0.361501
RMSEP	0.232212
$R^2_{\text{PREDIÇÃO}}$	0.181610
RMSECV	0.344261
$R^2_{\text{VALIDAÇÃO}}$	0.239502
VL	6

Fonte: Elaborada pela autora, 2023.

5.5.2 Construção dos modelos de PLS para lipídios com espectros pré-processadas

Na Tabela 7 são apresentados os resultados das figuras de méritos obtidas e os pré-processamentos realizados nas amostras leite de cabra para determinação da concentração de lipídios por espectroscopia NIR portátil.

Com base nos resultados obtidos RMSEC, RMSEP e RMSECV apresentaram maiores valores respectivamente 0.352400, 0.289003 e 0.358074 e 0.280168, 0.240506 e 0.328078 os menores valores.

Para $R^2_{\text{calibração}}$, os resultados variaram de 0.464545 a 0.172949. Em $R^2_{\text{predição}}$ os valores obtidos foram entre 0.153119 e 0.011547 e $R^2_{\text{validação}}$, os resultados foram de 0.276513 a 0.158616.

Observa-se que os resultados obtidos dos dados pré-processados foram satisfatórios para calibração, predição e validação, no entanto quando comparados ao modelo construídos dos dados brutos verifica-se que os mesmos sofreram pouco ou nenhum ajuste.

Ao analisar os valores dos coeficientes de correlação, nota-se que os resultados apresentados foram baixos. Este fato pode estar associado a não utilização de uma análise de referência, a exemplo da Cromatografia Líquida de Alta Eficiência, que consequentemente pode ter afetado a capacidade preditiva do modelo, pois o PLS utiliza diretamente a concentração e correlaciona com as informações espectrais (FERREIRA, et al., 2019).

Tabela 7 - Resultado das figuras de mérito para os espectros pré-processados na determinação da concentração de lipídios pré-processadas.

	RMSEC	R²_{CALIBRAÇÃO}	RMSEP	R²_{PREDIÇÃO}	RMSECV	R²_{VALIDAÇÃO}	VL
SNV	0.280168	0.464545	0.267507	0.153119	0.328078	0.276513	7
MSC	0.341570	0.204120	0.289003	0.011547	0.34868	0.191267	1
BASELINE OFFLINEAR	0.325009	0.317369	0.240506	0.087676	0.354591	0.199355	5
BASELINE OFFSET	0.319094	0.319697	0.267515	0.068874	0.330212	0.282139	5
SG J15P1D1	0.318738	0.322268	0.258344	0.130532	0.349989	0.194827	5
SG J15P2D2	0.329888	0.286746	0.270944	-0.060365	0.357066	0.176624	4
SG J19P1D1	0.352400	0.172949	0.268517	0.051722	0.358074	0.158616	1
SG J19P2D2	0.329473	0.288237	0.270801	-0.046926	0.353374	0.193218	4
SG J23P1D1	0.310774	0.357397	0.259312	0.108915	0.346890	0.211092	6
SG J23P2D1	0.352014	0.175533	0.269544	0.03720	0.357606	0.161596	1
SG J23P2D2	0.330001	0.278282	0.269344	0.020312	0.350275	0.198792	4

Legenda: R² - coeficiente de correlação, RMSEC - raiz quadrada do erro médio quadrático de calibração, RMSEP - raiz quadrada do erro médio quadrático de previsão, RMSECV - raiz quadrada do erro médio quadrático de validação cruzada, VL - variáveis latentes, SG - Savitzky Golay, J - janela, P – polinômio, D – derivada.

Fonte: Elaborada pela autora, 2023.

6 CONCLUSÃO

A técnica de NIR portátil utilizada para desenvolvimento dos modelos de calibração por PLS se mostrou mais promissora para determinação do percentual de adulteração de leite de cabra por adição de leite de vaca do que para a quantificação dos lipídios nas amostras estudadas.

Para o percentual de misturas de leite de vaca em leite de cabra o modelo SG J23P2D2 obteve o melhor resultado quando comparado aos outros modelos pré-processados. Já em comparação ao modelo construído com os dados brutos, observa-se que este apresenta resultados melhores para RMSEP, coeficiente de correlação e variáveis latentes.

