



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



Área de Concentração: Ciência de Alimentos

USO DE AMOSTRAS SINTÉTICAS E CALIBRAÇÃO MULTIVARIADA PARA
DETERMINAÇÃO DE CIPROFLOXACINA EM LEITE LIOFILIZADO VIA
ESPECTROSCOPIA VIBRACIONAL

Discente: Izabelle Christinne Pereira Brito

Orientador: Prof. Dr. Sérgio Augusto de Albuquerque Fernandes

Coorientador (a): Dra. Rosângela Silveira Barbosa

ITAPETINGA
BAHIA – BRASIL
MARÇO – 2026

IZABELLE CHRISTINNE PEREIRA DE BRITO

**USO DE AMOSTRAS SINTÉTICAS E CALIBRAÇÃO MULTIVARIADA PARA
DETERMINAÇÃO DE CIPROFLOXACINA EM LEITE LIOFILIZADO VIA
ESPECTROSCOPIA VIBRACIONAL**

Tese apresentada como parte das exigências para obtenção do título de Doutora em Engenharia e Ciência de Alimentos, no Programa de Pós-Graduação em Engenharia e Ciência de Alimentos da Universidade Estadual do Sudoeste da Bahia.

Orientador: Prof. Dr. Sérgio Augusto de Albuquerque Fernandes

Coorientador (a): Dra. Rosângela Silveira Barbosa

ITAPETINGA
BAHIA – BRASIL
MARÇO – 2026

637.127
B875u

Brito, Izabelle Christinne Pereira de.

Uso de amostras sintéticas e calibração multivariada para determinação de ciprofloxacina em leite liofilizado via espectroscopia vibracional. / Izabelle Christinne Pereira de Brito. – Itapetinga-BA: UESB, 2026.

128f.

Tese apresentada como parte das exigências para obtenção do título de Doutora em Engenharia e Ciência de Alimentos, no Programa de Pós-Graduação em Engenharia e Ciência de Alimentos da Universidade Estadual do Sudoeste da Bahia. Área de Concentração em Ciência de Alimentos. Sob a orientação do Prof. D. Sc. Sérgio Augusto de Albuquerque Fernandes e coorientação da D. Sc. Rosângela Silveira Barbosa.

1. Leite – Análise – Resíduo de antibiótico. 2. Leite – Espectroscopia. 3. Leite – Quimiometria. I. Universidade Estadual do Sudoeste da Bahia - Programa de Pós-Graduação de Doutorado em Engenharia e Ciências de Alimentos, *Campus* de Itapetinga. II. Fernandes, Sérgio Augusto de Albuquerque. III. Barbosa, Rosângela Silveira. IV. Título.

CDD (21): 637.127

Catálogo na Fonte:

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

Índice Sistemático para desdobramentos por Assunto:

1. Produtos lácteos – Controle de qualidade
2. Indústria láctea – Segurança alimentar

DECLARAÇÃO DE APROVAÇÃO

Título: “USO DE AMOSTRAS SINTÉTICAS E CALIBRAÇÃO MULTIVARIADA PARA DETERMINAÇÃO DE CIPROFLOXACINA EM LEITE LIOFILIZADO VIA ESPECTROSCOPIA VIBRACIONAL”

Autor (a): IZABELLE CHRISTINNE PEREIRA DE BRITO

Orientador (a): Prof. Dr. Sérgio Augusto de Albuquerque Fernandes

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

Documento assinado digitalmente
 **CAMILA MESSIAS BARBOSA SANTOS CALASANS**
Data: 06/03/2026 14:54:50-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Camila Messias Barbosa Santos
Calasans
(UESB)

Documento assinado digitalmente
 **BARBARA ELIZABETH ALVES DE MAGALHAES**
Data: 04/03/2026 12:15:39-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Bárbara Elizabeth Alves de Magalhães
(IFAL)

Documento assinado digitalmente
 **SERGIO DE SOUSA CASTRO**
Data: 14/03/2026 16:26:30-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Sérgio de Sousa Castro
(UESB)

Documento assinado digitalmente
 **WELLINGTON DA SILVA LYRA**
Data: 12/03/2026 07:13:26-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Wellington da Silva Lyra
(UFPB)

Documento assinado digitalmente
 **SERGIO AUGUSTO DE ALBUQUERQUE FERNANDES**
Data: 18/03/2026 11:53:46-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Sérgio Augusto de Albuquerque Fernandes
Orientador – UESB
Presidente da Banca

Itapetinga-BA, 04 de março de 2026.

“O insucesso é apenas uma oportunidade para recomeçar de novo com mais inteligência.”

Henry Ford

AGRADECIMENTOS

Agradeço primeiramente a Deus por sempre me dar forças e amparo em todos os momentos de dificuldades, sem Ele não conseguiria chegar até aqui.

A todos os meus familiares em especial aos meus pais Cláudia e Jovelino pelo suporte, amor e carinho durante essa jornada.

A Matheus por todo apoio, companheirismo e amor durante esses anos.

Ao meu orientador, professor Dr. Sérgio Augusto de Albuquerque Fernandes, pela orientação, amizade e companheirismo.

A minha coorientadora Dra. Rosângela Silveira Barbosa e toda a equipe da EMBRAPA – Clima Temperado, pela parceria e suporte.

A Universidade Estadual do Sudoeste da Bahia (UESB), por ter me possibilitado desenvolver este trabalho.

A Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) pelo fornecimento da bolsa incentivando cada vez mais a pesquisa e a ciência.

SUMÁRIO

LISTA DE TABELAS	x
LISTA DE FIGURAS	xi
LISTA DE EQUAÇÕES	xiii
LISTA DE SIGLAS E ABREVIATURAS	xiv
RESUMO	xv
ABSTRACT	xvi
CAPÍTULO I: REFERENCIAL TEÓRICO	17
1. Introdução	17
2. Revisão de literatura	18
2.1. Produção de leite no mundo e no Brasil	18
2.2. Leite: aspectos nutricionais e qualidade	20
2.3. Antibióticos de uso veterinário	21
2.3.1. Fluorquinolonas.....	22
2.3.1.1. Ciprofloxacina.....	23
2.4. Resistência Antimicrobiana	24
2.5. Legislação sobre medicamentos de uso veterinário em leite	25
2.6. Métodos de detecção e quantificação de resíduos de antibióticos em leite	26
2.6.1. Espectroscopia no Infravermelho Médio por Transformada de Fourier	28
2.6.2. Espectroscopia no Infravermelho Próximo - NIR.....	30
2.7. Quimiometria	32
2.7.1. Análise de Componentes Principais - PCA.....	33
2.7.2. Regressão por Componentes Principais - PCR	34
2.7.3. Regressão por Mínimos Quadrados Parciais - PLS	35
2.7.4. Análise Discriminante por Mínimos Quadrados Parciais – PLS - DA	36
2.8. Amostras Sintéticas.....	37
2.8.1. Geração de amostras sintéticas via PCA	39
3. Referências	40
CAPÍTULO II: OBJETIVOS	59
Objetivo Geral.....	59
Objetivos Específicos	59
CAPÍTULO III: ARTIGO I – USO DE AMOSTRAS SINTÉTICAS NA PREDIÇÃO QUIMIOMÉTRICA DE RESÍDUOS DE CIPROFLOXACINA EM LEITE EM PÓ OBTIDO POR LIOFILIZAÇÃO UTILIZANDO ESPECTROSCOPIAS VIBRACIONAIS	60
1. Introdução	60

2. Material e Métodos	63
2.1. Amostras	63
2.2 Delineamento experimental	63
2.3. Liofilização	64
2.4. Espectroscopia no infravermelho médio com Transformada de Fourier – FTIR- MIR	64
2.5. Espectroscopia no Infravermelho Próximo – NIR.....	64
2.6. Pré-tratamento dos dados.....	64
2.7. Geração das amostras sintéticas.....	64
2.8. Análises quimiométricas	66
2.8.1. Regressão de mínimos quadrados parciais - PLS.....	67
2.8.2 Regressão por Componentes Principais-PCR	69
3. Resultados e discussão	70
3.1. Geração das amostras sintéticas.....	70
3.2. Características espectrais obtidas por MIR.....	72
3.2.1. Leite liofilizado puro	72
3.2.2. Padrão de Ciprofloxacina	74
3.2.3. Amostras originais fortificadas e sintéticas.....	76
3.3. Características espectrais obtidas por NIR	77
3.3.1. Leite liofilizado puro	77
3.3.2. Padrão de Ciprofloxacina	79
3.3.3. Amostras originais fortificadas e sintéticas.....	80
3.4. Análise multivariadas	82
3.4.1. Regressão de Mínimos Quadrados Parciais – PLS	82
3.4.1.1. Avaliação da robustez do modelo PLS por validação Monte Carlo	90
3.4.2. Regressão por Componentes Principais-PCR	92
4. Conclusão	94
5. Referencias	94
CAPÍTULO IV: ARTIGO II - USO DE ESPECTROSCOPIAS VIBRACIONAIS E QUIMIOMETRIA PARA TRIAGEM RÁPIDA DE CIPROFLOXACINA EM LEITE EM PÓ OBTIDO POR LIOFILIZAÇÃO.....	101
1. Introdução.....	101
2. Material e Métodos.....	103
2.1. Amostras	103
2.2. Delineamento experimental	103
2.3. Liofilização	104
2.4. Espectroscopia no infravermelho médio (FTIR-MIR)	104
2.5. Espectroscopia no infravermelho próximo (NIR).....	104

2.6. Pré-processamento dos dados espectrais	104
2.7. Amostras sintéticas	104
2.8. Análises quimiométricas	105
2.8.1. Análise de Componentes Principais – PCA	105
2.8.2. Análise de Mínimos Quadrados Parciais - PLS-DA	106
3. Resultados e Discussão.....	106
3.1. Investigação espectral e procedimentos de pré-processamento.....	106
3.2. Análises quimiométricas.....	109
3.2.1. Análise de Componentes Principais.....	109
3.2.2. Análise de Mínimos Quadrados Parciais - PLS-DA	112
3.2.2.1. Avaliação da robustez do modelo PLS-DA por validação Monte Carlo	121
4. Conclusão	122
5. Referencias	122

LISTA DE TABELAS

CAPÍTULO I

Tabela 1: Limites Máximos de Resíduos de antimicrobianos em leite fluido e em pó.....	25
Tabela 2: Região do número de onda dos grupos funcionais no MIR.	30
Tabela 3: Região do comprimento de onda dos grupos funcionais no NIR.....	31

CAPÍTULO III

Tabela 1. Parâmetros de calibração dos modelos de regressão por mínimos quadrados parciais (PLS) obtidos a partir de espectros no infravermelho médio (MIR) e no infravermelho próximo (NIR) para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado de vacas da raça Jersey.....	83
Tabela 2. Métricas de desempenho dos dados de validação com e sem outliers.....	87
Tabela 3. Desempenho do modelo PLS-MIR-SNV com dados combinados avaliado por simulação de Monte Carlo (média $\pm$ DP).	90
Tabela 4. Parâmetros de calibração por regressão em componentes principais (PCR) obtidos a partir de espectros MIR e NIR para a predição da concentração de Ciprofloxacina em leite liofilizado de vacas Jersey.	93

CAPÍTULO IV

Tabela 1. Parâmetros de desempenho para os modelos PLS-DA usando espectroscopias MIR e NIR.	113
Tabela 2. Matriz de confusão para amostras de calibração e validação com base no modelo PLS-DA, considerando dados MIR.	119
Tabela 3. Desempenho do modelo PLS-DA-MIR-DSG com dados combinados avaliado por simulação de Monte Carlo (média $\pm$ DP).	121

LISTA DE FIGURAS

CAPÍTULO I

Figura 1. Produção de leite em 2024.	19
Figura 2. Produção de leite em pó 2022.	19
Figura 3. Estrutura básica das fluorquinolonas.	22
Figura 4. Estrutura química da Ciprofloxacina.	23

CAPÍTULO III

Figura 1. Código para geração das amostras sintéticas por espectros MIR e NIR.....	66
Figura 2. Geração de amostras sintéticas de leite de vacas da raça Jersey a partir de espectros no infravermelho médio (MIR, A) e no infravermelho próximo (NIR, B), projetadas no espaço de componentes principais (PCA).	71
Figura 3. Espectro médio no infravermelho médio (MIR) de amostras de leite liofilizado puro de vacas da raça Jersey.	73
Figura 4. Espectro no infravermelho médio por FTIR (MIR) do padrão de Ciprofloxacina.	75
Figura 5. Espectros médios no infravermelho médio (MIR) de amostras de leite liofilizado puro e fortificadas com diferentes concentrações do padrão de Ciprofloxacina, considerando dados originais (A) e sintéticos (B).	76
Figura 6. Espectro médio no infravermelho próximo (NIR) de amostras de leite liofilizado puro de vacas da raça Jersey.	78
Figura 7. Espectro no infravermelho próximo (NIR) do padrão de Ciprofloxacina.	79
Figura 8. Espectro médio no infravermelho próximo (NIR) de amostras de leite liofilizado de vacas da raça Jersey fortificadas com diferentes concentrações do padrão de Ciprofloxacina, considerando dados originais (A) e sintéticos (B).	81
Figura 9. Gráfico de correlação entre os valores reais e previstos pelo modelo PLS-MIR utilizando dados combinados e pré-processamento por SNV, para conjuntos de calibração (A) e validação (B).	86
Figura 10. <i>Loadings</i> das variáveis latentes (LV1–LV5) do modelo de regressão por mínimos quadrados parciais (PLS–MIR), desenvolvido a partir de espectros pré-processados por SNV, para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado de vacas da raça Jersey.....	88
Figura 11. Coeficientes de regressão do modelo de regressão por mínimos quadrados parciais (PLS–MIR), com pré-processamento por SNV, para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado de vacas da raça Jersey.	89
Figura 12. Variação das métricas de desempenho obtidas nas 50 iterações do teste de Monte Carlo, para o conjunto de calibração (A e B) e para o conjunto de validação (C e D).	91

CAPÍTULO IV

Figura 1. Espectro MIR processados por SNV (A), MSC (B), correção de linha de base (C) e DSG (D).	107
Figura 2. Espectro NIR processados por SNV (A), MSC (B), correção de linha de base (C) e DSG (D).	108

Figura 3. PCA dos espectros MIR (A, C) gráficos de <i>scores</i> das amostras originais e dos dados combinados, respectivamente e (B, D) <i>loadings</i> correspondentes às mesmas análises.	110
Figura 4. PCA dos espectros NIR (E, G) gráficos de <i>scores</i> das amostras originais e dos dados combinados, respectivamente e (F, H) <i>loadings</i> correspondentes às mesmas análises.....	111
Figura 5. Gráficos de dispersão dos valores Y preditos pelo modelo PLS-DA para as diferentes classes; A e C correspondem ao conjunto de calibração, enquanto B e D referem-se ao conjunto de validação externa.	118
Figura 6. Gráficos de <i>scores</i> (A) e <i>loadings</i> (B) do modelo PLS-DA.	120

LISTA DE EQUAÇÕES**CAPÍTULO I**

Equação 1	39
Equação 2	39

CAPÍTULO III

Equação 1	67
Equação 2	68
Equação 3	68
Equação 4	68
Equação 5	69
Equação 6	69
Equação 7	69

LISTA DE SIGLAS E ABREVIATURAS

CLAE – Cromatografia Líquida de Alta Eficiência

FAO - *Food and Agriculture Organization* - Organização das Nações Unidas para a Alimentação e Agricultura

FTIR – *Fourier Transform Infrared Spectroscopy* - Espectroscopia no Infravermelho Médio por Transformada de Fourier

LMR – Limite Máximo de Resíduo

MAPA - Ministério da Agricultura, Pecuária e Abastecimento

MIR – Espectroscopia no Infravermelho Médio

MSC – *Multiplicative Scatter Correction* - Correção de Espalhamento Multiplicativo

NIR – Espectroscopia no Infravermelho Próximo

OMS - Organização Mundial da Saúde

PAMVET - Programa de Análise de Resíduos de Medicamentos Veterinários em Alimentos de Origem Animal

PCA – *Principal Component Analysis* - Análise de Componentes Principais

PCR – *Principal Component Regression* - Regressão por Componentes Principais

PLS - *Partial Least Squares* – Regressão de Mínimos Quadrados Parciais

PLS-DA – *Partial Least Squares Discriminant Analysis* - Análise Discriminante por Mínimos Quadrados Parciais

r – Coeficiente de Correlação

RER - *Range Error Ratio* - Razão de Erro de Amplitude

RMSEC - *Root Mean Square Error of Calibration* - Erro Quadrático Médio de Calibração

RMSEP - *Root Mean Square Error of Prediction* - Erro Quadrático Médio de Predição

RPD - *Ratio of Performance to Deviation* - Desvio Preditivo Residual

SNV – *Standard Normal Variate* - Variável Normal Padrão

VLs – Variáveis Latentes

RESUMO

O uso inadequado de antibióticos em rebanhos leiteiros pode resultar na presença de resíduos no leite em concentrações superiores ao limite máximo de resíduo (LMR) estabelecido pela legislação, representando riscos à saúde dos consumidores e contribuindo para o desenvolvimento de resistência microbiana. Nesse contexto, cresce a demanda por métodos analíticos rápidos, não destrutivos e confiáveis, capazes de ampliar a frequência das análises sem comprometer a qualidade dos resultados. Entre elas, destacam-se a Espectroscopia no Infravermelho Médio com Transformada de Fourier (FTIR – *Fourier Transform Infrared Spectroscopy*) e a Espectroscopia no Infravermelho Próximo (NIR – *Near Infrared Spectroscopy*), consideradas promissoras por sua rapidez, baixo custo e caráter não destrutivo. O objetivo deste trabalho foi desenvolver uma metodologia para detecção e quantificação de resíduos de Ciprofloxacina em leite em pó obtido por liofilização, empregando técnicas espectroscópicas vibracionais associadas à quimiometria e à geração de amostras sintéticas. Foram coletadas amostras de leite de vacas da raça Jersey livres de antibióticos, às quais foram adicionadas concentrações conhecidas do padrão de Ciprofloxacina (25, 50, 75, 100, 125, 150, 175 e 200 $\mu\text{g L}^{-1}$). Em seguida, as amostras foram liofilizadas e analisadas por espectroscopia no infravermelho médio (MIR) e próximo (NIR). Com o intuito de ampliar o conjunto amostral e aumentar a robustez estatística dos modelos, amostras sintéticas foram geradas a partir da Análise de Componentes Principais (PCA – *Principal Component Analysis*) dos espectros MIR e NIR. Foram considerados dois conjuntos de dados: Originais e Combinados (dados originais combinados com sintéticos), submetidos a diferentes pré-processamentos espectrais, incluindo Variável Normal Padrão (SNV – *Standard Normal Variate*), Correção de Espalhamento Multiplicativo (MSC – *Multiplicative Scatter Correction*), Primeira Derivada de Savitzky–Golay e Correção de Linha de Base. Os conjuntos de dados foram avaliados por meio de análises quimiométricas de predição (PCR – *Principal Component Regression* / Regressão por Componentes Principais e PLS – *Partial Least Squares Regression* / Regressão por Mínimos Quadrados Parciais), de classificação (PLS-DA – *Partial Least Squares Discriminant Analysis* / Análise Discriminante por Mínimos Quadrados Parciais) e análise exploratória (PCA). Para as análises exploratória (PCA) e de classificação (PLS-DA), os dados (originais e combinados) foram organizados em três classes: classe 0 (sem adição de Ciprofloxacina), classe 1 (baixa concentração: 25–100 $\mu\text{g L}^{-1}$) e classe 2 (alta concentração: 125–200 $\mu\text{g L}^{-1}$). As técnicas espectroscópicas mostraram-se sensíveis para detectar as modificações químicas associadas à presença de Ciprofloxacina, podendo consolidar-se como métodos rápidos e confiáveis para investigação de resíduos deste fármaco em leite em pó liofilizado. O modelo PLS baseado em espectros MIR, utilizando dados combinados e pré-processamento SNV, apresentou elevada robustez e precisão, com coeficiente de correlação (r) > 0,90, Razão entre o Desempenho e o Desvio-padrão (RPD - *Ratio of Performance to Deviation*) > 4, Razão entre a Faixa e o Erro (RER - *Range Error Ratio*) > 14 e Erro Quadrático Médio da Raiz (RMSE - *Root Mean Square Error*) < 15 nas etapas de calibração e validação. A análise PCR não apresentou desempenho satisfatório em nenhum modelo testado, enquanto a PCA não demonstrou separação clara entre as classes de concentração do analito. Por outro lado, a análise PLS-DA, associada à espectroscopia MIR, ao pré-processamento de primeira derivada de Savitzky–Golay e à inclusão de amostras sintéticas, alcançou exatidão média de 0,994 no conjunto de calibração e 0,991 no conjunto de validação, com alta especificidade, sensibilidade e precisão em todas as classes. Conclui-se que a integração de técnicas espectroscópicas MIR e NIR com ferramentas quimiométricas avançadas e geração de amostras sintéticas constitui uma abordagem inovadora, sustentável e eficiente para a detecção e quantificação de resíduos de Ciprofloxacina em leite em pó liofilizado. Essa metodologia demonstra elevado potencial para aplicação em rotinas de controle de qualidade e segurança alimentar na indústria láctea, promovendo análises mais rápidas, econômicas e ambientalmente seguras.