Em relação ao modelo desenvolvido para concentração lipídios, tanto para os modelos com pré-processamentos quanto brutos, não foram apresentados resultados satisfatórios nas figuras de mérito analisada. Quando comparados aos dados brutos, o modelo utilizando SNV apresentou melhores valores para RMSEC, RMSECV e seus respectivos coeficientes de correlação. No entanto, os dados brutos apresentaram RMSEP e R^2 de predição melhores.

É de extrema importância ressaltar que as principais vantagens dos métodos aqui evidenciado, consiste na facilidade e agilidade na execução da análise como também na possibilidade de quantificá-lo por meio do modelo PLS, e que apesar da falta de ajuste em alguns casos, ainda é possível explorar este banco de dados com novos algoritmos para a obtenção de resultados mais satisfatórios.

REFERÊNCIAS

ADINEPOUR, F.; POURAMIN, S.; RASHIDINEJAD, A.; JAFARI, S. M. Fortification/enrichment of milk and dairy products by encapsulated bioactive ingredients. **Food Research International**, p. 111212, 2022. DOI:10.1016/j.foodres.2022.111212.

AGRIMONTI, C.; PIRONDINI, A.; MARMIROLI, M.; MARMIROLI, N. A quadruplex PCR (qPCR) assay for adulteration in dairy products. **Food Chemistry**, v. 187, p. 58-64, 2015. DOI: 10.1016/j.foodchem.2015.04.017.

AKINYMI, M. O.; BRAUN, D.; WINDISCH, P.; WARTH, B.; MORE, C. N. E. Assesment of muliple mycotoxins in raw milk of three different animal species in Nigeria. **Food Control**, v. 131, p. 108258, 2022. DOI: 10.1016/j.foodcont.2021.108258.

AKULICH, F.; ANAHIDEH, H.; SHEYYAB, M.; AMBRE, D. Explainable predictive modeling for limited spectral data. **Chemometrics and Intelligent Laboratory Systems**, v. 225, p. 104572, 2022. DOI: 10.1016/j.chemolab.2022.104572.

ALMEIDA, T. V. Detecção de adulteração em leite: análises de rotina e espectroscopia de infravermelho. **Seminário apresentado ao Curso de Mestrado em Ciência Animal da Escola de Veterinária e Zootecnia da Universidade Federal de Goiás, Goiânia**, 2013.

BREITKREITZ, M. C.; SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para planejamento de experimentos: avaliação das condições experimentais na determinação espectrofotométrica de ferro II com o-fenantrolina. Um tutorial, parte III. **Química Nova**, v. 37, p. 564-573, 2014. DOI: 10.5935/0100-4042.20140092.

BRERETON, R. G. **Chemometrics: Data Analysis for the Laboratory and Chemical Plant**. John Wiley & Sons, USA, 2003. 489 p.

BRERETON, R. G. **Chemometrics: data driven extraction for science**. 2nd. Hoboken: John Wiley & Sons, 2018.

BRERETON, R. G.; JANSEN, J.; LOPES, J.; MARINI, F.; POMERANTSEV, A.; RODIONOVA, O.; ROGER, J. M.; WALCZAK, B.; TAULER, R. Chemometrics in analytical chemistry—part II: modeling, validation, and applications. **Analytical and Bioanalytical Chemistry**, v. 410, n. 26, p. 6691-6704, 2018.

BRERETON, R. G.; JANSEN, J.; LOPES, J.; MARINI, F.; POMERANTSEV, A.; RODIONOVA, O.; ROGER, J. M.; WALCZAK, B.; TAULER, R. Chemometrics in analytical chemistry—part I: history, experimental design and data analysis tools. **Analytical and Bioanalytical Chemistry**, v. 409, n. 25, p. 5891-5899, 2017.

CAKIR, B.; TUNALI-AKBAY, T. Potential anticarcinogenic effect of goat milk-derived bioactive peptides on HCT-116 human colorectal carcinoma cell line. **Analytical Biochemistry**, v. 622, p. 114166, 2021. DOI: 10.1016/j.ab.2021.114166.

CHEN, X.; WANG, J.; STEVENSON, R. J.; ANG, X.; PENG, Y.; QUEK, V. Y. A

promising way to alter flavor notes of goat milk products. **Lebensmittel Wissenschaft Technologie**, v. 145, p. 111286, 2021. DOI: 10.1016/j.lwt.2021.111286.