Palavras-chave: MIR; NIR; Segurança Alimentar; Amostras Sintéticas; Quimiometria.

ABSTRACT

The inappropriate use of antibiotics in dairy herds may result in the presence of residues in milk at concentrations above the Maximum Residue Limit (MRL) established by legislation, posing risks to consumer health and contributing to the development of microbial resistance. In this context, the demand for rapid, non-destructive, and reliable analytical methods has increased, enabling a higher frequency of analyses without compromising result quality. Among these methods, Fourier Transform Mid-Infrared Spectroscopy (FTIR) and Near-Infrared Spectroscopy (NIR) stand out as promising techniques due to their speed, low cost, and non-destructive nature. The objective of this study was to develop a methodology for the detection and quantification of Ciprofloxacin residues in freeze-dried milk powder using vibrational spectroscopic techniques combined with chemometrics and synthetic sample generation. Milk samples were collected from Jersey cows free of antibiotics, to which known concentrations of a Ciprofloxacin standard were added (25, 50, 75, 100, 125, 150, 175, and 200 $\mu\text{g L}^{-1}$). The samples were then freeze-dried and analyzed by mid-infrared (MIR) and near-infrared (NIR) spectroscopy. To expand the dataset and increase the statistical robustness of the models, synthetic samples were generated from Principal Component Analysis (PCA) of the MIR and NIR spectra. Two datasets were considered: Original and Combined (original data combined with synthetic samples), which were subjected to different spectral preprocessing techniques, including Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), Savitzky–Golay First Derivative, and Baseline Correction. The datasets were evaluated using chemometric analyses for prediction (PCR – Principal Component Regression and PLS – Partial Least Squares Regression), classification (PLS-DA – Partial Least Squares Discriminant Analysis), and exploratory analysis (PCA). For the exploratory (PCA) and classification (PLS-DA) analyses, the datasets (original and combined) were organized into three classes: Class 0 (no Ciprofloxacin added), Class 1 (low concentration: 25–100 $\mu\text{g L}^{-1}$), and Class 2 (high concentration: 125–200 $\mu\text{g L}^{-1}$). The spectroscopic techniques proved to be sensitive in detecting chemical modifications associated with the presence of Ciprofloxacin and may become rapid and reliable methods for investigating residues of this drug in freeze-dried milk powder. The PLS model based on MIR spectra, using combined data and SNV preprocessing, showed high robustness and accuracy, with a correlation coefficient (r) > 0.90 , Ratio of Performance to Deviation (RPD) > 4 , Range Error Ratio (RER) > 14 , and Root Mean Square Error (RMSE) < 15 in both calibration and validation stages. The PCR analysis did not present satisfactory performance in any of the tested models, whereas PCA did not demonstrate clear separation between the analyte concentration classes. In contrast, PLS-DA analysis, associated with MIR spectroscopy, Savitzky–Golay first derivative preprocessing, and the inclusion of synthetic samples, achieved a mean accuracy of 0.994 in the calibration set and 0.991 in the validation set, with high specificity, sensitivity, and precision across all classes. It can be concluded that the integration of MIR and NIR spectroscopic techniques with advanced chemometric tools and synthetic sample generation represents an innovative, sustainable, and efficient approach for the detection and quantification of Ciprofloxacin residues in freeze-dried milk powder. This methodology demonstrates high potential for application in quality control and food safety routines in the dairy industry, promoting analyses that are faster, more economical, and environmentally safer.

Keywords: MIR; NIR; Food Safety; Synthetic Samples; Chemometrics.

CAPÍTULO I: REFERENCIAL TEÓRICO

1. Introdução

O leite é amplamente reconhecido como um dos alimentos mais completos e essenciais para a nutrição humana, por conter proteínas de alto valor biológico, lipídios, carboidratos, vitaminas e minerais. Seu elevado valor nutricional faz com que ele e seus derivados desempenhem papel fundamental na alimentação e na indústria de alimentos, sendo consumidos tanto em sua forma fluida quanto em produtos industrializados, como o leite em pó.

Entre os métodos de obtenção desse produto, destaca-se o *spray drying*, técnica amplamente utilizada pela indústria de laticínios por permitir a secagem rápida e eficiente do leite líquido. No entanto, esse processo envolve temperaturas elevadas que podem causar degradação de compostos termossensíveis, como vitaminas e proteínas, além de alterar características estruturais e funcionais de alguns componentes. Nesse contexto, o processo de liofilização surge como uma alternativa viável e vantajosa, pois preserva melhor as propriedades nutricionais, físico-químicas e bioquímicas do leite, sendo especialmente indicado para estudos laboratoriais que exigem alta precisão analítica e estabilidade das amostras.

Na produção leiteira moderna, é comum o uso de medicamentos veterinários, especialmente antibióticos, para fins terapêuticos e profiláticos. No entanto, o uso inadequado desses compostos pode resultar na presença de resíduos no leite, configurando risco à saúde pública e comprometendo a qualidade do produto. A Ciprofloxacina, pertencente à classe das fluoroquinolonas, é um dos antibióticos frequentemente utilizados em rebanhos leiteiros, e sua presença residual acima dos limites estabelecidos por órgãos reguladores representa uma preocupação crescente devido ao potencial de promover reações adversas e o desenvolvimento de resistência microbiana.

Diante desse cenário, a detecção e a quantificação de resíduos de antibióticos em leite demandam métodos analíticos altamente sensíveis, seletivos e reprodutíveis. Embora técnicas cromatográficas sejam tradicionalmente utilizadas como referência, elas apresentam limitações associadas ao alto custo, tempo de análise e necessidade de preparo complexo de amostras. Assim, métodos espectroscópicos, como os baseados nas regiões do infravermelho médio (MIR) e próximo (NIR), surgem como alternativas promissoras por serem técnicas rápidas, não destrutivas e ambientalmente sustentáveis, capazes de fornecer informações físico-químicas e estruturais de forma simultânea.

Apesar dos avanços, a aplicação dessas técnicas em matrizes complexas, como o leite,

ainda enfrenta desafios relacionados à variabilidade amostral e à construção de modelos quimiométricos robustos. A limitação no número de amostras e na representatividade dos conjuntos de dados pode comprometer o desempenho dos modelos preditivos. Nesse contexto, a geração de amostras sintéticas por meio da Análise de Componentes Principais (PCA - *Principal Component Analysis*) apresenta-se como uma alternativa eficiente para expandir a base de dados, preservando a estrutura multivariada original e aprimorando a estabilidade e a generalização dos modelos desenvolvidos.

Dessa forma, este trabalho tem como objetivo aplicar técnicas espectroscópicas (MIR e NIR) associadas a métodos quimiométricos e à geração de amostras sintéticas via PCA, a fim de desenvolver uma abordagem inovadora, sustentável e eficaz para a detecção, predição e classificação de resíduos de Ciprofloxacina em leite em pó liofilizado, contribuindo para o aprimoramento do controle de qualidade e da segurança alimentar na cadeia produtiva leiteira.

2. Revisão de literatura

2.1. Produção de leite no mundo e no Brasil

A produção mundial de leite, é liderada pelo leite de vaca, que representa 81% da produção total, seguida de 15% de leite de búfala. Os leites de cabra, ovelha e camelo totalizam 4% da produção global (LEITE, STOCK e RUBACK, 2022). A produção de leite de vaca é difundida em todas as regiões do mundo, e em 2024, a produção mundial foi de 673,29 milhões de toneladas. A Índia neste ano, destacou-se como o maior produtor leiteiro, sendo responsável por 31% da produção mundial, seguida da União Europeia e dos Estados Unidos da América (USDA, 2025). A figura 1, mostra os dados dos principais produtores de leite em 2024.

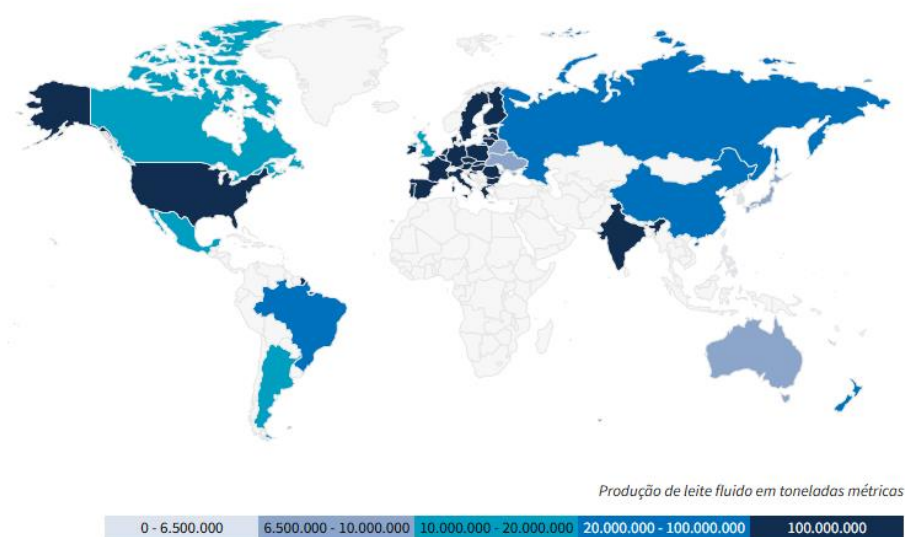


Figura 1. Produção de leite em 2024.

Fonte: USDA, 2025

O Brasil, no ano de 2024, ficou em 6º lugar no ranking dos maiores produtores de leite, produzindo 25 bilhões de litros (IBGE, 2025). Considerando que o Brasil tem mais de 1 milhão de produtores distribuídos em praticamente todos os municípios nacionais, deste modo estima-se que a cadeia produtiva do leite gere 4 milhões de empregos nos seus diferentes segmentos, resultando em valor bruto da produção superior a R\$ 49,9 bilhões (BRASIL, 2021).

A relação da produção mundial de leite em pó é apresentada na Figura 2. Dados referentes a produção de leite em pó em 2022, destaca a Nova Zelândia, Argentina, França, Alemanha e China como os principais produtores, produzindo mais de 62.000 toneladas anuais. Estes números revelam que estes países são ótimos exportadores do produto e possuem uma cadeia produtiva eficiente. Países da África, América Central e partes da Ásia registraram níveis baixos de produção indicando limitações na industrialização do leite. A produção intermediária (6.870 a 18.449 toneladas) foi registrada em países como Estados Unidos, Brasil, Austrália e Rússia, com capacidade de produção significativa, mas não em nível global. O Brasil apesar de ser um grande produtor de leite fluido, a industrialização da matéria-prima em leite em pó é mais limitada (FAO, 2022).

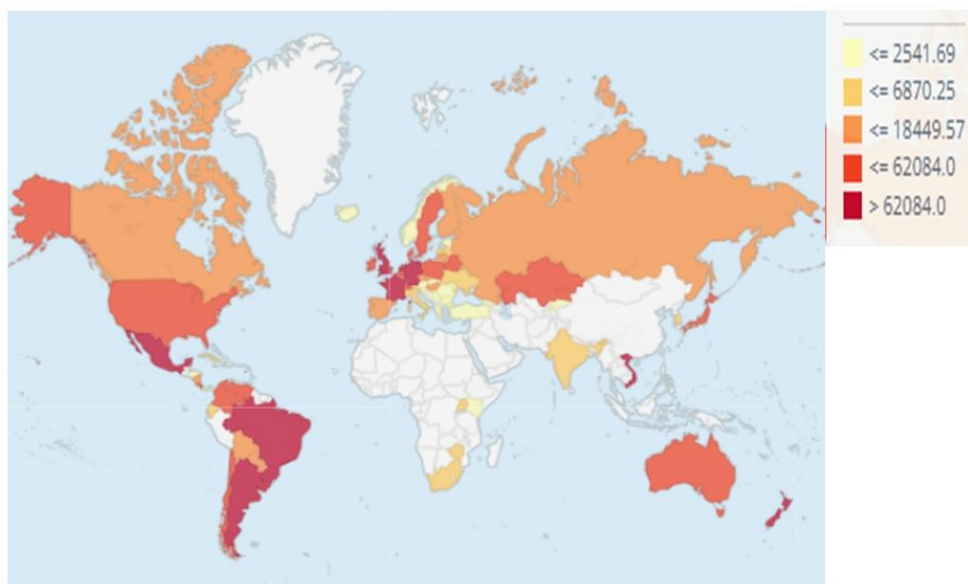


Figura 2. Produção de leite em pó 2022.

Fonte: FAO, 2022.

A produção leiteira exerce importante papel na economia nacional, assim, o Ministério do Desenvolvimento Agrário e Agricultura Familiar, tem como prioridade proteger e apoiar o

setor, afim de ampliar a competitividade e a qualidade da cadeia leiteira nacional (EMBRAPA, 2024). Além de exercer papel fundamental na economia, o leite também se destaca pela sua importância nutricional, sendo consumido mundialmente, por ser fonte de nutrientes essenciais para a alimentação humana.

2.2. Leite: aspectos nutricionais e qualidade

O leite de vaca e seus derivados se destacam por sua importância nutricional, pois fornecem aos consumidores, proteínas, minerais e vitaminas de alta qualidade (HASSEN et al., 2022). A composição nutricional desta matéria-prima dependerá de fatores intrínsecos e extrínsecos ao gado leiteiro, como: espécie; alimentação; ambiente de criação; saudabilidade e bem-estar do animal (BEDNARSKI e KUPCZYŃSKI, 2024).

Visando apenas à lucratividade e à produtividade, alguns produtores adulteram o leite por meio da adição de água, glicose e isolados de proteínas de soja e ervilha, utilizados para aumentar o volume do produto e substituir sua composição por componentes nutricionais inferiores (DU et al, 2018) além da adição deliberada de formaldeído para aumentar a vida útil do produto (COIMBRA et al., 2020).

Além das adulterações o leite fluido e em pó também podem sofrer contaminações físicas, químicas e biológicas, devido ao não cumprimento das Boas Práticas de Fabricação e Manejo. A contaminação física pode ocorrer pela presença de pelos, areia, pedra ou qualquer outro material estranho. A matéria-prima pode ser deteriorada por organismos microbiológicos, através da falta de higiene dos manipuladores, ao não higienizarem corretamente as mãos e a sala de ordenha, podendo ser fonte de transmissão de micro-organismos deteriorantes e patogênicos. Já a contaminação química pode acontecer, através da utilização de produtos de limpeza e da aplicação inadequada de medicamentos veterinários (RAHEEM e NIAMAH, 2021; CALAHORRANO-MORENO, 2022).

A contaminação por medicamentos veterinários está diretamente relacionada à ocorrência de doenças infecciosas que acometem o rebanho leiteiro, como as mastites clínicas e subclínicas. A mastite é uma das doenças mais frequente no rebanho leiteiro, e implica negativamente na produtividade e na economia (TOMMASONI et al., 2023). É causada por micro-organismos, que infectam o rebanho, causando a inflamação da glândula mamária e a elevação das células somáticas (ZIGO et al, 2021). A mastite clínica é caracterizada por sintomas visíveis, como aparecimento de grumos no leite e inflamação das tetas (SCHMENGER e KRÖMKER, 2020), já a subclínica é a forma mais prevalente da doença, que

conduz, na ausência de sinais clínicos. Assim, faz-se necessário a tomada de decisões corretas e a aplicação de medidas adequadas, que sejam capazes de minimizar os fatores externos e internos que aumentam a risco deste tipo de infecção (TOMMASONI et al., 2023).

Segundo Zigo et al. (2021) para prevenir a ocorrência de mastites no rebanho leiteiro é importante a implementação de boas práticas de higiene durante a ordenha, seguindo as seguintes etapas: definir uma meta de saúde da glândula mamária; garantir alojamento limpo e seco para os animais; adesão à ordem das vacas leiteiras na sala de ordenha (vacas paridas, grupos de produção, fim da lactação e tratadas); procedimento de ordenha escolhido corretamente; cuidado com o equipamento de ordenha. Porém, quando o rebanho é diagnosticado com essas enfermidades, o tratamento é realizado com a aplicação de antimicrobianos de uso veterinário (ZIGO et al., 2021). No entanto, a aplicação desses medicamentos, na maioria das vezes, é realizada pelos próprios produtores de maneira inadequada, o que favorece a contaminação do leite pela presença de resíduos de antimicrobianos.

2.3. Antibióticos de uso veterinário

A descoberta dos antibióticos em 1928, é considerado um marco histórico, sendo denominado como “drogas milagrosas”. A penicilina foi o primeiro antibiótico descoberto, através de estudos realizados por Alexander Fleming, e a partir deste antibiótico, várias pesquisas foram desenvolvidas em prol da obtenção de medicamentos anti-infecciosos, para o tratamento e controle de diversas doenças infecciosas (CANESCHI et al., 2023).

Os antibióticos são substâncias encontradas em micro-organismos que passam por processo de sintetização, para serem utilizados no tratamento de infecções em humanos e animais (QIAO et al., 2018). Os antibacterianos visam inibir vias ou estruturas essenciais para o desenvolvimento do agente infeccioso, ou seja, bactérias Gram-positivas e Gram-negativas (KIM et al., 2023). A escolha do antibacteriano, irá depender do seu alvo. A maioria dos antibióticos agem sobre as vias essenciais das bactérias, incluindo a replicação do DNA, a transcrição do RNA e a tradução das proteínas; e sobre as estruturas da parede celular e da membrana bacteriana (SILVER, 2016).

Estas substâncias podem ser classificadas com base na estrutura química em: β -lactâmicos, lincosamida, vancomicina, teicoplanina, ansamicina, macrolídeo, tetraciclina, aminoglicosídeo, fluoroquinolona e sulfonamida (HADDAD et al., 2024; LIMA et al. 2020; KIM et al., 2023).

Na criação animal, os antibióticos são utilizados não apenas para o tratamento de infecções, mas também como aditivos alimentares, promovendo a engorda dos animais, e consequentemente promove o aumento da produtividade (LOW et al., 2021). Neste sentido, as classes de antibióticos mais usuais na pecuária, são: sulfonamidas, tetraciclina, aminoglicosídeos, cefalosporinas (β -lactâmicos) e fluorquinolonas (BARROS e DI STASI, 2012; THAI et al., 2023).

2.3.1. Fluorquinolonas

Fluorquinolonas, é um grupo de antibióticos que visam inibir a ação de bactérias Gram-positivas e Gram-negativas, no tratamento de infecções graves e invasivas em humanos e animais (THAI et al., 2023). Este grupo de antibióticos é considerado um dos mais desenvolvidos e representam umas das classes mais efetivas de agentes antimicrobianos considerando os seguintes parâmetros: amplo espectro de atividade, administrações e distribuição nos tecidos infectados e o combate efetivo à bactérias (LUNGU et al., 2022).

As moléculas das fluorquinolonas são versáteis e permitem melhorias nas propriedades farmacocinéticas e farmacodinâmicas. A composição básica da classe das fluorquinolonas apresentada na Figura 3, baseia-se no compartilhamento de dois anéis com um grupo carboxila e um grupo cetona, a presença do flúor adicionado a estrutura e um átomo de nitrogênio (LUNGU et al., 2022).

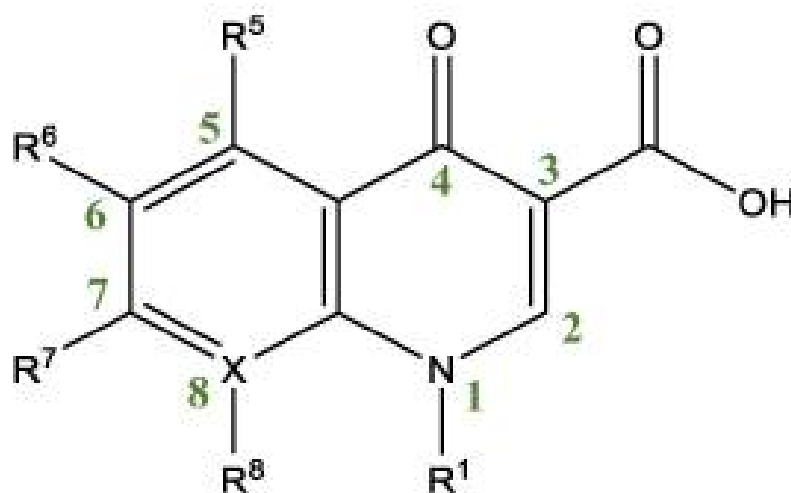


Figura 3. Estrutura básica das fluorquinolonas.