COHEN, P. A. The FDA and adulterated supplements—dereliction of duty. **JAMA network open**, v. 1, n. 6, p. e183329-e183329, 2018. DOI: 10.1001/jamanetworkopen.2018.3329.

DELGADO JÚNIOR, I. J.; STOCK, L. A.; SIQUEIRA, K. B. **Produção, composição e processamento de leite de cabra no Brasil**. 2020.

DIAZ, V. F.; KETELAERE, B.; AERNOUTS, B.; SAEYS, W. Cost-efficient unsupervised sample selection for multivariate calibration. **Chemometrics and Intelligent Laboratory Systems**, v. 215, p. 104352, 2021. DOI: 10.1016/j.chemolab.2021.104352.

DI PINTO, A.; TERIO, V.; MARCHETTI, P.; BOTTARO, M.; MOTTOLA, A.; BOZZO, G.; BONERBA, E.; CECI, E.; TANTILLO, G. DNA-based approach for species identification of goat-milk products. **Food Chemistry**, v. 229, p. 93-97, 2017. DOI: 10.1016/j.foodchem.2017.02.067.

DOS SANTOS, I.; BOSMAN, G.; ALEIXANDRE, J. L.; TOIT, W. Direct quantification of red wine phenolics using fluorescence spectroscopy with chemometrics. **Talanta**, v. 236, p. 122857, 2022. DOI: 10.1016/j.talanta.2021.122857.

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, p. 932-939, 2006. DOI: 10.1590/S0102-09352006000500032.

ELFADL, E.; REINBRECHT, C.; CLAUPEINA, W. Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide safflower germplasm collection. **Food Chemistry**, v. 351, p. 129310, 2010.

FERNANDES, M. F.; QUEIROGA, R. C. R. E.; MEDEIROS, A. N.; COSTA, R. G.; BOMFIM, M. A. D.; BRAGA, A. A. Características físico-químicas e perfil lipídico do leite de cabras mestiças Moxotó alimentadas com dietas suplementadas com óleo de semente de algodão ou de girassol. **Revista Brasileira de Zootecnia**, v. 37, p. 703-710, 2008. DOI: 10.1590/S1516-35982008000400017.

FARIA, I. D. L.; GOUVÊIA, M. M.; NETTO, A. D. P.; MARQUES, F. F. C. Determination of formaldehyde in bovine milk by micellar electrokinetic chromatography with diode array detection. **Lebensmittel Wissenschaft Technologie**, v. 163, p. 113473, 2022. DOI: 10.1016/j.lwt.2022.113473.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Quimiometria I: calibração multivariada, um tutorial. **Química Nova**, v. 22, p. 724-731, 1999. DOI: 10.1590/S0100-40421999000500016.

FERREIRA, M. M. C. **Quimiometria: conceitos, métodos e aplicações**. Campinas: Editora Unicamp, 2015.

FERREIRA, M. M. C.; MONTANARI, C. A.; GAUDIO, A. C. Seleção de variáveis em QSAR. **Química Nova**, v. 25, p. 439-448, 2002. DOI: 10.1590/S0100-40422002000300017.

FOROUTAN, A.; GUO, A. C.; VÁZQUEZ-FRESNO, R.; LIPFERT, M.; ZHANG, L.; ZHENG, J.; BADRAN, H.; BUDINSKI, Z.; MANDAL, R.; AMETAJ, B. N.; WISHART, D. S. Chemical composition of Commercial Cow's Milk. **Journal of Agricultural and Food Chemistry**, v. 67, n. 17, p. 4897-4914, 2019. DOI: 10.1021/acs.jafc.9b00204.

FRANÇA, J. A.; FELÍCIO, A. L. S. M.; PEREIRA, R. V.; SILVA JÚNIOR, J. C.; DIAS, L. S.; PINHEIRO, P. P.; SANTOS, J. C. F.; MARTINS, L. R. Um sistema para qualificação de gases baseado em espectrometria no infravermelho próximo. In: **Congresso Brasileiro de Automática-CBA**. 2019.

GALVÃO, R. K. H.; ARAÚJO, M. C. U.; JOSÉ, G. E.; PONTES, M. J. C.; SILVA, E. C.; SALDANHA, T. C. B. A method for calibration and validation subset partitioning. **Talanta**, v. 67, p. 736-740, 2005. DOI: 10.1016/j.talanta.2005.03.025.