Fonte: Kherroubi, Bacon e Rahman, 2024.

A classe das fluorquinolonas tem como objetivo inibir a enzima DNA girase de bactérias Gram-positivas e Gram-negativas. Ao serem aplicados para o tratamento de infecções os antibióticos desta classe formam um complexo fluorquinolonas-enzima-DNA, assim a síntese

do DNA é inibida e o DNA é clivado em ambas as fitas. A morte do agente infeccioso irá ocorrer dependendo da concentração do antibiótico (RUSU et al., 2023).

A classe das fluoroquinolonas inclui diversos antibióticos, porém os de segunda geração se destacam pelo amplo espectro de ação e elevada eficácia. Entre eles, destacam-se: Norfloxacino, Pefloxacino, Ofloxacina e Ciprofloxacina (LUNGU et al., 2022). Em animais, a Ciprofloxacina é um dos antibióticos mais usuais para o tratamento de infecções causadas por *E. coli*, *Salmonella*, *Pasteurella* e *Mycoplasma* (XIONG et al., 2023).

2.3.1.1. Ciprofloxacina

A Ciprofloxacina é um antibiótico de segunda geração das fluorquinolonas, é caracterizada por uma estrutura química típica da classe (Figura 4), contendo um núcleo quinolônico, formado por um átomo de flúor, responsável por melhorar a eficiência do antibiótico, um grupo carboxílico e um grupo cetona, que juntos são essenciais para inibir a replicação do DNA girase dos agentes infecciosos e possui também um anel piperazínico que amplia sua atividade contra bactéria Gram-negativas (KHWAZA, MLALA e ADERIBIGBE, 2024).

A Ciprofloxacina apresenta excelente atividade bactericida, principalmente contra micro-organismos aeróbicos Gram-negativos; o combate eficaz contra *Pseudomonas aeruginosa* é destaque em estudos (BAGGIO e ANANDA-RAJAH, 2021).

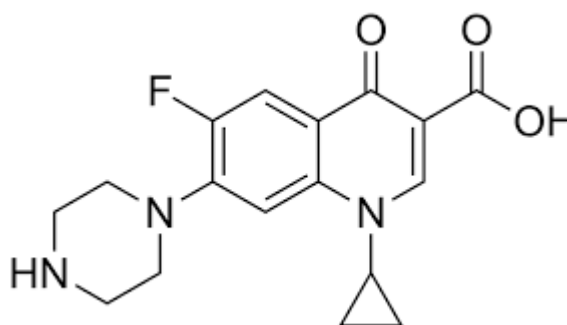


Figura 4. Estrutura química da Ciprofloxacina.

Fonte: Kherroubi, Bacon e Rahman, 2024.

A Ciprofloxacina é um medicamento veterinário comumente usados na aquicultura e pecuária. Estudos relatam o uso do medicamento contra mastite clínica em bovinos, salmonelose em frangos de corte e para aumentar a população de peixes (PELYUNTHA et al., 2022; MARINS et al., 2024; BUSARI et al., 2024)

Em rebanhos leiteiros a Ciprofloxacina é frequentemente usada no Brasil, por seu curto período de carência (2 dias) em outros países o período de carência pode chegar até 8 dias, e por essa finalidade seu uso é recomendado e aprovado em vários países (BULA – CIPROLAC OUROFINO SAÚDE ANIMAL LTDA, 2016; MAHMOOD et al., 2016). Mas o uso inadequado deste medicamento, pode conferir resistência antimicrobiana à classe dos antibióticos (VASCO et al., 2024).

2.4. Resistência Antimicrobiana

Resistência bacteriana, refere-se à perda da eficiência do antibiótico de inibir ou matar o micro-organismo infeccioso, em que o mesmo, adquire capacidade de suportar o efeito bactericida ou bacteriostático do antimicrobiano. Desta forma, o tratamento das infecções se torna mais difícil e aumenta o risco de ocorrência de outras doenças e a até a morte de animais e de humanos (CHINEMEREM et al., 2022).

Na criação animal os antibióticos são usados também como aditivos alimentares, que promovem o crescimento. Estima-se que 80% dos antibióticos produzidos são usados na produção alimentícia, e o uso indiscriminado destas drogas pode levar ao surgimento bactérias resistentes, e efeitos negativos a saúde do consumidor (LOW et al., 2021). A resistência antimicrobiana aos antibióticos é cada vez mais preocupante e abordada. Nesse sentido, o conceito de Saúde Única e as diretrizes de administração de antibióticos impõem que as indústrias produtoras de alimentos supervisionem o uso de medicamentos antimicrobianos em fazendas, a fim de diminuir o risco do surgimento de bactérias resistentes aos antibióticos (LAMMIE e HUGHES, 2016).

Várias classes de antibióticos estão associadas à disseminação da resistência antimicrobiana, o que gera implicações para a saúde pública e veterinária. Entre elas, destacam-se os macrolídeos e as fluorquinolonas, classificados como prioridade na lista da Organização Mundial da Saúde (OMS) de Antimicrobianos Criticamente Importantes para a Medicina Humana (RYAN, MCNEAL e CREDILLE, 2023).

O uso impróprio de antibióticos, como a superdosagem, o descumprimento do período de carência e o emprego para fins de engorda, contribui para a presença de resíduos desses medicamentos em produtos de origem animal, favorecendo o surgimento da resistência antimicrobiana, o que representa uma preocupação crescente para a saúde pública. Espera-se que essa resistência cause cerca de 10 milhões de mortes humanas por ano até 2050 e tenha um impacto econômico de quase US\$ 100 trilhões globalmente. No intuito de minimizar estes

impactos negativos, organismos nacionais e internacionais, são responsáveis por fiscalizar e implementar normas que visam diminuir a presença destes resíduos em alimentos (BASSETTI et al., 2017; MUTEEB et al., 2023).

2.5. Legislação sobre medicamentos de uso veterinário em leite

Para enfrentar a problemática, presença de resíduos de antibióticos em leite, organizações internacionais, como a Organização das Nações Unidas para a Alimentação e Agricultura (FAO) e a OMS, elaboraram políticas específicas. No Brasil, o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), juntamente com o Ministério da Saúde (MS), colabora com entidades internacionais na formulação dessas políticas. Entre as diversas ações dessas organizações, destacam-se a definição dos limites máximos de resíduos (LMR) para medicamentos veterinários permitidos (BRITO et al., 2024).

Segundo o Programa de Análise de Resíduos de Medicamentos Veterinários em Alimentos de Origem Animal – PAMVET, o LMR é a concentração máxima de resíduos provenientes do uso de um medicamento veterinário considerada aceitável. Esse limite é estabelecido com base na quantidade e no tipo de resíduos que não representam risco para a saúde humana, levando em conta a Ingestão Diária Aceitável (IDA). Além disso, os LMR consideram as boas práticas veterinárias e avaliam a presença de resíduos em alimentos de origem vegetal e/ou no ambiente (BRASIL, 2007).

A Tabela 1, apresenta os limites máximos de resíduos para os principais antibióticos veterinários permitidos.

Tabela 1: Limites Máximos de Resíduos de antimicrobianos em leite fluido e em pó.

Classe	Antibiótico	LMR ($\mu\text{g L}^{-1}$)	Referência
Fluorquinolonas	Ciprofloxacina	100	União Europeia
β-lactâmicos	Cefalosporina:		
	Ceftiofur	100	Mercosul
	Cefoperazone	50	Mercosul
Aminoglicosídeos	Neomicina	500	Mercosul
Tetraciclinas	Oxitetraciclina	100	Mercosul
Sulfonamidas	Sulfametazina	100	Mercosul

Fonte: Adaptado de BRASIL, 2007 e UNIÃO EUROPEIA, 2010.

A detecção e quantificação de resíduos em produtos de origem animal requerem métodos analíticos específicos, sendo os testes de triagem e outras abordagens analíticas amplamente empregados para essa finalidade (BRITO et al., 2024).

2.6. Métodos de detecção e quantificação de resíduos de antibióticos em leite

Para a obtenção de leite com qualidade e segurança, testes devem ser usados para detectar a presença de resíduos de antimicrobianos no produto. Espera-se que estes métodos sejam capazes de identificar uma ampla gama de antibióticos, com rapidez e baixo custo, mas a maioria, não atende estes requisitos (VERCELLI et al., 2023).

Métodos qualitativos, são largamente utilizados em fazendas produtoras de leite, para a investigação da presença de resíduos de antimicrobianos. São métodos simples, rápidos e de baixo custo, capazes de identificar a presença ou ausência dos resíduos. Testes como Delvotest® (BION et al., 2016; JAGESAR e KERKHOF, 2023), Charm MRL® (SHI et al., 2021; HASONOVÁ et al., 2025) e Snap Beta-Lactam® (DUREL, GALLINA E PELLET, 2019) baseiam-se em técnicas imunológicas e microbiológicas de identificação, mas são susceptíveis a resultados falso-positivos, fazendo-se necessária a confirmação de amostras positivas por métodos quantitativos (PACHECO-SILVA, SOUZA e CALDAS, 2014; GRABSK et al., 2019).

Diversas metodologias analíticas têm sido empregadas para a detecção e quantificação de resíduos de antimicrobianos em leite e produtos lácteos, variando quanto à seletividade, sensibilidade e aplicabilidade prática. Entre os métodos convencionais, destacam-se os ensaios microbiológicos, que se baseiam na inibição do crescimento bacteriano em presença de antimicrobianos. Esses testes apresentam a vantagem de detectar diferentes classes de antibióticos em uma única análise, sendo úteis como triagem inicial (VERCELLI et al., 2023). Entretanto, possuem baixa especificidade, limites de detecção elevados e demandam longos tempos de incubação, o que limita seu uso para análises quantitativas precisas (CHEN et al., 2018; WU et al., 2019).

Outra abordagem amplamente utilizada é o imunoensaio, como o teste ELISA (Ensaio imunoenzimático - *Enzyme-Linked Immunosorbent Assay*), que permite a detecção rápida e sensível de compostos específicos, sendo adequado para monitoramento de rotina (AHMED et al., 2020; GETAHUN et al., 2023). Contudo, esse método depende de anticorpos altamente específicos e pode sofrer reações cruzadas, além de apresentar limitações na quantificação exata

e na identificação de múltiplos resíduos simultaneamente (SOTNIKOV et al., 2021; ZHOU et al., 2018; CHEN et al., 2018).

Nos últimos anos, têm ganhado destaque os biossensores e outros métodos instrumentais alternativos, que oferecem rapidez, portabilidade e baixo custo operacional (SINGH et al., 2023). Os biossensores eletroquímicos e ópticos, por exemplo, demonstram sensibilidade adequada para triagens preliminares, mas ainda enfrentam desafios quanto à estabilidade, reprodutibilidade e validação para matrizes complexas como o leite (CHEN et al., 2018; ALIEV et al., 2022).

A Cromatografia Líquida de Alta Eficiência (CLAE) destaca-se entre os métodos quantitativos utilizados na análise de resíduos de antimicrobianos em leite devido à sua elevada sensibilidade, especificidade analítica e capacidade de quantificar compostos em níveis muito baixos. Essa técnica cromatográfica é amplamente empregada em estudos de determinação de antibióticos em alimentos de origem animal, pois permite separar eficientemente analitos em matrizes complexas e fornece dados quantitativos confiáveis quando adequada à matriz e ao analito de interesse. Além disso, variantes da CLAE acopladas a detectores sensíveis, como espectrometria de massa, ampliam ainda mais sua capacidade de detecção e identificação de múltiplos resíduos simultaneamente (PARMAR, CHAUBEY, GUPTA E BHARATH, 2021; VERCELLI et al., 2023). Apesar das inúmeras vantagens em utilizar a CLAE, este método possui algumas limitações que devem ser consideradas. A técnica cromatográfica requer instrumentação de alto custo, as análises são demoradas (necessita de preparo de amostras), faz uso de reagentes químicos, além da necessidade de colaboradores devidamente treinados (VERCELLI et al., 2023).

Neste contexto, técnicas espectroscópicas vibracionais surgem como alternativas promissoras para o controle rápido e não destrutivo de resíduos de antibióticos em leite líquido, em pó ou liofilizado. A Espectroscopia no Infravermelho Médio por Transformada de Fourier (FTIR) e a Espectroscopia no Infravermelho Próximo (NIR), especialmente quando associadas a ferramentas quimiométricas, possibilitam a avaliação simultânea de múltiplos constituintes da matriz e a detecção indireta de contaminantes por meio de alterações no perfil espectral. Essas técnicas oferecem vantagens como mínimo preparo de amostras, rapidez e redução no uso de reagentes químicos, tornando-as adequadas para aplicações de triagem e monitoramento (KURREY et al., 2019; FREITAS et al., 2021; CONCEIÇÃO et al., 2024).

Entretanto, apesar do avanço tecnológico, as espectroscopias vibracionais ainda apresentam limitações de sensibilidade quando comparadas às técnicas cromatográficas e dependem fortemente da modelagem quimiométrica para interpretação precisa dos espectros. Ainda assim, representam um campo em expansão, com potencial para se consolidar como ferramenta complementar ou substitutiva em análises de rotina, especialmente em laboratórios de controle de qualidade e em sistemas de monitoramento rápido de resíduos (FREITAS et al., 2021; SANTANA et al., 2025; SANTOS et al., 2026).

2.6.1. Espectroscopia no Infravermelho Médio por Transformada de Fourier

A Espectroscopia no Infravermelho Médio por Transformada de Fourier, se destaca como técnica que não necessita de preparo de amostras, tem alto rendimento, baixo custo e não requer produtos químicos perigosos, além de atuar como técnica qualitativa e quantitativa. O MIR é um método que permite a identificação de substâncias por meio de vibrações moleculares. A espectroscopia baseia-se na interação entre a radiação eletromagnética e a matéria, deste modo, os espectros de infravermelho fornecem a descrição de energia que a molécula absorve mediante excitação (HASHIMOTO et al., 2024; SAJI et al., 2024).

Os espectros de infravermelho médio (MIR) são obtidos a partir de interferogramas gerados por um interferômetro de Michelson, iluminado por uma fonte térmica de amplo espectro, geralmente um filamento de globo de Nernst ou SiC (HASHIMOTO et al., 2024). O interferômetro divide e recombina o feixe de radiação, criando um padrão de interferência dependente das diferenças ópticas de percurso. O sinal resultante, denominado interferograma, contém informações de todas as frequências absorvidas pela amostra. A Transformada de Fourier (FT), aplicada por meio de processamento matemático, converte esse sinal do domínio do tempo para o domínio da frequência, produzindo o espectro de absorção característico do material analisado (HOU et al., 2018).

No espectro resultante, cada banda de absorção corresponde à vibração de um grupo funcional específico, associada às transições entre níveis energéticos vibracionais de ligações químicas. As posições e intensidades dessas bandas são determinadas pela diferença de energia entre os estados vibracionais, sendo diretamente influenciadas pela força da ligação e pela massa dos átomos envolvidos (GUERRERO-PÉREZ e PATIENCE, 2020; FADLELMOULA et al., 2022). Assim, o espectro MIR fornece uma impressão digital molecular da amostra, permitindo a identificação qualitativa e, mediante calibração quimiométrica, a quantificação de compostos orgânicos e inorgânicos (FREITAS et al., 2021; SANTANA et al., 2025).

Nos sistemas modernos, a aquisição dos espectros MIR é frequentemente realizada utilizando o acessório ATR (*Attenuated Total Reflectance*), que permite a análise direta de líquidos, pastas ou sólidos sem preparo prévio de amostra. O princípio do ATR baseia-se na reflexão interna total da radiação infravermelha em um cristal de alto índice de refração, como diamante, germânio ou seleneto de zinco (ZnSe) (MOHAMED et al., 2017; TKACHENKO e NIEDZIELSKI, 2022). Quando o feixe de luz incide no cristal e sofre reflexão interna total, uma pequena fração da radiação penetra na amostra, formando uma onda evanescente que interage com a superfície do material. A absorção dessa radiação pela amostra é então registrada pelo detector, resultando em um espectro representativo de sua composição superficial (HAAS E MIZAIKOFF, 2016; MOHAMED et al., 2017; MENDES e DUARTE, 2021).

A principal vantagem do FTIR-ATR é a simplicidade e rapidez da análise, eliminando etapas como moagem, secagem ou prensagem em pastilhas de KBr, exigidas em sistemas de transmissão convencionais. Além disso, a técnica permite obter espectros reprodutíveis de amostras heterogêneas, úmidas ou viscosas, características comuns ao leite e a produtos lácteos (MENDES e DUARTE, 2021). Essa característica torna o FTIR-ATR amplamente aplicável em estudos de alimentos e, mais recentemente, na detecção de resíduos de antimicrobianos (FREITAS et al., 2021; SANTANA et al., 2025; SANTOS et al., 2026).

Pesquisas recentes demonstram que o uso do MIR, aliado a estratégias quimiométricas, permite detectar pequenas alterações no perfil espectral do leite decorrentes da presença de antibióticos, mesmo em baixas concentrações (SAJI et al., 2024; FREITAS et al., 2021; SANTANA et al., 2025; SANTOS et al., 2026). Essas alterações geralmente ocorrem em regiões espectrais associadas às vibrações de grupos carbonila (C=O), amida (N-H e C-N) e hidroxila (O-H), os quais sofrem influência da interação dos antimicrobianos com proteínas e lipídios da matriz (GUERRERO-PÉREZ e PATIENCE, 2020; SANTOS et al., 2026).

Além disso, a espectroscopia no infravermelho médio apresenta excelente repetibilidade, baixo consumo energético e requer mínima manutenção instrumental, o que o torna adequado para monitoramento rápido em laboratórios de rotina ou ambientes industriais (COQUEIRO et al., 2024; FREITAS et al., 2025). Quando acoplado a métodos de aprendizado de máquina e validação multivariada, tem se mostrado uma alternativa sustentável e confiável para detecção e quantificação de resíduos e avaliação da autenticidade do leite e derivados (SAJI et al., 2024; SANTANA et al., 2025).

Os espectros obtidos na espectroscopia no infravermelho médio (MIR) são baseados nas vibrações fundamentais que estão relacionadas aos grupos funcionais presentes nas amostras e podem ser observados e registrados entre 4000 e 400 cm^{-1} (SAJI et al., 2024).

A espectroscopia no infravermelho médio é utilizada para detecção de contaminantes, fraudes e até mesmo toxinas em produtos farmacêuticos, alimentos e na medicina humana (SU e LEE, 2020; SONG et al., 2020; AHMAD e AYUB, 2022) e identificação pode ser associada aos grupos funcionais presentes nas amostras (Tabela 2).

Tabela 2. Região do número de onda dos grupos funcionais no MIR.

Função Orgânica	Grupo Funcional	Região do número de onda (cm^{-1})
Éter	C–O	1150-1070
Éster	C–O	1310-1210
	C=O	1750-1730
Cetona	C=O	1750-1660
Aldeído	C–H	2850-2700
	C=O	1740-1685
Álcool	C–O	1200-1100
	O–H	3600-3200
Amina	C–N	1360-1180
Alceno	C=C	1680-1620
Alcano	C–H	2950-2850
	C–C	1200

Fonte: Adaptado de HAGE, 2012.

É importante destacar que essa tabela é genérica. Dependendo da matriz e de outras condições, esses picos podem ser deslocados, sobrepostos e/ou mudar de intensidade.

2.6.2. Espectroscopia no Infravermelho Próximo - NIR

A espectroscopia no infravermelho próximo (NIR) destaca-se como uma técnica analítica de baixo custo, não destrutiva e de fácil manuseio, sendo amplamente utilizada em investigações químicas e industriais. Essa metodologia permite a identificação e caracterização

de substâncias com base nas absorções correspondentes a sobretons e combinações das bandas fundamentais de vibração molecular, principalmente associadas aos grupos funcionais C–H, N–H e O–H (BELLON-MAUREL et al., 2011). A Tabela 3, apresenta os comprimentos de onda referentes aos grupos funcionais caracterizados por NIR.

Tabela 3. Região do comprimento de onda dos grupos funcionais no NIR.

Função Orgânica	Comprimento de onda (nm)	Grupo Funcional
Alcano	1210	C–H
Alcano	1355–1400	C–H
Amida	1435	N–H
Álcool / Fenol	1480	O–H
Amida	1500	N–H
Álcool / Fenol	1540	O–H
Amida	1565	N–H
Alcano	1610–1725	C–H
Alcano	1780	C–H
Álcool / Éter	1820	O–H + C–O
Éter	1955	C–O
Amida	2050	N–H + amida secundária
Álcool / Fenol	2075	O–H
Amida	2155	Amida primária + amida terciária
Álcool / Fenol	2255	O–H
Amida	2295	N–H + CO

Fonte: Adaptado de BEĆ, GRABSKA e HUCK, 2020

Deve-se enfatizar que a tabela tem caráter ilustrativo, pois a atribuição de bandas no NIR é notoriamente desafiadora, em função da alta sobreposição de bandas harmônicas e combinações vibracionais típicas dessa região espectral.

No contexto das análises por espectroscopia no infravermelho próximo, destacam-se os espectrômetros NIR, amplamente utilizados para análises rápidas e não destrutivas de diferentes tipos de amostras. Esses equipamentos operam, geralmente, na faixa espectral de aproximadamente 1100 a 2500 nm e são aplicáveis à análise de amostras sólidas, pastosas ou líquidas. Entre as tecnologias empregadas nesses instrumentos, destaca-se o sistema de

varredura por monocromador rotativo, responsável pela seleção de comprimentos de onda específicos da fonte de luz por meio de uma grade de difração de alta precisão (COQUEIRO et al., 2024; FREITAS et al., 2025; CANDEIAS et al., 2025).

A Espectroscopia no Infravermelho Próximo tem sido amplamente empregada em diferentes áreas, como saúde, agricultura e alimentos, devido à sua versatilidade, rapidez e caráter não destrutivo. Em alimentos, destaca-se como uma ferramenta poderosa para identificação de fraudes, monitoramento da qualidade e detecção de contaminantes, incluindo resíduos de antibióticos (ERLINGE et al., 2021; DU et al., 2022). Apesar de a região espectral do NIR conter informações químicas relevantes, a alta sobreposição espectral e a natureza difusa das bandas tornam a interpretação direta dos espectros desafiadora, exigindo o uso de ferramentas quimiométricas avançadas para extração e modelagem dos dados (LUIZ et al., 2018).

2.7. Quimiometria

A quimiometria é uma área interdisciplinar que combina princípios da estatística, matemática e ciência da computação para extrair informações relevantes de dados químicos complexos (ROHMAN et al., 2020; DUPONT et al., 2020). O termo foi introduzido na década de 1970 e, desde então, tem se consolidado como uma ferramenta essencial na análise química moderna, especialmente em técnicas espectroscópicas e cromatográficas (OLIVERI, MALEGORI e CASALE, 2020; COQUEIRO et al., 2024; FREITAS et al., 2025).

A crescente complexidade dos sistemas analíticos e o grande volume de dados gerados por instrumentos modernos tornaram inviável a interpretação puramente visual ou univariada (DUPONT et al., 2020). Nesse contexto, a quimiometria permite modelar relações entre variáveis dependentes e independentes, identificar padrões ocultos e construir modelos preditivos robustos (OLIVERI, MALEGORI e CASALE, 2020; HOUHOU e BOCKLITZ, 2021). Assim, sua aplicação abrange desde o controle de qualidade industrial e autenticação de produtos até a elucidação de processos físico-químicos e o monitoramento ambiental.

O emprego de ferramentas quimiométricas tem sido amplamente difundido em diversas técnicas analíticas, como a Cromatografia Líquida de Alta Eficiência (CLAE) (PAGES-REBULL et al., 2023; ZHANG et al., 2021), a Espectroscopia no Infravermelho Próximo (ZHU et al., 2019; TEYE et al., 2019), a Espectroscopia no Infravermelho Médio (FATTAHI et al., 2024; NAVID e AKBARI-HASANJANI, 2025) e a Espectroscopia Raman (NAVID e AKBARI-HASANJANI, 2025; SANTOS et al., 2023), entre outras. Nessas abordagens, a

quimiometria contribui para a extração e interpretação de grandes matrizes de dados, permitindo a distinção entre amostras, a quantificação simultânea de compostos e a detecção de adulterações.

Na área de alimentos, a quimiometria tem se mostrado uma aliada estratégica no controle de qualidade, autenticação e rastreabilidade de produtos, especialmente em matrizes complexas como o leite e seus derivados (NTAKATSANE et al., 2020; ZADE et al., 2023). A combinação de técnicas espectroscópicas (como NIR e MIR) com métodos quimiométricos possibilita análises rápidas, não destrutivas e com mínima preparação de amostras, viabilizando a detecção de fraudes, a avaliação de parâmetros nutricionais e a identificação de resíduos de antibióticos (NTAKATSANE et al., 2020; ZADE et al., 2023; SANTANA et al., 2025; SANTOS et al., 2026). Além disso, modelos multivariados permitem o monitoramento contínuo da produção leiteira, assegurando a conformidade com padrões regulatórios e contribuindo para a segurança alimentar (FREITAS et al., 2021).

Entre os principais métodos quimiométricos empregados, destacam-se a Análise de Componentes Principais (PCA), utilizada para explorar e reduzir a dimensionalidade dos dados, identificando agrupamentos naturais e amostras atípicas (BIANCOLILLO e MARINI, 2018); a Regressão por Componentes Principais (PCR), que combina a decomposição dos dados com a regressão linear (LEE et al., 2020); a Regressão por Mínimos Quadrados Parciais (PLS), voltada para calibração quantitativa (GARCIA et al., 2015); e sua variação discriminante, a Análise Discriminante por Mínimos Quadrados Parciais, amplamente aplicada em classificação de amostras e reconhecimento de padrões (FOLLI et al., 2024). Esses métodos constituem a base da análise multivariada moderna, sendo fundamentais para a interpretação e predição em sistemas químicos complexos.

2.7.1. Análise de Componentes Principais - PCA

A Análise de Componentes Principais (*Principal Component Analysis* – PCA) é uma das ferramentas mais empregadas em quimiometria para a exploração de dados multivariados, sendo essencial na etapa de investigação inicial e no pré-processamento de modelos supervisionados (JOLLIFFE e CADIMA, 2016; BIANCOLILLO e MARINI, 2018). Desenvolvida originalmente por Karl Pearson em 1901 e posteriormente formalizada por Harold Hotelling em 1933, a PCA tem como objetivo reduzir a dimensionalidade de um conjunto de dados, preservando a maior variância possível (PÖGE e REINECKE, 2021).

Em termos conceituais, a PCA transforma um conjunto original de variáveis correlacionadas em um novo conjunto de variáveis ortogonais denominadas componentes principais (PCs). Cada componente é uma combinação linear das variáveis originais e é calculada de forma a capturar o máximo de variância remanescente nos dados, sob a restrição de ortogonalidade em relação às componentes anteriores (ANOWAR, SADAOUUI e SELIM, 2021). Assim, a primeira componente (PC1) representa a maior parte da variabilidade total, a segunda (PC2) descreve a maior variância residual e assim sucessivamente, até que a variância cumulativa desejada seja alcançada.

A aplicação da PCA permite identificar tendências, agrupamentos e *outliers*, sendo amplamente utilizada na análise de dados espectroscópicos, como Espectroscopia no Infravermelho Próximo, Espectroscopia no Infravermelho Médio e Espectroscopia Raman, para distinguir amostras com base em diferenças químicas sutis (COQUEIRO et al., 2024; FREITAS et al., 2025; NAVID e AKBARI-HASANJANI, 2025). Além disso, os *scores* e *loadings* obtidos na decomposição auxiliam na interpretação físico-química dos dados: os *scores* representam as amostras em um novo espaço reduzido de variáveis (componentes), enquanto os *loadings* indicam a contribuição de cada variável original na formação dessas componentes (ANOWAR, SADAOUUI e SELIM, 2021; PÖGE e REINECKE, 2021).

2.7.2. Regressão por Componentes Principais - PCR

A Regressão por Componentes Principais (*Principal Component Regression* - PCR) é uma técnica multivariada amplamente utilizada na quimiometria para modelagem e predição em sistemas complexos, sendo particularmente eficaz quando há alta correlação entre variáveis explicativas (multicolinearidade) ou quando o número de variáveis supera o número de amostras (BEGDACHE et al., 2019; AGARWAL et al., 2019). O método combina a Análise de Componentes Principais (PCA) com a regressão linear múltipla, substituindo as variáveis originais por um conjunto reduzido de componentes principais ortogonais, que retêm a maior parte da variância dos dados e eliminam redundâncias (MAYERHÖFER et al., 2025).

Na PCR, as componentes principais são inicialmente obtidas a partir da decomposição da matriz de preditores, e apenas as primeiras componentes (que explicam a maior proporção da variabilidade de X) são utilizadas para ajustar o modelo de regressão com a variável resposta. Essa abordagem reduz o impacto do ruído instrumental e da colinearidade entre as variáveis, tornando o modelo mais estável e interpretável (BEGDACHE et al., 2019; AGARWAL et al., 2019; MAYERHÖFER et al., 2025). Contudo, diferentemente da Regressão por Mínimos Quadrados Parciais (PLS), a PCR não considera diretamente a correlação entre os preditores e

a variável dependente durante a extração das componentes, o que pode resultar na inclusão de informações que pouco contribuem para a capacidade preditiva do modelo (LIU et al., 2022).

Apesar dessa limitação, a PCR continua sendo uma das técnicas mais aplicadas em análises espectroscópicas e controle de qualidade de alimentos, permitindo a predição simultânea de múltiplos constituintes químicos e a avaliação de parâmetros físico-químicos de forma rápida e não destrutiva (MOHD SALLEH e HASSAN, 2019; FLORES-VALDEZ et al., 2020).

2.7.3. Regressão por Mínimos Quadrados Parciais - PLS

A Regressão por Mínimos Quadrados Parciais (*Partial Least Squares Regression* - PLS) é uma das técnicas mais amplamente utilizadas em quimiometria para modelagem de dados espectrais e construção de modelos preditivos multivariados (LIMA et al., 2020; GRELET et al., 2024). Desenvolvida originalmente por Herman Wold na década de 1970 (WESTLAND et al., 2019). Sua principal vantagem reside na capacidade de lidar com dados altamente colineares e com número de variáveis (p) superior ao número de amostras (n), condição típica de espectros vibracionais obtidos por técnicas como Espectroscopia no Infravermelho Próximo e Espectroscopia no Infravermelho Médio (JESUS et al., 2020; PEREIRA et al., 2021; FREITAS et al., 2025).

Diferentemente da regressão linear múltipla tradicional, a PLS não busca ajustar diretamente os coeficientes das variáveis originais, mas extrai variáveis latentes (VLs) ou componentes principais que maximizam simultaneamente a covariância entre os espectros (matriz X) e a variável resposta (matriz Y). Assim, o método projeta X e Y em um novo espaço de menor dimensionalidade, preservando a variabilidade relevante para a predição (LIMA et al., 2020; COQUEIRO et al., 2024; FREITAS et al., 2025).

Do ponto de vista prático, a PLS apresenta elevada robustez e capacidade preditiva, mesmo em conjuntos de dados com ruídos experimentais, sobreposição de bandas espectrais e variabilidade instrumental. Por essa razão, tem sido amplamente aplicada em estudos envolvendo determinação de constituintes químicos, controle de qualidade e monitoramento de processos industriais (LIMA et al., 2020; COQUEIRO et al., 2024). Em sistemas complexos como o leite e seus derivados, a PLS permite quantificar simultaneamente múltiplos analitos e detectar pequenas variações composicionais associadas a fatores tecnológicos ou adulterações (PEREIRA et al., 2021; GRELET et al., 2024).

Além disso, a eficiência da modelagem PLS depende criticamente da seleção adequada do número de variáveis latentes. A avaliação do desempenho do modelo é realizada com base em métricas como o erro quadrático médio de calibração (RMSEC), validação (RMSECV), previsão (RMSEP), o coeficiente de correlação (r) e o Desvio Residual da Predição (*Residual Prediction Deviation* -RPD), indicadores essenciais para aferir a confiabilidade e aplicabilidade do modelo (JESUS et al., 2020; LIMA et al., 2020; FREITAS et al., 2025). De forma complementar, abordagens modernas, como PLS Discriminante (PLS-DA), permitem a classificação de amostras em diferentes grupos, tornando o método versátil tanto para fins quantitativos quanto qualitativos (FOLLI et al., 2024).

2.7.4. Análise Discriminante por Mínimos Quadrados Parciais – PLS - DA

A Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) é uma extensão da Regressão PLS voltada para problemas de classificação multivariada, sendo amplamente utilizada em quimiometria para identificar padrões e distinguir classes de amostras com base em dados espectrais complexos (FOLLI et al., 2024). O método foi introduzido a partir da adaptação do algoritmo PLS, substituindo a variável resposta contínua (Y) por uma matriz categórica binária, que codifica as diferentes classes sob investigação (RUIZ-PEREZ et al., 2020; LASALVIA, CAPOZZI E PERNA, 2022). Essa modificação permite que a PLS-DA modele a relação entre os espectros (X) e as categorias (Y) de forma supervisionada, extraindo variáveis latentes que maximizam a separação entre grupos e minimizam a variabilidade intraclasse.

A principal vantagem da PLS-DA é sua capacidade de lidar com dados colineares e de alta dimensionalidade, típicos de espectros NIR, MIR ou Raman, mantendo uma estrutura interpretável baseada em *loadings* e *scores* (SUN et al., 2020; ASRI et al., 2022). Essa abordagem tem se mostrado eficaz na discriminação de amostras adulteradas, identificação de origem geográfica, autenticação de produtos e detecção de contaminantes, especialmente em matrizes alimentares complexas como leite, mel e azeite (AKBARI, BAIGBABAEI e SHAHIDI, 2020; PEREIRA et al., 2021; EL MAOUARDI et al., 2024).

A avaliação do desempenho de modelos PLS-DA é geralmente realizada por meio de validação cruzada e conjuntos independentes de teste, utilizando métricas como acurácia, sensibilidade, especificidade, precisão e análise de confusão (FOLLI et al., 2024; COQUEIRO et al., 2024). Embora seja uma técnica poderosa e interpretável, a PLS-DA pode apresentar risco de *overfitting* quando aplicada a conjuntos de dados com poucas amostras ou alto desequilíbrio entre classes, exigindo uma validação criteriosa e estratégias complementares,

como o uso de Monte Carlo ou permutações aleatórias para verificação da significância estatística (FOLLI et al., 2024; TODOROV, SEARLE-WHITE e GERBER, 2020). Ainda assim, a PLS-DA permanece como uma das ferramentas mais robustas e acessíveis para classificação multivariada em estudos espectroscópicos e de quimiometria aplicada.

É importante destacar que modelos multivariados eficazes dependem de uma quantidade representativa de dados para que possam aprender padrões reais e generalizáveis. Quando o número de amostras disponíveis é pequeno, a diversidade dos dados é comprometida, favorecendo o surgimento de modelos sobreajustados (*overfitting*), que apresentam bom desempenho apenas nas amostras usadas para treinamento, mas falham ao serem aplicados a novos conjuntos (SOLTANA, SABETZADEH e BRIAND, 2017; LI et al., 2017). Isso ocorre porque o modelo não consegue capturar toda a variabilidade inerente à matriz analisada, prejudicando sua capacidade de previsão e sua utilidade prática (BABYAK, 2004). Uma abordagem viável para contornar essa limitação é a utilização de dados sintéticos, que complementam o conjunto original e ajudam a ampliar o espaço amostral de forma controlada.

2.8. Amostras Sintéticas

A geração de amostras sintéticas tem se consolidado como uma estratégia promissora para ampliar a robustez e a representatividade de modelos quimiométricos, especialmente quando há limitação no número de amostras reais disponíveis (FOLLI et al., 2024). Esses dados artificiais são produzidos de forma controlada, a partir da estrutura estatística do conjunto original, mantendo as relações multivariadas entre as variáveis. Essa abordagem permite reduzir o risco de sobreajuste e melhorar a capacidade preditiva dos modelos multivariados, como PCR, PLS e PLS-DA, principalmente em contextos de alta dimensionalidade e baixo número amostral (SOLTANA, SABETZADEH e BRIAND, 2017; LI et al., 2017; FOLLI et al., 2024).

As amostras sintéticas podem ser geradas a partir de diferentes tipos de conjuntos de dados, como espectros obtidos por Espectroscopias NIR, MIR, Raman, dados cromatográficos, imagens hiperespectrais dentre outros (LIU et al., 2018; SALES et al., 2023; CHEN et al., 2025; FOLLI et al., 2024). No caso dos espectros de MIR e NIR, essas técnicas fornecem matrizes de alta dimensionalidade com forte correlação entre variáveis, cenário ideal para o emprego de modelagens multivariadas e geração de amostras sintéticas (SALES et al., 2023; FOLLI et al., 2024). O processo visa manter a estrutura de correlação entre as bandas espectrais, de modo a garantir que as amostras artificiais representem realisticamente as variações químicas e físicas dos sistemas analisados (CHEN et al., 2017; FIGUEIRA e VAZ, 2022; SUN et al., 2024).

Conhecer a natureza destes dados é fundamental para a escolha do método de criação das amostras sintéticas.

As técnicas mais utilizadas para geração de amostras sintéticas incluem SMOTE (Técnica de Superamostragem Sintética da Minoria - *Synthetic Minority Over-sampling Technique*), Algoritmos GAN (Redes Geradoras Adversárias - *Generative Adversarial Networks*), Interpolação Aleatória Linear, Ruído Gaussiano e Análise de Componentes Principais (PCA). O SMOTE baseia-se na criação de novas amostras por interpolação entre observações reais da classe minoritária, equilibrando conjuntos de dados desbalanceados (ARSLAN et al., 2019; KUMMER et al., 2022). Os algoritmos GAN utilizam redes neurais adversárias (uma geradora e uma discriminadora) que competem entre si até que a geradora produza amostras sintéticas indistinguíveis das reais (TANG et al., 2025; SIKIRU, AERNOUTS e ERKINBAEV, 2026). A Interpolação Aleatória Linear combina pares de amostras reais com pesos aleatórios, gerando novos pontos dentro do espaço amostral original (LI et al., 2020; WAN et al., 2023). A adição de Ruído Gaussiano, por sua vez, insere pequenas variações aleatórias controladas aos dados originais, preservando a distribuição estatística (MOMENY et al., 2023). Já a PCA gera amostras no espaço latente das componentes principais, simulando novas observações com base na variabilidade explicada do conjunto original (FIGUEIRA e VAZ, 2022; FOLLI et al., 2024).

Na área de alimentos, a utilização de amostras sintéticas pode contribuir para o desenvolvimento de modelos de classificação e calibração mais estáveis e generalizáveis. Estudos recentes demonstram a eficácia dessa estratégia em análises de autenticidade, detecção de fraudes e segurança alimentar (GEORGOULI et al., 2018; GENG et al., 2022; ZHANG et al., 2023; MOMENY et al., 2023; GAO, XIAO e DENG, Z, 2024). Contudo, até o momento, não há registros na literatura de estudos que empreguem geração de amostras sintéticas especificamente para a detecção de resíduos de antibióticos em leite, embora técnicas de aumento de dados e aprendizado de máquina já venham sendo aplicadas com sucesso na modelagem físico-química e detecção de contaminantes em leite (WANG et al., 2025; SIKIRU, AERNOUTS e ERKINBAEV, 2026). Essa lacuna evidencia uma oportunidade relevante de pesquisa, especialmente para aprimorar modelos de triagem e monitoramento de resíduos em matrizes complexas como o leite.

Assim, o uso de amostras sintéticas em quimiometria aplicada a alimentos representa uma abordagem eficiente para lidar com limitações amostrais, reduzindo custos e ampliando a robustez de modelos preditivos. No entanto, é fundamental garantir que o método de geração

preserve a estrutura estatística dos dados e que as amostras artificiais sejam validadas cuidadosamente, a fim de evitar a introdução de viés.