GARCIA, M. B. E. de O. **Predição de curva de destilação de gasolina de pirólise a partir de resultados de cromatografia em fase gasosa e calibração multivariada**. Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, 2021.

GOLINELLI, L. P.; CARVALHO, A. C.; CASAES, R. S.; LOPES, C. S. C.; DELIZA, R.; PASCHOALIN, V. M. F.; SILVA, J. T. Sensory analysis and species-specific PCR detect bovine milk adulteration of frescal (fresh) goat cheese. **Journal of Dairy Science**, v. 97, n. 11, p. 6693-6699, 2014. DOI: 10.3168/jds.2014-7990.

GONG, H.; GAO, J.; WANG, Y.; LUO, Q.W.; GUO, K. R.; REN, F. Z.; MAO, X. Y. Identification of novel peptides from goat milk casein that ameliorate high-glucose-induced insulin resistance in HepG2 cells. **Journal of Dairy Science**, v. 103, n. 6, p. 4907-4918, 2020. DOI: 10.3168/jds.2019-17513.

GRASSI, S.; ALAMPRESE, C. Advances in NIR spectroscopy applied to process analytical technology in food industries. **Current Opinion in Food Science**, v. 22, p. 17-21, 2018. DOI: 10.1016/j.cofs.2017.12.008.

GU, F.; LIANG, S.; ZHU, S.; LIU, J.; SUN, H. Multi-omics revealed the effects of rumen-protected methionine on the nutrient profile of milk in dairy cows. **Food Research International**, v. 149, p. 110682, 2021. DOI: 10.1016/j.foodres.2021.110682.

HAN, B. A fosfoproteômica revela que os leites de camelo e cabra melhoram a homeostase da glicose em ratos diabéticos induzidos por HDF/STZ através da ativação do eixo AMPK hepático e GSK3-GYS. **Food Research International**, v. 157, p. 111254, 2022.

HARROU, F.; KADRI, F.; CHAABANE, S.; TAHON, C.; SUN, Y. Improved principal component analysis for anomaly detection: Application to an emergency department. **Computers & Industrial Engineering**, v. 88, p. 63-77, 2015. DOI:

10.1016/j.cie.2015.06.020.

CLARKE, H. J.; MANNION D. T.; O'SULLIVAN, M. G.; KERRY, J. P.; KILCAWLEY, K. N. Development of a headspace solid-phase microextraction gas chromatography mass spectrometry method for the quantification of volatiles associated with lipid oxidation in whole milk powder using response surface methodology. **Food Chemistry**, v. 292, p. 75-80, 2019. DOI: 10.1016/j.foodchem.2019.04.027.

KOSTIĆ, A.Z.; MILINČIĆ, D. D.; STANISAVLJEVIĆ, N. S.; GAŠIĆ, U. M.; LEVIĆ, S.; KOJIĆ, M. O.; TEŠIĆ, J. L.; NEDOVIĆ, N. V.; BARAC, M. B.; PEŠIĆ, M. P. Bioacessibilidade de polifenóis e propriedades antioxidantes de leite de cabra desnatado tratado termicamente digerido in vitro enriquecido com pólen. **Food Chemistry**, v.351, p. 129310. 2021. DOI: 10.1016/j.foodchem.2021.129310.

LI, Q.; ZHAO, Y.; ZHU, D.; PANG, X.; LIU, Y.; FREW, R.; CHEN, G. Lipidomics profiling of goat milk, soymilk and bovine milk by UPLC-Q-Exactive Orbitrap Mass Spectrometry. **Food Chemistry**, v. 224, p. 302-309, 2017. DOI: 10.1016/j.foodchem.2016.12.083.

LIU, H.; GUO, H.; ZHAO, Q.; QIN, Y.;ZHANG, J. Lipidomics analysis for identifying the geographical origin and lactation stage of goat milk. **Food Chemistry**, v. 309, p. 125765, 2020. DOI: 10.1016/j.foodchem.2019.125765.

LIU, Y.; CAI, J.; ZHANG, F. Functional comparison of breast milk, cow milk and goat milk based on changes in the intestinal flora of mice. **Lebensmittel Wissenschaft Technologie**, v. 150, p. 111976, 2021. DOI: 10.1016/j.lwt.2021.111976.

LIU, Z.; LI, C.; PRYCE, J.; ROCHFORT, S. Comprehensive characterization of bovine milk lipids: Triglycerides. **Acs Omega**, v. 5, n. 21, p. 12573-12582, 2020. DOI: 10.1021/acsomega.0c01841.

LYRA, W. S.; SILVA, E. C.; ARAÚJO, M. C. U.; FRAGOSO, W. D.; VÉRAS, G. Classificação periódica: um exemplo didático para ensinar análise de componentes principais. **Química Nova**, 2010. v. 33, n. 7, p. 1594–1597. DOI: 10.1590/S0100-40422010000700030.

LUZ, E. R. **Predição de propriedades de gasolinas usando espectroscopia FTIR e regressão por mínimos quadrados parciais**. Dissertação de Mestrado, PUC-Rio, 2003.

MAPA – MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO. **Instrução Normativa Nº 76, de 26 de novembro de 2018**. Aprovar os Regulamentos Técnicos que fixam a identidade e as características de qualidade que devem apresentar o leite cru refrigerado, o leite pasteurizado, o leite pasteurizado tipo A, na forma desta Instrução Normativa e do Anexo único.

MARIANI, N. C. T.; TEXEIRA, G. H. A.; LIMA, K. M. G.; MORGENSTERN, T. B.; NARDINI, V.; JÚNIOR, L. C. C. NIRS and iSPA-PLS for predicting total anthocyanin content in jaboticaba fruit. **Food Chemistry**, v. 174, p. 643–648, 2015. DOI: 10.1016/j.foodchem.2014.11.008.

MARQUES, I. T. O.; VASCONCELOS, F. R.; ALVES, J. P. M.; MONTENEGRO, A. R.; FERNANDES, C. C. L.; OLIVEIRA, F. B. B.; SILVA, C. P.; NAGANO, C. S.; FIGUEIREDO, S. C.; BESERRA, F. J.; MOURA, A. L.; RONDINA, D. Proteome of milk fat globule membrane and mammary gland tissue in goat fed different lipid supplementation. **Small Ruminant Research**, v. 199, p. 106378, 2021. DOI: 10.1016/j.smallrumres.2021.106378.

MARTINS, M. E. G.; RODRIGUES, J. F. Coeficiente de correlação amostral. **Revista de Ciência Elementar**, v. 2, n. 2, p. 1-2, 2014. DOI: 10.24927/rce2014.042.

MOCHÃO, J. V. C. **Modelação multivariada conjugada com espectroscopia NIR para a otimização de uma formulação contendo elevado teor em Vitamina C**. Dissertação de Mestrado, Universidade de Lisboa, 2015.

MOUFID, M.; TIEBE, C.; EL BARI, N.; FAKRA, D. A. H.; BARTHOLMAI, M.; BOUCHIKHI, B. Pollution parameters evaluation of wastewater collected at different treatment stages from wastewater treatment plant based on E-nose and E-tongue systems combined with chemometric techniques. **Chemometrics and Intelligent Laboratory Systems**, v. 227, p. 104593, 2022. DOI: 10.1016/j.chemolab.2022.104593.

NAMONJE-KAPEMBWA, T.; CHIWAWA, H.; SITKO, N. Analysis of Goat Production and Marketing Among Smallholder Farmers Zambia. **Small Ruminant Research**, p. 106620, 2022. DOI: 10.1016/j.smallrumres.2022.106620.

NAYIK, G. A.; JAGDALE, Y.; GAIKWAD, S. A.; DEVKATTE, A. N.; DAR, A. H.; DEZMIREAN, D. S.; BOBIS, O.; RANJHA, M. M.; ANSARI, M. J.; HEMEG, H. A.; ALOTAIB, S. S. Recent Insights Into Processing Approaches and Potential Health Benefits of Goat Milk and Its Products: A Review. **Frontiers in Nutrition**, v. 8, p. 1033, 2021. DOI: 10.3389/fnut.2021.789117.

NEPA – NÚCLEO DE ESTUDOS E PESQUISAS EM ALIMENTAÇÃO. **Tabela Brasileira de Composição de Alimentos (TACO) 4^a ed.** Campinas: NEPA – UNICAMP, 2011.