2.8.1. Geração de amostras sintéticas via PCA

No contexto da geração de amostras sintéticas, a Análise de Componentes Principais é utilizada para projetar os dados originais em um espaço latente, definido pelas componentes principais (PCs). Inicialmente, o conjunto de dados original é centralizado e, se necessário, escalonado, garantindo que todas as variáveis contribuam de forma equilibrada para a decomposição (NIU et al., 2024). Em seguida, os dados são decompostos em duas matrizes fundamentais: a matriz de *scores* (T), que representa as amostras no espaço das componentes principais, e a matriz de *loadings* (P), que descreve o peso de cada variável original em cada componente (FOLLI et al., 2024; GAO et al., 2024). Essa decomposição é expressa pela equação 1, onde E representa a matriz de resíduos.

$$X = T \cdot P^T + E$$

Equação 1

Com base nessa estrutura, as amostras sintéticas são geradas no espaço dos *scores* (T). Cada componente principal, por capturar parte da variância total dos dados, pode ser descrita por uma distribuição estatística, com média e variância determinadas pelos *scores* originais. Assim, o processo de geração consiste em simular novos *scores* artificiais para cada componente. O número de componentes utilizados pode ser escolhido de forma a preservar a maior parte da variância explicada, garantindo que as novas amostras mantenham a coerência estatística do conjunto original (FOLLI et al., 2024; GAO et al., 2024).

Após a simulação, as novas observações são reconstruídas no espaço original das variáveis utilizando a relação inversa da decomposição (Equação 2) em que X_{sint} representa a matriz de amostras sintéticas e T_{sint} os escores gerados artificialmente (GAO et al., 2024). Essa etapa devolve as amostras ao domínio original (por exemplo, espectros de MIR ou NIR), preservando as correlações entre variáveis e a estrutura multivariada típica dos dados espectroscópicos. Em seguida, as amostras sintéticas podem ser incorporadas ao conjunto real, ampliando o espaço amostral.

$$X_{sint} = T_{sint} \cdot P^T$$

Equação 2

Assim, a geração de amostras sintéticas por PCA representa uma forma robusta e controlada de aumento de dados, capaz de aumentar a representatividade amostral e melhorar o desempenho de modelos multivariados em análises de alimentos e outras matrizes químicas complexas.

3. Referências

- ADEGOKE, O. A., QUADRI, M. O. Novel spectrophotometric determinations of some cephalosporins following azo dye formation with p-dimethylaminobenzaldehyde. **Arabian Journal of Chemistry**, v. 9, p. S1272-S1282, 2016.
- AGARWAL, A., SHAH, D., SHEN, D., SONG, D. On robustness of principal component regression. **Advances in Neural Information Processing Systems**, v. 32, 2019.
- AHMAD, A. AYUB, H. Fourier transform infrared spectroscopy (FTIR) technique for food analysis and authentication. In: **Nondestructive quality assessment techniques for fresh fruits and vegetables**. Singapore: Springer Nature Singapore, 2022. p. 103-142.
- AHMED, S., NING, J., PENG, D., CHEN, T., AHMAD, I., ALI, A., YUAN, Z. Current advances in immunoassays for the detection of antibiotics residues: A review. **Food and Agricultural Immunology**, v. 31, n. 1, p. 268-290, 2020.
- AKBARI, E., BAIGBABAELI, A., SHAHIDI, M. Determination of the floral origin of honey based on its phenolic profile and physicochemical properties coupled with chemometrics. **International Journal of Food Properties**, v. 23, n. 1, p. 506-519, 2020.
- AL-AWSEYA, R. A., HASSAN, M. F., FATAH, A. F. M. A., KHALIL, W. F., BASHA, W. A., EL-TARABILI, R. M., GABR, E. M. Efficacy and Safety of Ceftiofur for Treating Serious Respiratory Diseases in Cattle: Clinical, Histopathological, and Microbiological Assessments. **Journal of Advanced Veterinary Research**, v. 13, n. 10, p. 1995-1999, 2023.
- ALI, M. S., NA, S. H., MOON, B. Y., KANG, H. Y., KANG, H. S., KIM, S. J., LIM, S. K. Antimicrobial resistance profiles and molecular characteristics of extended-spectrum β -lactamase-producing *Salmonella enterica* serovar Typhimurium isolates from food animals during 2010–2021 in South Korea. **Foodborne Pathogens and Disease**, v. 21, n. 10, p. 634-642, 2024.
- ALIEV, T. A., BELYAEV, V. E., POMYTKINA, A. V., NESTEROV, P. V., SHITYAKOV, S. V., SADOVNYCHIY, R. V., SKORB, E. V. The development of an electrochemical sensor

for antibiotics in milk based on machine learning algorithms. **arXiv preprint arXiv:2212.04422**, 2022.

ANOWAR, F., SADAoui, S., SELIM, B. Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne). **Computer Science Review**, v. 40, p. 100378, 2021.

ARSLAN, M., GUZEL, M., DEMIRCI, M., OZDEMIR, S. SMOTE and gaussian noise based sensor data augmentation. In: **2019 4th international conference on computer science and engineering (UBMK)**. IEEE, 2019. p. 1-5.

ASRI, M. N. M., VERMA, R., MAHAT, N. A., NOR, N. A. M., DESA, W. N. S. M., ISMAIL, D. Discrimination and source correspondence of black gel inks using Raman spectroscopy and chemometric analysis with UMAP and PLS-DA. **Chemometrics and Intelligent Laboratory Systems**, v. 225, p. 104557, 2022.

AURILIO, C., SANSONE, P., BARBARISI, M., POTA, V., GIACCARI, L. G., COPPOLINO, F., PACE, M. C. Mechanisms of action of carbapenem resistance. **Antibiotics**, v. 11, n. 3, p. 421, 2022.

BABYAK, M. A. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. **Biopsychosocial Science and medicine**, v. 66, n. 3, p. 411-421, 2004.

BAGGIO, D., ANANDA-RAJAH, M. R. Fluoroquinolone antibiotics and adverse events. **Australian prescriber**, v. 44, n. 5, p. 161, 2021.

BARROS, C. M., DI STASI, L. C. Farmacologia veterinária. **1a. ed**, 2012.

BASSETTI, M., POULAKOU, G., RUPPE, E., BOUZA, E., VAN HAL, S. J., BRINK, A. Antimicrobial resistance in the next 30 years, humankind, bugs and drugs: a visionary approach. **Intensive care medicine**, v. 43, p. 1464-1475, 2017.

BEĆ, K. B., GRABSKA, J., HUCK, C. W. Near-infrared spectroscopy in bio-applications. **Molecules**, v. 25, n. 12, p. 2948, 2020.

BEDNARSKI, M., KUPCZYŃSKI, R. Factors Affecting Milk Productivity, Milk Quality and Dairy Cow Health. **Animals**, v. 14, n. 24, p. 3707, 2024.

BEGDACHE, L., KIANMEHR, H., SABOUNCHI, N., MARSZALEK, A., DOLMA, N. Principal component regression of academic performance, substance use and sleep quality

in relation to risk of anxiety and depression in young adults. **Trends in Neuroscience and Education**, v. 15, p. 29-37, 2019.

BELLON-MAUREL, Véronique; MCBRATNEY, Alex. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils—Critical review and research perspectives. **Soil Biology and Biochemistry**, v. 43, n. 7, p. 1398-1410, 2011.

BIANCOLILLO, A; MARINI, F. Chemometric methods for spectroscopy-based pharmaceutical analysis. *Frontiers in chemistry*, v. 6, p. 576, 2018.

BION, C., BECK HENZELIN, A., QU, Y., PIZZOCRI, G., BOLZONI, G., BUFFOLI, E. Analysis of 27 antibiotic residues in raw cow's milk and milk-based products—validation of Delvotest® T. **Food Additives & Contaminants: Part A**, v. 33, n. 1, p. 54-59, 2016.

BRASIL. ANVISA-Agência Nacional de Vigilância Sanitária. Programa de Análise de Resíduos de Medicamentos Veterinários em Alimentos de Origem Animal - PAMVet - Relatório 2006/2007 - Monitoramento de Resíduos em Leite Exposto ao Consumo, 2007.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Valor Bruto da Produção Agropecuária. Ministério da Agricultura, Pecuária e Abastecimento, Brasília, DF, 2021. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/valor-bruto-da-producao-agropecuaria-vbp>. Acesso em: 13 de dezembro de 2022.

BRITO, I. C. P., DA SILVA SANTOS, H. L., SANTOS, K. E. O., DE ALBUQUERQUE FERNANDES, S. A. Modification and validation of miniaturized QuEChERS method for multi-residual determination of antibiotics in milk from a tropical region. **Microchemical Journal**, v. 207, p. 112025, 2024.

BRUNI, S. A. R.; DE OLIVEIRA, V. M. A. T.; FERNANDEZ, A. S. T.; SAKAI, O. A.; MARÇO, P. H.; VALDERRAMA, P. Attenuated total reflectance Fourier transform (ATRFTIR) spectroscopy and chemometrics for organic cinnamon evaluation. *Food Chemistry*, v. 365, p. 130466, 2021.

BRUNO, D. R., CLEALE, R. M., JARDON, G., SHORT, T., MILLS, B., PEDRAZA, J. R. Outcomes after treatment of nonsevere gram-negative clinical mastitis with ceftiofur hydrochloride for 2 or 5 days compared with negative control. **Journal of Dairy Science**, v. 107, n. 4, p. 2390-2405, 2024.

BULA – CIPROLAC OUROFINO SAÚDE ANIMAL LTDA. Indicações, modo de uso e período de carência. Disponível em: Bulário Veterinário. [s.l.]: Ouro Fino; 2016. p. [sem paginação]. Disponível em: Bulário Veterinário. Acesso em: 13 jun. 2025.

BUSARI, A. A., EFEJENE, I. O., OLAYEMI, S. O., ORORORO, O. C., & EGBUNE, E. O. Evaluation of antibiotic use and analysis of ciprofloxacin and gentamicin residue in fish samples from farms in Lagos, Nigeria. **Environmental Monitoring and Assessment**, v. 196, n. 2, p. 127, 2024.

CAIS-SOKOLIŃSKA, D., TEICHERT, J., GAWAŁEK, J. JOLANTA. Foaming and other functional properties of freeze-dried water milk. **Food**, v. 12, n. 11, pág. 2274, 2023.

CALAHORRANO-MORENO, M. B, ORDOÑEZ-BAILON, J. J, BAQUERIZO-CRESPO, R. J, DUEÑAS-RIVADENEIRA, A. A, MONTENEGRO, M. C. B., RODRÍGUEZ-DÍAZ, J. M. Contaminants in the cow's milk we consume? Pasteurization and other technologies in the elimination of contaminants. **F1000Research**, v. 11, p. 91, 2022.

CANDEIAS, D. N. C., SILVA, K. M., PEREIRA, H. S., BEZERRA, L. P., DA SILVA, J. D. S., DE SOUSA FERNANDES, D. D., DINIZ, P. H. G. D. Geographical origin authentication of instant coffee from southern Bahia using MIR and NIR spectroscopy coupled with DD-SIMCA. **Food Chemistry**, v. 479, p. 143698, 2025.

CANESCHI, A., BARDHI, A., BARBAROSSA, A., ZAGHINI, A. The use of antibiotics and antimicrobial resistance in veterinary medicine, a complex phenomenon: A narrative review. **Antibiotics**, v. 12, n. 3, p. 487, 2023.

CASARRUBIAS-TORRES, L. M., MEZA-MÁRQUEZ, O. G., OSORIO-REVILLA, G., GALLARDO-VELAZQUEZ, T. Mid-infrared spectroscopy and multivariate analysis for determination of tetracycline residues in cow's milk. **Acta Veterinaria Brno**, v. 87, n. 2, p. 181-188, 2018.

CHAUHAN, S., POWAR, P., MEHRA, R. A review on nutritional advantages and nutraceutical properties of cow and goat milk. **Int. J. Appl. Res**, v. 7, n. 10, p. 101-105, 2021.

CHEN, H., TAN, C., QIU, S., TAN, C., XUE, Z. Qualitative and quantitative detection of tartrazine in tea powder by mid-infrared spectroscopy and an ensemble of virtual sample generation-based calibration. **Journal of Food Composition and Analysis**, p. 108096, 2025.

CHEN, T., CHENG, G., AHMED, S., WANG, Y., WANG, X., HAO, H., YUAN, Z. New methodologies in screening of antibiotic residues in animal-derived foods: Biosensors. **Talanta**, v. 175, p. 435-442, 2017.

CHEN, Z. S., ZHU, B., HE, Y. L., YU, L. A. A PSO based virtual sample generation method for small sample sets: Applications to regression datasets. **Engineering Applications of Artificial Intelligence**, v. 59, p. 236-243, 2017.

CHINEMEREM NWOBODO, D., UGWU, M. C., OLISELOKE ANIE, C., AL-OUQAILI, M. T., CHINEDU IKEM, J., VICTOR CHIGOZIE, U., SAKI, M. Antibiotic resistance: The challenges and some emerging strategies for tackling a global menace. **Journal of clinical laboratory analysis**, v. 36, n. 9, p. e24655, 2022.

CHRISTENSEN, S. B. Drugs that changed society: History and current status of the early antibiotics: Salvarsan, sulfonamides, and β -lactams. **Molecules**, v. 26, n. 19, p. 6057, 2021.

ČIKOVSKÝ, S., ŠVÁB, P., HANÁK, P. Hybrid PCA-Based and Machine Learning Approaches for Signal-Based Interference Detection and Anomaly Classification Under Synthetic Data Conditions. **Sensors**, v. 25, n. 24, p. 7585, 2025.

COELHO-BARROS, E. A.; SIMÕES, P. A.; ACHCAR, J. A.; MARTINEZ, E. Z.; SHIMANO, A. C. Métodos de estimação em regressão linear múltipla: aplicação a dados clínicos. **Revista Colombiana de Estadística**, v. 31, n. 1, p. 111-129, jun. 2008.

COIMBRA, P. T., BATHAZAR, C. F., GUIMARÃES, J. T., COUTINHO, N. M., PIMENTEL, T. C., NETO, R. P., CRUZ, A. G. Detection of formaldehyde in raw milk by time domain nuclear magnetic resonance and chemometrics. **Food Control**, v. 110, p. 107006, 2020.

CONCEIÇÃO, D. G., REIS, L. C. C., SILVA VIRGENS, E., JESUS, J. C., BAUER, L. C., CHAVES, M. A., FERRÃO, S. P. B. Thermal, structural and spectroscopic properties of powder sheep milk obtained by different drying methods. **International Dairy Journal**, v. 150, p. 105845, 2024.

COQUEIRO, J. S., LIMA, A. B. S., JESUS, J. C., SILVA, R. R., FERRÃO, S. P. B., SANTOS, L. S. Ensuring authenticity of cinnamon powder: Detection of adulteration with coffee husk and corn meal using NIR, MIR spectroscopy and chemometrics. **Food Control**, v. 166, p. 110681, 2024.

DONG, L., MENG, L., LIU, H., WU, H., SCHROYEN, M., ZHENG, N., WANG, J. Effect of cephalosporin treatment on the microbiota and antibiotic resistance genes in feces of dairy cows with clinical mastitis. **Antibiotics**, v. 11, n. 1, p. 117, 2022.

DU, L., LU, W., CAI, Z. J., BAO, L., HARTMANN, C., GAO, B., YU, L. L. Rapid detection of milk adulteration using intact protein flow injection mass spectrometric fingerprints combined with chemometrics. **Food chemistry**, v. 240, p. 573-578, 2018.

DU, Z., TIAN, W., TILLEY, M., WANG, D., ZHANG, G., LI, Y. Quantitative assessment of wheat quality using near-infrared spectroscopy: A comprehensive review. **Comprehensive Reviews in Food Science and Food Safety**, v. 21, n. 3, p. 2956-3009, 2022.

DUPONT, M. F., ELBOURNE, A., COZZOLINO, D., CHAPMAN, J., TRUONG, V. K., CRAWFORD, R. J., LATHAM, K. Chemometrics for environmental monitoring: a review. **Analytical Methods**, v. 12, n. 38, p. 4597-4620, 2020.

DUREL, Luc; GALLINA, Guglielmo; PELLET, Terence. Assessment of ceftiofur residues in cow milk using commercial screening test kits. **Veterinary Record Open**, v. 6, n. 1, p. e000329, 2019.

EL MAOUARDI, M., BRAEKELEER, K., BOUKLOUZE, A., VANDER HEYDEN, Y. Comparison of Near-Infrared and Mid-Infrared spectroscopy for the identification and quantification of argan oil adulteration through PCA, PLS-DA and PLS. **Food Control**, v. 165, p. 110671, 2024.

EMBRAPA – Anuário Leite, 2024. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1164754/anoario-leite-2024-avaliacao-genetica-multirracial>. Acesso em 02 de abril de 2025.

ERLINGE, D., MAEHARA, A., BEN-YEHUDA, O., BØTKER, H. E., MAENG, M., KJØLLER-HANSEN, L., ZHOU, Z. Identification of vulnerable plaques and patients by intracoronary near-infrared spectroscopy and ultrasound (PROSPECT II): a prospective natural history study. **The Lancet**, v. 397, n. 10278, p. 985-995, 2021.

FADLELMOULA, A., PINHO, D., CARVALHO, V. H., CATARINO, S. O., MINAS, G. Fourier transform infrared (FTIR) spectroscopy to analyse human blood over the last 20 years: a review towards lab-on-a-chip devices. **Micromachines**, v. 13, n. 2, p. 187, 2022.

FAN, S., FOSTER, D., MILLER, W. G., OSBORNE, J., KATHARIOU, S. Impact of ceftiofur administration in steers on the prevalence and antimicrobial resistance of *Campylobacter* spp. **Microorganisms**, v. 9, n. 2, p. 318, 2021.

FAO – Food and Agriculture Organization of the United Nations. *FAOSTAT – Production quantities of whole milk powder by country, 2022*. Disponível em: <https://www.fao.org/faostat/en/#data/QCL/visualize>. Acesso em: 16 jun. 2025.

FATTAHI, S. H., KAZEMI, A., KHOJASTEHNZHAND, M., ROOSTAEI, M., MAHMOUDI, A. The classification of Iranian wheat flour varieties using FT-MIR spectroscopy and chemometrics methods. **Expert Systems with Applications**, v. 239, p. 122175, 2024.

FIGUEIRA, A., VAZ, B. Survey on synthetic data generation, evaluation methods and GANs. **Mathematics**, v. 10, n. 15, p. 2733, 2022.

FLORES-VALDEZ, M., MEZA-MÁRQUEZ, O. G., OSORIO-REVILLA, G., GALLARDO-VELÁZQUEZ, T. Identification and quantification of adulterants in coffee (*Coffea arabica* L.) using FT-MIR spectroscopy coupled with chemometrics. **Foods**, v. 9, n. 7, p. 851, 2020.

FOLLI, G. S., NASCIMENTO, M. H., LOVATTI, B. P., ROMÃO, W., FILGUEIRAS, P. R. A generation of synthetic samples and artificial outliers via principal component analysis and evaluation of predictive capability in binary classification models. **Chemometrics and Intelligent Laboratory Systems**, v. 251, p. 105154, 2024.

FOSTER, D. M., JACOB, M. E., FARMER, K. A., CALLAHAN, B. J., THERIOT, C. M., KATHARIOU, S., PAPICH, M. G. Ceftiofur formulation differentially affects the intestinal drug concentration, resistance of fecal *Escherichia coli*, and the microbiome of steers. **PLoS One**, v. 14, n. 10, p. e0223378, 2019.

FRANZOI, M., NIERO, G., VISENTIN, G., PENASA, M., CASSANDRO, M., MARCHI, M. Variation of detailed protein composition of cow milk predicted from a large database of mid-infrared spectra. **Animals**, v. 9, n. 4, p. 176, 2019.

FREITAS OLIVEIRA, T., SANTOS, G. R., JESUS, J. C., COQUEIRO, J. S., SILVA, R. R., FERRÃO, S. P. B., SANTOS, L. S. Assessing the authenticity of the sport supplement creatine using spectroscopic techniques and multivariate analysis. **Food Chemistry**, v. 484, p. 144282, 2025.

FREITAS, A. G. M., MINHO, L. A. C., DE MAGALHÃES, B. E. A., DOS SANTOS, W. N. L., SANTOS, L. S., DE ALBUQUERQUE FERNANDES, S. A. Infrared spectroscopy combined with random forest to determine tylosin residues in powdered milk. **Food Chemistry**, v. 365, p. 130477, 2021.

GAO, S., JIANG, X., ZHANG, Y., LIU, X., XIONG, Q., CAI, Z. Superpixelwise PCA based data augmentation for hyperspectral image classification. **Multimedia Tools and Applications**, v. 83, n. 34, p. 81209-81229, 2024.