NETO, B. B.; SCARMINIO, I. S.; BRUNS, R. E. 20 anos de quimiometria no Brasil. **Química Nova**, São Paulo, v. 29, n. 6, p. 1401-1406, 2006. DOI: 10.1590/S0100-40422006000600042.

OLIVEIRA, A. A.; SIQUEIRA, P. H.; NISGOSKI, S.; BOLZON DE MUNIZ, G. I.; FERREIRA, J. H. Identificação de madeiras utilizando a espectrometria no infravermelho próximo e redes neurais artificiais. **TEMA**, São Carlos, v. 16, p. 81-95, 2015. DOI: 10.5540/tema.2015.016.02.0081.

OLIVEIRA, A.P. D.; ALMEIDA, T. J. O.; SANTOS, T. M. B.; DIAS, F. S. Symbiotic goat milk ice cream with umbu fortified with autochthonous goat cheese lactic acid bacteria. **Lebensmittel Wissenschaft Technologie**, v. 141, p. 110888, 2021. DOI: 10.1016/j.lwt.2021.110888.

OLIVEIRA, E. C. S. **Utilização da RMN Associada a Métodos Quimiométricos na**

Caracterização de Petr6leos e Derivados. Tese de Doutorado, Universidade Federal do Esp6rito Santo, 2016.

PEREIRA, E. V. S.; FERNANDES, D. D. S.; ARAÚJO, M. C. U.; DINIZ, P. H. G. D.; MACIEL, M. I. S. Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms.

Lebensmittel Wissenschaft Technologie, v. 127, p. 109427, 2020. DOI: 10.1016/j.lwt.2020.109427.

PEREIRA, E.V.S. **Novas estrat6gias anal6ticas baseadas em espectroscopia no infravermelho pr6ximo e imagens digitais para identifica76o e quantifica76o de adultera76es em leite caprino.** Tese de Doutorado, Universidade Federal da Para6iba, 2020.

PRIETO, N.; PAWLUCZYK, O.; DUGAN, M. E. R.; AALHUS, L. J. A Review of the principles and applications of near-infrared spectroscopy to characterize meat, fat, and meat products. **Applied Spectroscopy**, v. 71, n. 7, p. 1403-1426, 2017. DOI: 10.1177/0003702817709299.

PIETRZAK-FIE7CKO, R.; KAMELSKA-SADOWSKA, A. M. The comparison of nutritional value of human milk with other mammals' milk. **Nutrients**, v. 12, n. 5, p. 1404, 2020. DOI: 10.3390/nu12051404.

PULINA, G.; MILÁN, M. J.; LAVÍN, D.; THEODORIDIS, A.; MORIN, E.; CAPOTE, J.; THOMAS, D. L.; FRANCESCONI, A. H. D.; CAJA, G. Current production trends, farm structures, and economics of the dairy sheep and goat sectors. **Journal of Dairy Science**, v. 101, n. 8, p. 6715-6729, 2018. DOI: 10.3168/jds.2017-14015.

RINNAN, Å.; VAN DEN BERG, F.; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **TrAC Trends in Analytical Chemistry**, v. 28, n. 10, p. 1201-1222, 2009. DOI: 10.1016/j.trac.2009.07.007.

RODRIGUES, A. N. P.; GIVISIEZ, P.; 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. 2749-2752, 2012. DOI: 10.3168/jds.2011-5235.

ROHMAN, A.; MAN, Y. B. C. Fourier transform infrared (FTIR) spectroscopy for analysis of extra virgin olive oil adulterated with palm oil. **Food Research International**, v. 43, n. 3, p. 886-892, 2010. DOI: 10.1016/j.foodres.2009.12.006.

SALA, O. **Fundamentos da Espectroscopia Raman e no Infravermelho.** S6o Paulo: Editora da UNESP, 1996.

SHAO, X. BIAN, X.; LIU, J.; ZHANG, M.; CAI, W. Multivariate calibration methods in near infrared spectroscopic analysis. **Analytical Methods**, v. 2, n. 11, p. 1662-1666, 2010. DOI: 10.1039/c0ay00421a.

SHARMA, N.; SINGH, S. P.; BHARADWAJ, A.; NATESAN, R. Pregnancy associated

glycoproteins as a potential marker for diagnosis of early pregnancy in goats: A scoping reviewing. **Asian Pacific Journal of Reproduction**, v. 9, n. 6, p. 255, 2020. DOI: 10.4103/2305-0500.298772.

SKOOG, D. A.; HOLLER, F. J.; CROUCH, S. R. **Principles of instrumental analysis**. Cengage learning, Seventh edition, Boston, 2017.

SUH, J. H. Critical review: metabolomics in dairy science-evaluation of milk and milk product quality. **Food Research International**, p. 110984, 2022. DOI: 10.1016/j.foodres.2022.110984.

SOUSA, E. R.; FROTA, C. S.; COSTA, C. H. C.; SILVA, G. S.; SAMPAIO, D. G.; Avaliação da oxidação e de parâmetros de qualidade do óleo de Babacu por espectroscopia no infravermelho médio com transformada de Fourier (FTIR) e calibração multivariada. **Revista Virtual de Química**, v. 11, p. 849-865, 2019.

TARAPOULOUZI, M.; ORTONE, V.; CINTI, S. Heavy metals detection at chemometrics-powered electrochemical (bio) sensors. **Talanta**, p. 123410, 2022. DOI: doi.org/10.1016/j.talanta.2022.123410.

TEYE, E.; ANYIDOHO, E.; AGBEMAFLE, R.; SAM-AMOA, L. K.; ELLIOT, CRIS. Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. **Infrared Physics & Technology**, v. 104, p. 103127, 2020. DOI: 10.1016/j.infrared.2019.103127.

VANZELLA, E.; SANTOS, W. S. O controle de qualidade, por meio das ferramentas BPF e APPCC, em uma linha de produção de uma indústria de alimentos. **Destarte**, v. 5, n. 2, p. 76-90, 2020.

VERRUCK, S.; DANTAS, A.; PRUDENCIO, E. S. Functionality of the components from goat's milk, recent advances for functional dairy products development and its implications on human health. **Journal of Functional Foods**, v. 52, p. 243-257, 2019. DOI: 10.1016/j.jff.2018.11.017.

WANG, H.; WANG, C.; GUO, M. Sucessões autogênicas de bactérias e fungos em grãos de kefir de diferentes origens quando subcultivados em leite de cabra. **Food Research International**, v. 138, p. 109784, 2020. DOI: 10.1016/j.foodres.2020.109784.

WOLD, S.; ESBENSEN, K.; GELADI, P. Principal component analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 2, n. 1-3, p. 37-52, 1987. DOI: 10.1016/0169-7439(87)80084-9.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 58, n. 2, p. 109-130, 2001. DOI: 10.1016/S0169-7439(01)00155-1.

YUAN, T.; WEI, W.; ZHANG, X.; WANG, L.; DAI, X.; REN, C.; WANG, X.; JIN, Q. Medium-and long-chain triacylglycerols composition in preterm and full-term human milk across different lactation stages. **Lebensmittel Wissenschaft Technologie**, v. 142,

p. 110907, 2021. DOI: 10.1016/j.lwt.2021.110907.

ZHANG, H.; XU, Y.; ZHAO, C.; XUE, Y.; TAN, D.; WANG, S.; JIA, M.; WU, H.; MA, A.; CHEN, G. Milk lipids characterization in relation to different heat treatments using lipidomics. **Food Research International**, p. 111345, 2022. DOI: 10.1016/j.foodres.2022.111345.

ZHANG, R.; ZHU, Z.; JIA, W. Molecular mechanism associated with the use of magnetic fermentation in modulating the dietary lipid composition and nutritional quality of goat milk. **Food Chemistry**, v. 366, p. 130554, 2022. DOI: 10.1016/j.foodchem.2021.130554.

ZHAO, L.; WANG, J.; MAO, X. Composition and interfacial properties play key roles in different lipid digestion between goat and cow milk fat globules in vitro. **Food Chemistry**, v. 374, p. 131538, 2022. DOI: 10.1016/j.foodchem.2021.131538.

ZHIQIANG, H. Analysis on the Development Status and Prospects of My Country's Dairy Industry. **China Dairy**, v. 21, p. 20-23, 2021. DOI: 10.12377/1671-4393.21.04.05.

ZHANG, R.; JIA, W.; SHI, L. A Comprehensive Review on the Development of Foodomics-Based Approaches to Evaluate the Quality Degradation of Different Food Products. **Food Reviews International**, p. 1-20, 2022. DOI: 10.1080/87559129.2022.2077362.