GAO, X., XIAO, Z., DENG, Z. High accuracy food image classification via vision transformer with data augmentation and feature augmentation. **Journal of food engineering**, v. 365, p. 111833, 2024.

GARCIA, S. R.P.; CHAVES-NETO, A.; ORO, S. R.; MAFIOLETI, T. R.; NEUMANN-JÚNIOR, C. Modelo de regressão por mínimos quadrados parciais para dados de monitoramento de barragens. **Revista Engenho**, v. 11, p. 1-20, 2015.

GENG, Z., DUAN, X., LI, J., CHU, C., HAN, Y. Risk prediction model for food safety based on improved random forest integrating virtual sample. **Engineering Applications of Artificial Intelligence**, v. 116, p. 105352, 2022.

GEORGOULI, K., OSORIO, M. T., MARTINEZ DEL RINCON, J., KOIDIS, A. Data augmentation in food science: Synthesising spectroscopic data of vegetable oils for performance enhancement. **Journal of Chemometrics**, v. 32, n. 6, p. e3004, 2018.

GETAHUN, M., ABEBE, R. B., SENDEKIE, A. K., WOLDEYOHANIS, A. E., KASAHUN, A. E. Evaluation of antibiotics residues in milk and meat using different analytical methods. **International journal of analytical chemistry**, v. 2023, n. 1, p. 4380261, 2023.

GRELET, C., LARSEN, T., CROWE, M. A., WATHES, D. C., FERRIS, C. P., INGVAERTSEN, K. L. Prediction of key milk biomarkers in dairy cows through milk mid-infrared spectra and international collaborations. **Journal of Dairy Science**, v. 107, n. 3, p. 1669-1684, 2024.

GUERRERO-PÉREZ, M. O., PATIENCE, G. S. Experimental methods in chemical engineering: Fourier transform infrared spectroscopy—FTIR. **The Canadian Journal of Chemical Engineering**, v. 98, n. 1, p. 25-33, 2020.

HAAS, J., MIZAIKOFF, B. Advances in mid-infrared spectroscopy for chemical analysis. **Annual Review of Analytical Chemistry**, v. 9, n. 1, p. 45-68, 2016.

HADDAD, M. F., ABDULLAH, B. A., ALOBEIDI, H. A., SAADI, A. M., HADDAD, M. F. Antibiotic classification, mechanisms, and indications: A review. **International Journal of Medical and All Body Health Research**, v. 5, n. 3, p. 39-46, 2024.

HARDEFELDT, L. Y., PRESCOTT, J. F. Beta-lactam Antibiotics: Cephalosporins. **Antimicrobial Therapy in Veterinary Medicine**, p. 143-167, 2024.

HASHIMOTO, K., HOROSHKO, D. B., KOLOBOV, M. I., MICHAEL, Y., GEFEN, Z., CHEKHOVA, M. V. Fourier-transform infrared spectroscopy with undetected photons from high-gain spontaneous parametric down-conversion. **Communications Physics**, v. 7, n. 1, p. 217, 2024.

HASOŇOVÁ, L., SAMKOVA, E., BALDIKOVA, E., REINDL, K., NEJESCHLEBOVA, H., HONESOVÁ, S. J., HANUŠ, O. Occurrence and seasonal variation of aflatoxin M1 in raw cow milk in the Czech Republic, Central Europe. **Food Control**, v. 170, p. 111028, 2025.

HASSEN, A., AHMED, R., ALAM, M. S., CHAVULA, P., MOHAMMED, S. S., DAWID, A. The effect of feed supplementation on cow milk productivity and quality: a brief study. **Int. J. Agric. Vet. Sci**, v. 4, n. 1, p. 13-25, 2022.

HELBLING, I. M., BUSATTO, C. A., KARP, F., ISLAN, G. A., ESTENOZ, D. A., LUNA, J. A. An analysis of the microencapsulation of ceftiofur in chitosan particles using the spray drying technology. **Carbohydrate Polymers**, v. 234, p. 115922, 2020.

HOU, X., LV, S., CHEN, Z., XIAO, F. Applications of Fourier transform infrared spectroscopy technologies on asphalt materials. **Measurement**, v. 121, p. 304-316, 2018.

HOUHOU, R., BOCKLITZ, T. Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. **Analytical Science Advances**, v. 2, n. 3-4, p. 128-141, 2021.

HU, J., XU, Z., LI, M., HE, Y., LIU, Y. Discriminant analysis and quantitative study of antibiotics in infant milk powder based on hyperspectral detection. **Vibrational Spectroscopy**, v. 114, p. 103244, 2021.

HU, J., XU, Z., LI, M., HE, Y., LIU, Y. Discriminant analysis and quantitative study of antibiotics in infant milk powder based on hyperspectral detection. **Vibrational Spectroscopy**, v. 114, p. 103244, 2021.

IBGE - 2024 registra recorde no abate de bovinos, frangos e suínos.
<https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de->

noticias/releases/42898-2024-registra-recorde-no-abate-de-bovinos-frangos-e-suinos#:~:text=A%20aquisi%C3%A7%C3%A3o%20de%20leite%20cru,3%C2%B0%20trimestre%20de%202024. Acesso em: 01 de abril de 2025.

JAGESAR, D. C., KERKHOF, J. Validation of Delvotest® Fast BT for Detection of β -Lactams and Tetracyclines in Raw Cow's Milk: Performance Tested Method SM 102104. **Journal of AOAC International**, v. 106, n. 1, p. 88-98, 2023.

JESUS, J. C. D., SILVA, G. J., GONÇALVES, B. H. R., SOUZA, M. R. D., SANTOS, L. S., FERRÃO, S. P. Quick identification of the time of maturation of artisanal minas cheese by FTIR-ATR spectroscopy and multivariate techniques. **Journal of the Brazilian Chemical Society**, v. 31, n. 10, p. 2000-2011, 2020.

JOLLIFFE, I. T., CADIMA, J. Principal component analysis: a review and recent developments. **Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences**, v. 374, n. 2065, p. 20150202, 2016.

KHWAZA, V., MLALA, S., & ADERIBIGBE, B. A. Advancements in synthetic strategies and biological effects of ciprofloxacin derivatives: A Review. **International Journal of Molecular Sciences**, v. 25, n. 9, p. 4919, 2024.

KIM, D., KIM, S., KWON, Y., KIM, Y., PARK, H., KWAK, K., KANG, L. W. Structural insights for β -lactam antibiotics. **Biomolecules & Therapeutics**, v. 31, n. 2, p. 141, 2023.

KUMAR, A., PANDA, A. K., SHARMA, N. Determination of antibiotic residues in bovine milk by HPLC-DAD and assessment of human health risks in Northwestern Himalayan region, India. **Journal of Food Science and Technology**, v. 59, n. 1, p. 95-104, 2022.

KUMMER, A., RUPPERT, T., MEDVEGY, T., ABONYI, J. Machine learning-based software sensors for machine state monitoring-The role of SMOTE-based data augmentation. **Results in Engineering**, v. 16, p. 100778, 2022.

LAMMIE, S. L., HUGHES, J. M. Antimicrobial resistance, food safety, and one health: the need for convergence. **Annual review of food science and technology**, v. 7, n. 1, p. 287-312, 2016.

LASALVIA, M., CAPOZZI, V., PERNA, G. A comparison of PCA-LDA and PLS-DA techniques for classification of vibrational spectra. **Applied Sciences**, v. 12, n. 11, p. 5345, 2022.

- LEE, K. M., YARBROUGH, D., KOZMAN, M. M., HERRMAN, T. J., PARK, J., WANG, R., KUROUSKI, D. Rapid detection and prediction of chlortetracycline and oxytetracycline in animal feed using surface-enhanced Raman spectroscopy (SERS). **Food Control**, v. 114, p. 107243, 2020.
- LI, D. C., HU, S. C., LIN, L. S., YEH, C. W. Detecting representative data and generating synthetic samples to improve learning accuracy with imbalanced data sets. **PloS one**, v. 12, n. 8, p. e0181853, 2017.
- LI, S., LIU, W., LIU, M., CHEN, Y., ZHANG, F., WANG, X. A sensitive lateral flow immunoassay relying on time-resolved fluorescent microspheres immune probe for determination of ceftiofur and its metabolite. **Talanta**, v. 271, p. 125580, 2024.
- LI, Z., GUO, J., JIAO, W., XU, P., LIU, B., ZHAO, X. Random linear interpolation data augmentation for person re-identification. **Multimedia Tools and Applications**, v. 79, n. 7, p. 4931-4947, 2020.
- LIAO, M., WU, J., LI, Y., LU, X., TAN, H., CHEN, S., ZHANG, R. Prevalence and persistence of ceftiofur-resistant *Escherichia coli* in a chicken layer breeding program. **Animals**, v. 13, n. 1, p. 90, 2022.
- LIMA, A. B. S., BATISTA, A. S., JESUS, J. C., JESUS SILVA, J., DE ARAÚJO, A. C. M., SANTOS, L. S. Fast quantitative detection of black pepper and cumin adulterations by near-infrared spectroscopy and multivariate modeling. **Food Control**, v. 107, p. 106802, 2020.
- LIMA, L. M., DA SILVA, B. N. M., BARBOSA, G., BARREIRO, E. J. An overview from a medicinal chemistry perspective. **European journal of medicinal chemistry**, v. 208, p. 112829, 2020.
- LIN, X., KÜCK, U. Cephalosporins as key lead generation beta-lactam antibiotics. **Applied microbiology and biotechnology**, v. 106, n. 24, p. 8007-8020, 2022.
- LIU, C., ZHANG, X., NGUYEN, T. T., LIU, J., WU, T., LEE, E., TU, X. M. Partial least squares regression and principal component analysis: similarity and differences between two popular variable reduction approaches. **General psychiatry**, v. 35, n. 1, p. e100662, 2022.
- LIU, X., SUN, Q., MENG, Y., FU, M., BOURENNANE, S. Hyperspectral image classification based on parameter-optimized 3D-CNNs combined with transfer learning and virtual samples. **Remote Sensing**, v. 10, n. 9, p. 1425, 2018.

LOW, C. X., TAN, L. T. H., AB MUTALIB, N. S., PUSPARAJAH, P., GOH, B. H., CHAN, K. G., LEE, L. H. Unveiling the impact of antibiotics and alternative methods for animal husbandry: A review. **Antibiotics**, v. 10, n. 5, p. 578, 2021.

LUIZ, L. D. C., BELL, M. J. V., ROCHA, R. A. D., LEAL, N. L., ANJOS, V. D. C. D. Detection of veterinary antimicrobial residues in milk through near-infrared absorption spectroscopy. **Journal of Spectroscopy**, v. 2018, n. 1, p. 5152832, 2018.

LUNGU, I. A., MOLDOVAN, O. L., BIRIȘ, V., RUSU, A. Fluoroquinolones hybrid molecules as promising antibacterial agents in the fight against antibacterial resistance. **Pharmaceutics**, v. 14, n. 8, p. 1749, 2022.

MAHMOOD, T., ABBAS, M., ILYAS, S., AFZAL, N., NAWAZ, R. Quantification of fluoroquinolone (enrofloxacin, norfloxacin and ciprofloxacin) residues in cow milk. **Ijcsbs**, v. 10, p. 10-15, 2016.

MARINS, L. et al. Negatively controlled, randomized clinical trial to evaluate a culture-based selective antimicrobial treatment of non-severe clinical mastitis on behavior of dairy cows. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 76, n. 5, p. e13129, 2024.

MAYERHÖFER, T. G., ILCHENKO, O., KUTSYK, A., POPP, J. Complex-Valued Chemometrics in Spectroscopy: Principal Component Regression. **Applied Spectroscopy**, p. 00037028251393273, 2025.

MENDES, E., DUARTE, N. Mid-infrared spectroscopy as a valuable tool to tackle food analysis: A literature review on coffee, dairies, honey, olive oil and wine. **Foods**, v. 10, n. 2, p. 477, 2021.

MENTA, P. R., FERNANDES, L., PRIM, J., DE OLIVEIRA, E., LIMA, F., GALVÃO, K. N., MACHADO, V. S. A randomized controlled trial evaluating the efficacy of systemic ceftiofur administration for metritis therapy in dairy cows and the effect of metritis cure on economically important outcomes. **Journal of Dairy Science**, v. 107, n. 9, p. 7092-7105, 2024.

MOHAMED, M. A., JAAFAR, J., ISMAIL, A. F., OTHMAN, M. H. D., RAHMAN, M. A. Fourier transform infrared (FTIR) spectroscopy. In: **Membrane characterization**. elsevier, 2017. p. 3-29.

MOHD SALLEH, N. A., HASSAN, M. S. Discrimination of Lard and other Edible Fats after Heating Treatments using Partial Least Square Regression (PLSR), Principal Component

Regression (PCR) and Linear Support Vector Machine Regression (SVMR). In: **Journal of Physics: Conference Series**. IOP Publishing, 2019. p. 012114.

MOMENY, M., NESHAT, A. A., JAHANBAKHSHI, A., MAHMOUDI, M., AMPATZIDIS, Y., RADEVA, P. Grading and fraud detection of saffron via learning-to-augment incorporated Inception-v4 CNN. **Food Control**, v. 147, p. 109554, 2023.

MOTEVALIZADEH, S. M.; MOLLENHAUER, K.; WETEKAM, J. FTIR spectroscopy and multivariate discriminant analysis for classifying bituminous mastics: Exploring aging states and mastic composition. **Construction and Building Materials**, v. 438, p. 137188, 2024.

MOUDGIL, P., BEDI, J. S., AULAKH, R. S., GILL, J. P. S., KUMAR, A. Validation of HPLC multi-residue method for determination of fluoroquinolones, tetracycline, sulphonamides and chloramphenicol residues in bovine milk. **Food Analytical Methods**, v. 12, n. 2, p. 338-346, 2019.

MUTEEB, G., REHMAN, M. T., SHAHWAN, M., AATIF, M. Origin of antibiotics and antibiotic resistance, and their impacts on drug development: A narrative review. **Pharmaceuticals**, v. 16, n. 11, p. 1615, 2023.

NAVID, P., AKBARI-HASANJANI, H. R. Application of Raman and ATR-MIR Spectroscopy in Colorectal Cancer Diagnosis Combined with Chemometrics Techniques: A Review. **Critical Reviews in Analytical Chemistry**, p. 1-22, 2025.

NIU, W., HUERTAS, A., LI, J., STILLER, B. Enhancing synthetic data generation for class imbalance and privacy preservation. In: **2024 IEEE International Conference on Big Data (BigData)**. IEEE, 2024. p. 7104-7109.

NTAKATSANE, M., CHEN, P., LIU, J., MOSEBI, P., XU, L., MATEBESI, P., WANG, Y. Multi-dimensional fluorescence spectroscopy coupled with chemometrics in rapid antibiotic detection and discrimination. **Journal of Food Measurement and Characterization**, v. 14, n. 4, p. 1892-1900, 2020.

OLIVERI, P., MALEGORI, C., CASALE, M. Chemometrics: Multivariate analysis of chemical data. In: **Chemical analysis of food**. Academic Press, 2020. p. 33-76.

PACHECO-SILVA, E., SOUZA, J. R., CALDAS, E. D. Resíduos de medicamentos veterinários em leite e ovos. **Química nova**, v. 37, p. 111-122, 2014.

PAGES-REBULL, J., PÉREZ-RÀFOLS, C., SERRANO, N., DEL VALLE, M., DÍAZ-CRUZ, J. M. Classification and authentication of spices and aromatic herbs by means of HPLC-UV and chemometrics. **Food Bioscience**, v. 52, p. 102401, 2023.

PAIANO, R. B., MORRISON, E. I., LEBLANC, S. J. Randomized clinical trial of ketoprofen or ceftiofur for treatment of metritis in dairy cows. **Journal of Dairy Science**, v. 107, n. 10, p. 8366-8377, 2024.

PELYUNTHA, W., SANGUANKIAT, A., KOVITVADHI, A., VONGKAMJAN, K. Broad lytic spectrum of novel Salmonella phages on ciprofloxacin-resistant Salmonella contaminated in the broiler production chain. **Veterinary World**, v. 15, n. 8, p. 2039, 2022.

PEREIRA, E. V. D., FERNANDES, D. D. S., DE ARAÚJO, M. C. U., DINIZ, P. H. G. D., MACIEL, M. I. S. In-situ authentication of goat milk in terms of its adulteration with cow milk using a low-cost portable NIR spectrophotometer. **Microchemical Journal**, v. 163, p. 105885, 2021.

PÖGE, A., REINECKE, J. Principal component analysis. In: **Handbook of Computational Social Science, Volume 2**. Routledge, 2021. p. 322-333.

QIAO, M., YING, G. G., SINGER, A. C., ZHU, Y. G. Review of antibiotic resistance in China and its environment. **Environment international**, v. 110, p. 160-172, 2018.

RAHEEM, WS, NIAMAH, A. Contamination methods of milk with pesticides residues and veterinary drugs. In: **IOP Conference Series: Earth and Environmental Science**. IOP Publishing, 2021. p. 012003.

RODRÍGUEZ-HERNÁNDEZ, P., NÚÑEZ-SÁNCHEZ, N., MOLINA-GAY, S., RODRÍGUEZ-ESTÉVEZ, V., CARDOSO-TOSET, F. Near infrared spectroscopy analysis as a screening tool to classify milk from bovine subclinical mastitis and promote pathogen-based therapy. **Applied Food Research**, v. 5, n. 1, p. 100651, 2025.

ROHMAN, A., GHAZALI, M. A. I. B., WINDARSIH, A., RIYANTO, S., YUSOF, F. M., MUSTAFA, S. Comprehensive review on application of FTIR spectroscopy coupled with chemometrics for authentication analysis of fats and oils in the food products. **Molecules**, v. 25, n. 22, p. 5485, 2020.

RUIZ-PEREZ, D., GUAN, H., MADHIVANAN, P., MATHEE, K., NARASIMHAN, G. So you think you can PLS-DA?. **BMC bioinformatics**, v. 21, n. Suppl 1, p. 2, 2020.

- RUSU, A., MUNTEANU, A. C., ARBĂNAȘI, E. M., UIVAROSI, V. Overview of side-effects of antibacterial fluoroquinolones: new drugs versus old drugs, a step forward in the safety profile?. **Pharmaceutics**, v. 15, n. 3, p. 804, 2023.
- RYAN, C. A., MCNEAL, C. D., CREDILLE, B. C. Ceftiofur use and antimicrobial stewardship in the horse. **Equine Veterinary Journal**, v. 55, n. 6, p. 944-961, 2023.
- SAJI, R., RAMANI, A., GANDHI, K., SETH, R., SHARMA, R. Application of FTIR spectroscopy in dairy products: A systematic review. **Food and Humanity**, v. 2, p. 100239, 2024.
- SALES, R. D. F., BARBOSA-PATRÍCIO, L. C., DA SILVA, N. C., E BRITO, L. R., DA SILVA, M. E. F., PIMENTEL, M. F. Gasoline discrimination using infrared spectroscopy and virtual samples based on measurement uncertainty. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 303, p. 123248, 2023.
- SANTANA, V. R., MAGALHÃES, B. E. A. D., MINHO, L. A. C., FREITAS, A. G. M. D., SOUZA, R. D., BATISTA, A. S., FERNANDES, S. A. D. A. Random forest and infrared spectroscopy to detect contamination of freeze-dried milk by ivermectin residues. **Ciência Rural**, v. 55, n. 12, p. e20240536, 2025.
- SANTOS, C. P. D. S., BARBOSA, R. S., SOUZA, R. D., FERNANDES, S. A. A. Mid-infrared spectroscopy combined with chemometrics for the detection of ciprofloxacin and Bronopol® residues in milk. **Food Chemistry**, v. 504, p. 147826, 2026.
- SANTOS, D. P., SENA, M. M., ALMEIDA, M. R., MAZALI, I. O., OLIVIERI, A. C., VILLA, J. E. Unraveling surface-enhanced Raman spectroscopy results through chemometrics and machine learning: principles, progress, and trends. **Analytical and bioanalytical chemistry**, v. 415, n. 18, p. 3945-3966, 2023.
- SCHMENGER, A., KRÖMKER, V. Characterization, cure rates and associated risks of clinical mastitis in Northern Germany. **Veterinary sciences**, v. 7, n. 4, p. 170, 2020.
- SHAWISH, R. R., ELTAHAN, A., ELBAYOUMI, Z. H. Prevalence, Population Structure and Antibiotic Resistance Patterns of Ceftiofur Resistant Enterobacteriaceae from Chicken Meat and Fish. **Journal of Current Veterinary Research**, v. 5, n. 2, p. 27-39, 2023.
- SHI, R., YU, Z., HO, H., WU, W., WANG, J., TIAN, X., HAN, R. Survey of aflatoxin M1 in commercial liquid milk products in China. **Journal of Food Protection**, v. 84, n. 2, p. 200-203, 2021.

SIKIRU, Y., AERNOUTS, B., & ERKINBAEV, C. Conditional Generation of Near-Infrared Spectra Using a Transformer-Based GAN for Spectral Data Augmentation in Milk Composition. **Food Control**, p. 112042, 2026.

SILVER, L. L. Appropriate targets for antibacterial drugs. **Cold Spring Harbor perspectives in medicine**, v. 6, n. 12, p. a030239, 2016.

SINGH, B., BHAT, A., DUTTA, L., PATI, K. R., KORPAN, Y., DAHIYA, I. Electrochemical biosensors for the detection of antibiotics in milk: recent trends and future perspectives. **Biosensors**, v. 13, n. 9, p. 867, 2023.

SOLTANA, G., SABETZADEH, M., BRIAND, L. C. Synthetic data generation for statistical testing. In: **2017 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE)**. IEEE, 2017. p. 872-882.

SONG, Y., CONG, Y., WANG, B., ZHANG, N. Applications of Fourier transform infrared spectroscopy to pharmaceutical preparations. **Expert opinion on drug delivery**, v. 17, n. 4, p. 551-571, 2020.

SOTNIKOV, D. V., ZHERDEV, A. V., ZVEREVA, E. A., EREMIN, S. A., DZANTIEV, B. B. Changing cross-reactivity for different immunoassays using the same antibodies: Theoretical description and experimental confirmation. **Applied Sciences**, v. 11, n. 14, p. 6581, 2021.

SU, K. Y., LEE, W. L. Fourier transform infrared spectroscopy as a cancer screening and diagnostic tool: A review and prospects. **Cancers**, v. 12, n. 1, p. 115, 2020.

SUN, F., CHEN, Y., WANG, K. Y., WANG, S. M., LIANG, S. W. Identification of genuine and adulterated pinellia ternata by mid-infrared (MIR) and near-infrared (NIR) spectroscopy with partial least squares-discriminant analysis (PLS-DA). **Analytical Letters**, v. 53, n. 6, p. 937-959, 2020.

SUN, Z., YAO, Q., JIN, H., XU, Y., HANG, W., CHEN, H., SHEN, X. A novel in-situ sensor calibration method for building thermal systems based on virtual samples and autoencoder. **Energy**, v. 297, p. 131314, 2024.

TANG, M., XU, K., SUN, Y., ZHANG, Y., SHE, W., HU, J., TIE, J. Enhancing Infrared Spectroscopy Analysis Based on Data Augmentation and Deep Learning. In: **International Conference on Intelligent Computing**. Singapore: Springer Nature Singapore, 2025. p. 268-280.

TEIXEIRA, R. C., LUIZ, L. C., JUNQUEIRA, G. M. A., BELL, M. J. V., ANJOS, V. C. Detection of antibiotic residues in Cow's milk: A theoretical and experimental vibrational study. **Journal of Molecular Structure**, v. 1215, p. 128221, 2020.

TEMPINI, P. N., ALY, S. S., KARLE, B. M., PEREIRA, R. V. Multidrug residues and antimicrobial resistance patterns in waste milk from dairy farms in Central California. **Journal of dairy science**, v. 101, n. 9, p. 8110-8122, 2018.

TEYE, E., AMUAH, C. L., MCGRATH, T., ELLIOTT, C. Innovative and rapid analysis for rice authenticity using hand-held NIR spectrometry and chemometrics. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 217, p. 147-154, 2019.

THAI, V. A., THUY, N. T., PANDIT, B., VO, T. K. Q., KHEDULKAR, A. P. Fluoroquinolones: Fate, effects on the environment and selected removal methods. **Journal of Cleaner Production**, v. 418, p. 137762, 2023.

TIE, Y., DUCHATEAU, C., VAN DE STEENE, S., MEES, C., DE BRAEKELEER, K., DE BEER, T., DECONINCK, E. Spectroscopic techniques combined with chemometrics for fast on-site characterization of suspected illegal antimicrobials. **Talanta**, v. 217, p. 121026, 2020.

TKACHENKO, Y., NIEDZIELSKI, P. FTIR as a method for qualitative assessment of solid samples in geochemical research: a review. **Molecules**, v. 27, n. 24, p. 8846, 2022.

TODOROV, H., SEARLE-WHITE, E., GERBER, S. Applying univariate vs. multivariate statistics to investigate therapeutic efficacy in (pre) clinical trials: A Monte Carlo simulation study on the example of a controlled preclinical neurotrauma trial. **PLoS One**, v. 15, n. 3, p. e0230798, 2020.

TOMMASONI, C., FIORE, E., LISUZZO, A., GIANESELLA, M. Mastitis in dairy cattle: On-farm diagnostics and future perspectives. **Animals**, v. 13, n. 15, p. 2538, 2023.

USDA – United States Department of Agriculture. **Production - Fluid Milk**. Disponível em: <https://www.fas.usda.gov/data/production/commodity/0223000>. Acesso em: 01 de abril de 2025.

VASCO, K. A., BOWCUTT, B., CARBONELL, S., SOUZA, L., ROBISON, C., ABUELO, A., MANNING, S. D. Selection of antibiotic-resistant bacterial populations in the dairy cow gut following intramuscular ceftiofur treatment for metritis. **Journal of Dairy Science**, v. 107, n. 12, p. 11536-11549, 2024.

VERCELLI, C., AMADORI, M., GAMBINO, G., RE, G. A review on the most frequently used methods to detect antibiotic residues in bovine raw milk. **International Dairy Journal**, v. 144, p. 105695, 2023.

WAN, D., LU, R., XU, T., SHEN, S., LANG, X., REN, Z. Random interpolation resize: A free image data augmentation method for object detection in industry. **Expert Systems with Applications**, v. 228, p. 120355, 2023.

WANG, K. Z., HAN, C. Y., FENG, X. Z., HAN, G. C., KRAATZ, H. B. Electrochemical sensor based on the synergistic effect of copper nanoparticles and poly-L-histidine for cefquinome detection and its oxidation mechanism. **Microchemical Journal**, v. 208, p. 112451, 2025.

WANG, X., WANG, T., JI, R., CHEN, H., QIN, H., HUANG, Z. Identification of goat milk adulterated with cow milk based on total synchronous fluorescence spectroscopy combined with CNN. **Food Analytical Methods**, v. 18, n. 2, p. 202-217, 2025.

WEDEL, C., ATAMER, Z., DETTLING, A., WENNING, M., SCHERER, S., HINRICHS, J. Towards low-spore milk powders: A review on microbiological challenges of dairy powder production with focus on aerobic mesophilic and thermophilic spores. **International Dairy Journal**, v. 126, p. 105252, 2022.

WESTLAND, J. C. Partial least squares path analysis. **Structural equation models: From paths to networks**, p. 17-38, 2019.

WU, Q., PENG, D., LIU, Q., SHABBIR, M. A. B., SAJID, A., LIU, Z., YUAN, Z. A novel microbiological method in microtiter plates for screening seven kinds of widely used antibiotics residues in milk, chicken egg and honey. **Frontiers in Microbiology**, v. 10, p. 436, 2019.

XIONG, Y., ZHANG, D., YE, C., WANG, Y., DENG, X., DENG, D., HUANG, G. Ultra-sensitive detection of ciprofloxacin hydrochloride in milk by molecularly imprinted electrochemical sensor based on S-CoFe-MOFs/AuNPs. **Journal of Food Composition and Analysis**, v. 122, p. 105439, 2023.

YAN, Y. H., LI, G., LI, G. B. Principles and current strategies targeting metallo- β -lactamase mediated antibacterial resistance. **Medicinal Research Reviews**, v. 40, n. 5, p. 1558-1592, 2020.

ZADE, S. V., NEYMEYR, K., SAWALL, M., KHOSHAYAND, M. R., ABDOLLAHI, H. Identification and quantification of ciprofloxacin, enrofloxacin and difloxacin residues in milk

using excitation-emission fluorescence and second-order standard addition methods. **Chemometrics and Intelligent Laboratory Systems**, v. 242, p. 104988, 2023.

ZAHARI, N. I. N., ENGKU ABD RAHMAN, E. N. S., IREKEOLA, A. A., AHMED, N., RABAAN, A. A., ALOTAIBI, J., YEAN, C. Y. A review of the resistance mechanisms for β -lactams, macrolides and fluoroquinolones among *Streptococcus pneumoniae*. **Medicina**, v. 59, n. 11, p. 1927, 2023.

ZHANG, Y., WU, M., XI, J., PAN, C., XU, Z., XIA, W., ZHANG, W. Multiple-fingerprint analysis of *Poria cocos* polysaccharide by HPLC combined with chemometrics methods. **Journal of Pharmaceutical and Biomedical Analysis**, v. 198, p. 114012, 2021.

ZHANG, Y., YILMAZ, A., POPA, M., BREWSTER, C. Analysis of the impact of data augmentation on the performance of deep learning models in multispectral food authenticity identification. In: **2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS)**. IEEE, 2023. p. 823-832.

ZHOU, J., NIE, W., CHEN, Y., YANG, C., GONG, L., ZHANG, C., FENG, X. Quadruplex gold immunochromatographic assay for four families of antibiotic residues in milk. **Food Chemistry**, v. 256, p. 304-310, 2018.

ZHU, M. Z., WEN, B., WU, H., LI, J., LIN, H., LI, Q., LIU, Z. The Quality Control of Tea by Near-Infrared Reflectance (NIR) Spectroscopy and Chemometrics. **Journal of Spectroscopy**, v. 2019, n. 1, p. 8129648, 2019.

ZIGO, F., VASIL', M., ONDRAŠOVIČOVÁ, S., VÝROSTKOVÁ, J., BUJOK, J., PECKA-KIELB, E. Maintaining optimal mammary gland health and prevention of mastitis. **Frontiers in veterinary science**, v. 8, p. 607311, 2021.

CAPÍTULO II: OBJETIVOS

Objetivo Geral

Aplicar técnicas espectroscópicas (MIR e NIR) associadas a métodos quimiométricos e à geração de amostras sintéticas por PCA, visando desenvolver uma abordagem inovadora, sustentável e eficaz para a detecção, predição e classificação de resíduos de Ciprofloxacina em leite em pó obtido por liofilização.

Objetivos Específicos

- Analisar espectros nas regiões do infravermelho médio (MIR) e próximo (NIR);
- Gerar amostras sintéticas por meio da Análise de Componentes Principais (PCA);
- Aplicar pré-processamentos espectrais e métodos quimiométricos para desenvolver modelos preditivos e classificatório;
- Comparar o desempenho dos modelos de predição e classificação.

CAPÍTULO III: ARTIGO I – USO DE AMOSTRAS SINTÉTICAS NA PREDIÇÃO QUIMIOMÉTRICA DE RESÍDUOS DE CIPROFLOXACINA EM LEITE EM PÓ OBTIDO POR LIOFILIZAÇÃO UTILIZANDO ESPECTROSCOPIAS VIBRACIONAIS

1. Introdução

Os antibióticos são substâncias de origem natural ou sintética amplamente utilizadas no tratamento e na prevenção de infecções em humanos e animais. Na pecuária leiteira, esses fármacos são empregados principalmente no controle de enfermidades como a mastite (inflamação a glândula mamária) de animais lactantes, destacando-se as fluoroquinolonas, entre elas a Ciprofloxacina. Em função de seu amplo espectro de ação contra microrganismos patogênicos, seu uso é comum na pecuária leiteira (QIAO et al., 2018; XIONG et al., 2023). Entretanto, o uso inadequado desses medicamentos, incluindo aplicações profiláticas indiscriminadas ou o descumprimento do período de carência, pode resultar na presença de resíduos em produtos de origem animal, como o leite e seus derivados. Isso representa riscos à saúde pública.

A ingestão de resíduos de antibióticos está associada a diversos efeitos adversos à saúde humana, incluindo reações alérgicas, toxicidade e, sobretudo, o desenvolvimento da resistência antimicrobiana, considerada um dos principais desafios sanitários globais da atualidade (KALUNKE et al., 2018). Diante desse cenário, organismos internacionais, como a Organização das Nações Unidas para a Alimentação e Agricultura (FAO) e a Organização Mundial da Saúde (OMS), estabeleceram diretrizes rigorosas para o controle de resíduos de medicamentos veterinários em alimentos. No Brasil, o Ministério da Agricultura, Pecuária e Abastecimento (Mapa), em conjunto com o Ministério da Saúde (MS), atua na regulamentação e fiscalização desses limites, seguindo organismos internacionais. De acordo com a legislação o limite máximo de resíduo (LMR) para a Ciprofloxacina no leite é de $100 \mu\text{g L}^{-1}$, sendo o produto considerado impróprio para consumo quando esse valor é excedido (UNIÃO EUROPEIA, 2010; BRITO et al., 2024).

A determinação de resíduos de antibióticos em leite é tradicionalmente realizada por técnicas analíticas consolidadas, como a Cromatografia Líquida de Alta Eficiência (CLAE) e a Eletroforese Capilar, que apresentam elevada sensibilidade e seletividade. Contudo, esses métodos requerem preparo extensivo das amostras, utilização de reagentes químicos, elevando o tempo analítico, limitando seu uso rotineiro em larga escala (BRITO et al., 2024). Em função da resposta analítica lenta, o número de amostras analisadas é baixo, o que tem causado críticas ao sistema avaliativo.

Assim, embora a literatura científica seja escassa quanto à avaliação direta da percepção de produtores e laticinistas sobre as perdas econômicas decorrentes da demora nas análises oficiais de resíduos de antibióticos em leite, estudos qualitativos e de percepção fornecem subsídios relevantes para essa discussão. Estudos indicam que os produtores de leite em diferentes contextos produtivos, relatam custos econômicos em função das rotinas operacionais que oneram o custo de produção (BERNIER, et al., 2022; WEMETTE et al., 2020). Nesse sentido, sistemas de controle baseados em análises laboratoriais confirmatórias, especialmente métodos cromatográficos, embora tecnicamente robustos, podem ser percebidos como entraves operacionais quando não oferecem respostas em tempo compatível com a dinâmica da cadeia produtiva.

Dessa forma, é plausível inferir que a demora entre a coleta da amostra e a liberação do resultado analítico oficial intensifique as perdas econômicas percebidas pelos produtores, uma vez que o leite permanece retido ou é descartado. Ainda que os estudos disponíveis não abordem explicitamente o tempo de resposta dos laboratórios oficiais, estes estudos sustentam a argumentação de que atrasos nos processos de avaliação de resíduos podem agravar custos diretos e indiretos, reforçando a percepção negativa sobre os sistemas de controle vigentes e evidenciando uma lacuna relevante na literatura quanto à eficiência dos programas oficiais de monitoramento de resíduos de antibióticos no leite (BERNIER et al., 2022).

Nesse contexto, técnicas espectroscópicas no infravermelho têm se destacado como alternativas rápidas, não destrutivas e ambientalmente sustentáveis. A espectroscopia no infravermelho próximo (NIR) é amplamente empregada em análises rápidas e aplicações industriais; entretanto, suas bandas espectrais estão associadas principalmente a sobretons e combinações vibracionais, o que resulta em menor especificidade química. Ainda assim, diversos estudos relatam seu potencial para a determinação de adulterantes em matrizes complexas (RODRÍGUEZ-HERNÁNDEZ et al., 2025; PEREIRA et al., 2024). A espectroscopia no infravermelho médio (MIR), especialmente quando acoplada à transformada de Fourier (FTIR), fornece bandas fundamentais diretamente relacionadas às vibrações moleculares, permitindo maior sensibilidade e especificidade na detecção e quantificação de compostos químicos, como resíduos de antibióticos, em leite em pó (HU et al., 2021; FREITAS et al., 2021; SANTANA et al., 2025).

No presente estudo, optou-se pela utilização de leite em pó liofilizado em detrimento do leite em pó obtido por *spray drying* (FREITAS et al., 2021; SANTANA et al., 2025; PRIYADARSINI et al., 2025). A liofilização é um processo conduzido sob baixas temperaturas e pressão reduzida, o que minimiza alterações físico-químicas e estruturais da matriz,

preservando compostos termossensíveis (CAIS-SOKOLIŃSKA, TEICHERT e GAWAŁEK, 2023; ZHOU et al., 2025). De acordo com o estudo de Santana et al. (2025), o uso de leite em pó liofilizado proporciona maior controle experimental, reprodutibilidade e confiabilidade na construção de modelos quimiométricos preditivos.

Assim como nos estudos de Santana et al. (2025) e Freitas et al. (2021), as amostras analisadas neste trabalho não foram adulteradas, mas fortificadas com padrão de antibiótico em concentrações conhecidas e controladas. A fortificação é uma prática amplamente empregada em estudos analíticos e quimiométricos, pois permite avaliar de forma sistemática o desempenho, a sensibilidade e a capacidade preditiva dos modelos desenvolvidos (BRITO et al., 2024; FREITAS et al., 2021; ARSAND, 2016).

A associação das espectroscopias MIR e NIR com ferramentas quimiométricas tem se mostrado altamente promissora para o desenvolvimento de modelos preditivos capazes de estimar concentrações de analitos a partir de dados espectrais. Métodos quimiométricos como a Regressão dos Mínimos Quadrados Parciais (PLS) e a Regressão por Componentes Principais (PCR) são amplamente utilizados na construção de modelos de predição, permitindo a extração de informações relevantes de conjuntos de dados espectrais de alta dimensionalidade (FREITAS et al., 2025; CHRISTOPHE et al., 2025; COQUEIRO et al., 2024).

Em análises espectrais no MIR e NIR, caracterizadas por variáveis altamente correlacionadas, a aplicação de pré-processamentos é essencial para a construção de modelos preditivos robustos (RINNAN, VAN DEN BERG e ENGELSEN, 2009; SANTOS et al., 2026). Técnicas como correção de linha de base, variável normal padrão (SNV, *Standard Normal Variate*), derivadas de Savitzky–Golay e correção de espalhamento multiplicativo (MSC, *Multiplicative Scatter Correction*) são empregadas para reduzir ruídos instrumentais e efeitos físicos indesejáveis, além de realçar variações espectrais associadas à concentração do analito, contribuindo para a melhoria do desempenho dos modelos quimiométricos (CANDEIAS et al., 2025; SANTOS et al., 2026).

A robustez dos modelos preditivos também depende da representatividade do conjunto de dados, sendo que a limitação no número de amostras reais pode comprometer seu desempenho (HAO, LI e ZHANG, 2023). Nesse cenário, a geração de amostras sintéticas tem se mostrado estratégia eficiente para expandir conjuntos de dados espectrais e aumentar a variabilidade estatística. Além disso, fatores econômicos e ambientais, como a redução de custos analíticos e a minimização do descarte de amostras, reforçam sua aplicação (DONG et al., 2022; KULKO, HANUS e ELSER, 2025; HELLEQUIN et al., 2025). A escolha do método

para geração de amostras sintéticas é uma etapa crucial para a confiabilidade dos dados, pois depende da natureza do conjunto original, dos objetivos do estudo e das técnicas estatísticas ou quimiométricas empregadas. Entre os métodos mais utilizados destacam-se o SMOTE, os algoritmos baseados em redes adversariais generativas (GANs), a interpolação linear aleatória, a adição de ruído gaussiano e a Análise de Componentes Principais (PCA) (HAO, LI e ZHANG, 2023). Dentre essas abordagens, no presente estudo a PCA se destaca por sua capacidade de preservar a estrutura de variância e as correlações espectrais dos dados originais, sendo especialmente adequada para dados MIR e NIR de alta dimensionalidade (FIGUEIRA e VAZ, 2022; FOLLI et al., 2024).

Diante desse contexto, o objetivo deste estudo foi o de usar as espectroscopias vibracionais MIR e NIR associada a métodos quimiométricos, empregando geração de amostras sintéticas via PCA, para o desenvolvimento de modelos preditivos robustos na determinação de Ciprofloxacina em leite em pó liofilizado.

2. Material e Métodos

2.1. Amostras

Foram coletadas amostras de leite cru de vacas da raça Jersey, que não receberam tratamento com antibióticos nos últimos 60 dias, pertencentes ao rebanho da Embrapa Clima Temperado, localizada no município de Pelotas, Rio Grande do Sul. Um litro de leite foi coletado do tanque de refrigeração e conduzido ao laboratório para posterior análise.

2.2 Delineamento experimental

O padrão analítico do antibiótico Ciprofloxacina foi empregado em sua forma sólida (pó). Primeiramente, preparou-se uma solução estoque de Ciprofloxacina em leite, com concentração de $10.000 \mu\text{g L}^{-1}$, elaborada imediatamente antes do uso. A partir dessa solução, realizaram-se diluições sucessivas até atingir uma concentração correspondente ao dobro do Limite Máximo de Resíduos (LMR) estabelecido para o composto, obtendo-se assim as seguintes concentrações: 25, 50, 75, 100, 125, 150, 175 e $200 \mu\text{g L}^{-1}$. A amostra controle (0), denominada amostra pura, não recebeu adição do antibiótico. A fortificação das amostras foi conduzida durante cinco dias consecutivos para cada nível de concentração, assegurando a homogeneização por agitação manual durante 1 minuto. Em seguida, as amostras foram encaminhadas para liofilização no Laboratório de Análise de Alimentos da Embrapa Clima Temperado e, posteriormente, analisadas no Laboratório do Centro de Estudos em Leite (CEL), vinculado ao Departamento de Tecnologia Rural e Animal.

2.3. Liofilização

As amostras de leite, com e sem adição de padrão, foram transferidas para frascos eppendorf® de 2 mL e congeladas a -20 °C para posterior liofilização à -48 °C $\pm$ 2 °C e 0,100 mBar por 48 horas em liofilizador de bancada L108, LioTop (São Paulo, Brasil).

2.4. Espectroscopia no infravermelho médio com Transformada de Fourier – FTIR- MIR

As amostras de leite liofilizado puras e fortificadas com padrão de Ciprofloxacina foram analisados na região do infravermelho médio de 4000-650 cm^{-1} em equipamento Agilent, modelo Cary 630 (Santa Clara, USA). Os dados espectroscópicos foram obtidos com os softwares Microlab e Resolution Pro, da Agilent (Santa Clara, USA), e usados para posterior caracterização dos espectros. Para a análise, uma pequena quantidade de cada amostra foi depositada diretamente sobre a superfície do cristal de diamante. Em seguida, foram coletados os espectros no modo de absorbância por meio da incidência dos raios infravermelhos, utilizando a técnica de reflectância total atenuada (ATR).

2.5. Espectroscopia no Infravermelho Próximo – NIR

Para as análises por espectroscopia no infravermelho próximo, foram coletados os espectros através de espectrômetro NIR (SpectraStar 2500XL, Unity Scientific, Brookfield, CT, EUA) equipado com uma lâmpada de halogênio de tungstênio como fonte de luz e um detector de arseneto de índio-gálio (InGaAs). Os sinais foram gerados no modo de refletância (%R) e transformados em absorbância usando $\log 1/R$. A faixa de comprimento de onda equivalente a 1100 a 2500 nm, em intervalos de 1 nm. Cada espectro foi obtido como resultado médio de 64 varreduras com 1400 pontos.

2.6. Pré-tratamento dos dados

Os dados obtidos para MIR e NIR, foram pré-processados, para remoção de variações sistemáticas do instrumento e até mesmo de condições ambientais. O pré-tratamento consistiu na correção de linha de base, utilizando o software RStudio versão 2024.12.1 (COQUEIRO et al., 2024).

2.7. Geração das amostras sintéticas

O procedimento adotado seguiu a metodologia proposta por Folli et al. (2024), na qual pequenas perturbações são introduzidas nos *scores* dos componentes principais, permitindo a criação de amostras sintéticas realistas e estatisticamente consistentes com o conjunto de dados original.

Inicialmente, os espectros foram organizados em grupos de acordo com a concentração de Ciprofloxacina (0, 25, 50, 75, 100, 125, 150, 175 e 200 $\mu\text{g L}^{-1}$), de modo a permitir a geração de amostras sintéticas por classe (por exemplo, Classe_200). Cada classe, composta por cinco amostras reais, foi organizada em planilhas eletrônicas no Microsoft Excel e, posteriormente, importada individualmente para o ambiente MATLAB® R2025b (MathWorks, Natick, MA, EUA). A geração das amostras sintéticas foi realizada utilizando a linguagem MATLAB, com o uso de funções nativas para leitura e escrita de dados (*readmatrix* e *writematrix*) e para aplicação da Análise de Componentes Principais (PCA).

A geração de amostras sintéticas foi realizada a partir dos espectros (sem pré-processamento prévio) reais utilizando Análise de Componentes Principais (PCA), com o objetivo de aumentar o número de amostras preservando a estrutura multivariada dos dados espectrais. Os espectros foram organizados em uma matriz $\mathbf{X}$ de dimensão $n \times p$, em que n corresponde ao número de amostras reais e p ao número de variáveis espectrais.

Inicialmente, os dados espectrais foram centrados na média, subtraindo-se, de cada variável, sua respectiva média calculada a partir das amostras reais. Em seguida, a PCA foi aplicada à matriz centrada, sendo retidos os quatro primeiros componentes principais (PCs), número escolhido de modo a ser inferior ao total de amostras reais e suficiente para representar a variabilidade relevante do conjunto de dados. Os *scores* obtidos para cada amostra real nos PCs selecionados foram utilizados como ponto de partida para a geração das amostras sintéticas. Para cada amostra real, foram geradas cinco amostras sintéticas por meio da adição de ruído aleatório aos *scores* originais. O ruído foi modelado como variável aleatória com distribuição normal, escalado por um fator proporcional (10%) ao desvio padrão dos *scores* reais em cada componente principal, garantindo que a variabilidade introduzida permanecesse compatível com a variabilidade experimental típica de dados espectrais do tipo MIR e NIR.

Os novos conjuntos de *scores* sintéticos foram então reconvertidos ao espaço espectral original por meio da reconstrução linear utilizando os pesos (*loadings*) dos componentes principais selecionados. Após a reconstrução, a média espectral previamente subtraída foi adicionada novamente, resultando nos espectros sintéticos finais.

As amostras sintéticas geradas foram posteriormente combinadas às amostras reais, formando um conjunto de dados ampliado, que foi exportado em formato de planilha para uso em etapas posteriores de modelagem quimiométrica. Todas as etapas são descritas no código abaixo (Figura 1).

```

clear; clc; close all;

1. CONFIGURAÇÕES
arquivo = Dados MIR';
n_sint_por_amostra = 5;
nPC = 4; fator_ruido = 0.1; alpha = 0.95;

2. LEITURA DOS DADOS
X = readmatrix(arquivo);
[n_real, n_vars] = size(X);

if nPC >= n_real
    error
end

fprintf('Amostras reais: %d\n', n_real);
fprintf('Variáveis espectrais: %d\n', n_vars);

3. PCA (DADOS REAIS)
mu = mean(X, 1); % média espectral
Xc = X - mu; % centragem (como prcomp(center = TRUE))

[coeff, score, latent] = pca(Xc, 'NumComponents', nPC);

% Desvio padrão real dos scores (como no R)
sd_scores = std(score(:,1:nPC), 0, 1);

4. GERAÇÃO DAS SINTÉTICAS
Xsyn = [];
Tsyn = [];

for i = 1:n_real
    for j = 1:n_sint_por_amostra

        ruído = randn(1,nPC) .* (fator_ruido * sd_scores);
        novo_score = score(i,1:nPC) + ruído;

        % Reconstrução espectral
        x_rec = novo_score * coeff(:,1:nPC);
        x_final = x_rec + mu;

        Xsyn = [Xsyn; x_final];
        Tsyn = [Tsyn; novo_score];
    end
end

fprintf('Amostras sintéticas geradas: %d\n', size(Xsyn,1));

5. CONJUNTO FINAL
X_final = [X; Xsyn];

6. SALVAR EM EXCEL
arquivo_saida = 'sinteMIR_concentraçãoX.xlsx';
writematrix(X_final, arquivo_saida);

disp(['✓ Arquivo salvo com sucesso: ', arquivo_saida]);

```

Figura 1. Código para geração das amostras sintéticas por espectros MIR e NIR.

A distribuição das amostras foi analisada por meio do gráfico de *scores* $PC1 \times PC2$, no qual também foi construída uma elipse de confiança de 95%, baseada na estatística T^2 de Hotelling calculada a partir das amostras reais, permitindo verificar se as amostras sintéticas permaneciam dentro da região de variabilidade multivariada dos dados experimentais.

2.8. Análises quimiométricas

A organização dos dados para análises foi realizada em planilhas do Excel. As análises estatísticas multivariadas foram realizadas no *software* MATLAB® R2025b (MathWorks, Natick, MA, EUA) e foram realizadas para o conjunto de dados reais e para o novo conjunto de dados (reais combinadas com sintéticas).

Antes da aplicação dos métodos quimiométricos, os dados espectrais foram submetidos a diferentes estratégias de pré-processamento, com o objetivo de reduzir efeitos físicos não relacionados à composição química e melhorar a qualidade da informação espectral. Foram avaliados quatro pré-processamentos distintos: (i) *Standard Normal Variate*-normalização da variável padrão (SNV), utilizado para corrigir variações de espalhamento e diferenças de caminho óptico; (ii) primeira derivada de Savitzky–Golay, com polinômios de segunda ordem, utilizando uma janela de 15 pontos aplicada para realçar bandas espectrais e minimizar efeitos de linha de base; (iii) *Multiplicative Scatter Correction* - correção multiplicativa de dispersão

(MSC), empregada para compensar variações multiplicativas e aditivas associadas ao espalhamento da radiação; e (iv) correção de linha de base, utilizada para remover deslocamentos sistemáticos do sinal espectral. A comparação entre os diferentes pré-processamentos permitiu avaliar sua influência no desempenho das análises multivariadas subsequentes (CANDEIAS et al., 2025).

2.8.1. Regressão de mínimos quadrados parciais - PLS

A regressão por mínimos quadrados parciais (PLS) foi empregada para avaliar a capacidade do modelo em prever as concentrações de Ciprofloxacina em amostras de leite liofilizado, utilizando os espectros completos nas regiões MIR e NIR. A metodologia desta análise foi baseada nos estudos de Freitas et al. (2025) e Coqueiro et al. (2024).

Os dados foram divididos em conjuntos de calibração (60%) e validação externa (40%) por meio do algoritmo de amostragem de Kennard–Stone (Folli et al., 2024), assegurando a distribuição representativa do espaço espectral. As variáveis independentes (**X**), correspondentes aos comprimentos de onda, foram utilizadas para modelar as variáveis dependentes (**y**), representadas pelas concentrações de Ciprofloxacina.

O número ótimo de variáveis latentes do modelo PLS foi definido por validação cruzada, com base no menor valor da raiz do erro médio quadrático de validação. Após o ajuste do modelo, a presença de possíveis *outliers* foi avaliada no conjunto de calibração por meio da análise dos resíduos de predição e da distância de Hotelling (T^2), com nível de confiança de 95%. A distância de Hotelling foi calculada conforme a Equação (1):

$$T^2 = t^T \cdot S^{-1} \cdot t$$

Equação 1

em que **t** representa o vetor de *scores* da amostra e **S** a matriz de covariância.

Para investigar a influência dessas amostras no desempenho do modelo, análises PLS foram realizadas com e sem a exclusão dos *outliers* identificados. Os modelos foram comparados a partir das métricas: erro médio quadrático de predição (RMSEP), a correlação (*r*) e viés, visando avaliar a robustez preditiva e a presença de erro sistemático (COQUEIRO et al., 2024; LLETÍ et al., 2005).

A fórmula do RMSEP é dada por:

$$RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Equação 2

Em que:

y_i = valor experimental (real) da amostra i ;

$\hat{y}_i$ = valor predito pelo modelo para a amostra i ;

n = número total de amostras no conjunto de predição.

O coeficiente de correlação (r) foi determinado conforme a Equação (3):

$$r = \frac{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{\hat{y}})^2}{\sqrt{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{\hat{y}})^2}}$$

Equação 3

Em que:

r = coeficiente de correlação entre os valores experimentais e os valores preditos pelo modelo;

y_i = valor experimental (real) da amostra i ;

$\hat{y}_i$ = valor predito pelo modelo para a amostra i ;

$\bar{y}$ = média dos valores experimentais;

$\bar{\hat{y}}$ = média dos valores preditos pelo modelo;

n = número total de amostras avaliadas.

O viés foi determinado conforme a Equação (4):

$$Viés = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$$

Equação 4

Em que:

Viés = indica a presença de erro sistemático nas predições;

y_i = valor experimental (real) da amostra i ;

$\hat{y}_i$ = valor predito pelo modelo para a amostra i ;

n = número total de amostras no conjunto de predição.

O desempenho dos modelos PLS foi avaliado por meio do coeficiente de correlação (r), da raiz do erro médio quadrático de calibração (RMSEC) e de predição (RMSEP). A aplicabilidade analítica dos modelos foi avaliada pelas razões RPD (Razão de Desempenho em Relação ao Desvio) e RER (Razão de Erro Relativo).

O RMSEC foi determinado pela Equação (5):

$$RMSEC = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Equação 5

Em que:

y_i = valor experimental (real) da amostra i ;

$\hat{y}_i$ = valor predito pelo modelo para a amostra i ;

n = número total de amostras no conjunto de calibração.

O RPD foi determinado pela Equação (6):

$$RPD = \frac{DP}{RMSE}$$

Equação 6

Em que:

DP = desvio padrão dos valores experimentais do conjunto de referência;

RMSE = raiz do erro médio quadrático de predição ou calibração.

Por fim, o RER foi determinado pela Equação (7):

$$RER = \frac{\text{Valor máximo real} - \text{Valor mínimo real}}{RMSE}$$

Equação 7

Em que:

Valor máximo e mínimo = intervalo dos valores de referência do conjunto de dados;

RMSE = raiz do erro médio quadrático de predição ou calibração.

2.8.2 Regressão por Componentes Principais-PCR

Os espectros MIR e NIR foram organizados em uma matriz (**X**), contendo as variáveis espectrais, e um vetor resposta (**y**), correspondente às concentrações do analito. Os dados foram divididos em conjuntos de calibração (60%) e validação externa (40%) por meio do algoritmo de amostragem de Kennard–Stone (FOLLI et al., 2024), assegurando distribuição representativa do espaço espectral. A metodologia desta análise foi baseada no estudo de Mouazen et al. (2010).

O modelo de Regressão por Componentes Principais (PCR) foi construído a partir da decomposição em valores singulares (SVD) da matriz (**X**) calibrada. Foram avaliados modelos contendo até 15 componentes principais, sendo o número ótimo selecionado com base no menor erro quadrático médio de previsão (RMSEP). O desempenho do modelo foi avaliado com as mesmas métricas utilizadas na seção anterior.

3. Resultados e discussão

3.1. Geração das amostras sintéticas

As amostras sintéticas foram geradas com base em 45 amostras reais, utilizando análise de componentes principais (PCA) seguida de reconstrução. Para cada amostra real, foram produzidas cinco amostras sintéticas, totalizando 225 amostras sintéticas. A distribuição das amostras sintéticas geradas via PCA para os conjuntos espectrais MIR (A) e NIR (B), permitem avaliar a fidelidade da geração de dados em relação às amostras reais em diferentes concentrações (Figura 2).

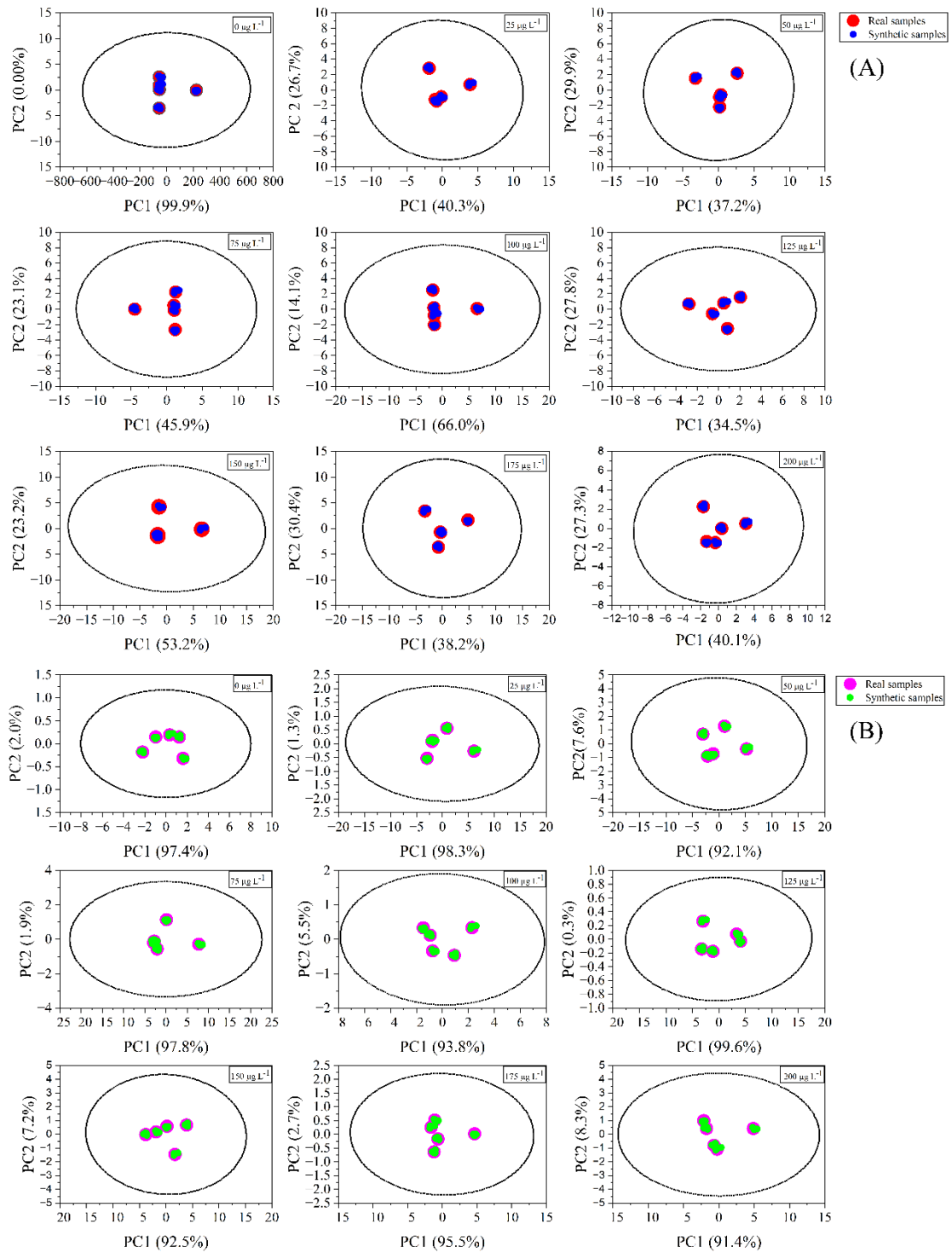


Figura 2. Geração de amostras sintéticas de leite de vacas da raça Jersey a partir de espectros no infravermelho médio (A) e no infravermelho próximo (B), projetadas no espaço de componentes principais (PCA).

Em ambos os casos, observa-se que os pontos correspondentes às amostras sintéticas (azuis no MIR e verdes no NIR) se posicionam próximos às amostras reais, indicando que o modelo PCA foi eficiente em preservar a estrutura multivariada original dos espectros. As

elipses de confiança (95%) delimitam a variabilidade natural dos dados reais dentro de cada classe de concentração. O fato de as amostras sintéticas se encontrarem majoritariamente dentro dessas elipses demonstra que a variabilidade introduzida pelo modelo foi controlada e representativa, sem gerar amostras fora do espaço químico experimental (FOLLI et al., 2024).

É importante destacar que o espaço PCA do MIR apresenta maior separabilidade entre as classes e elipses ligeiramente mais amplas, refletindo maior sensibilidade às variações de concentração do padrão de Ciprofloxacina. No NIR, o agrupamento é mais compacto e linear ao longo do PC1, evidenciando comportamento mais colinear, mas ainda adequado para a geração de amostras sintéticas representativas.

O resultado alcançado demonstra que a geração das amostras sintéticas foi eficaz e adequada, aumentando a base de dados sem introduzir variações artificiais, podendo favorecer o desempenho de modelos de predição multivariada (HEINE et al., 2023; HELLEQUIN et al., 2025).

3.2. Características espectrais obtidas por MIR

3.2.1. Leite liofilizado puro

Os espectros no FTIR (4000 a 650 cm^{-1}) do leite seguem o padrão esperado, com as faixas bem determinadas e de acordo com a literatura científica (Figura 3) (DINÇ et al., 2023; SANTOS et al., 2026).

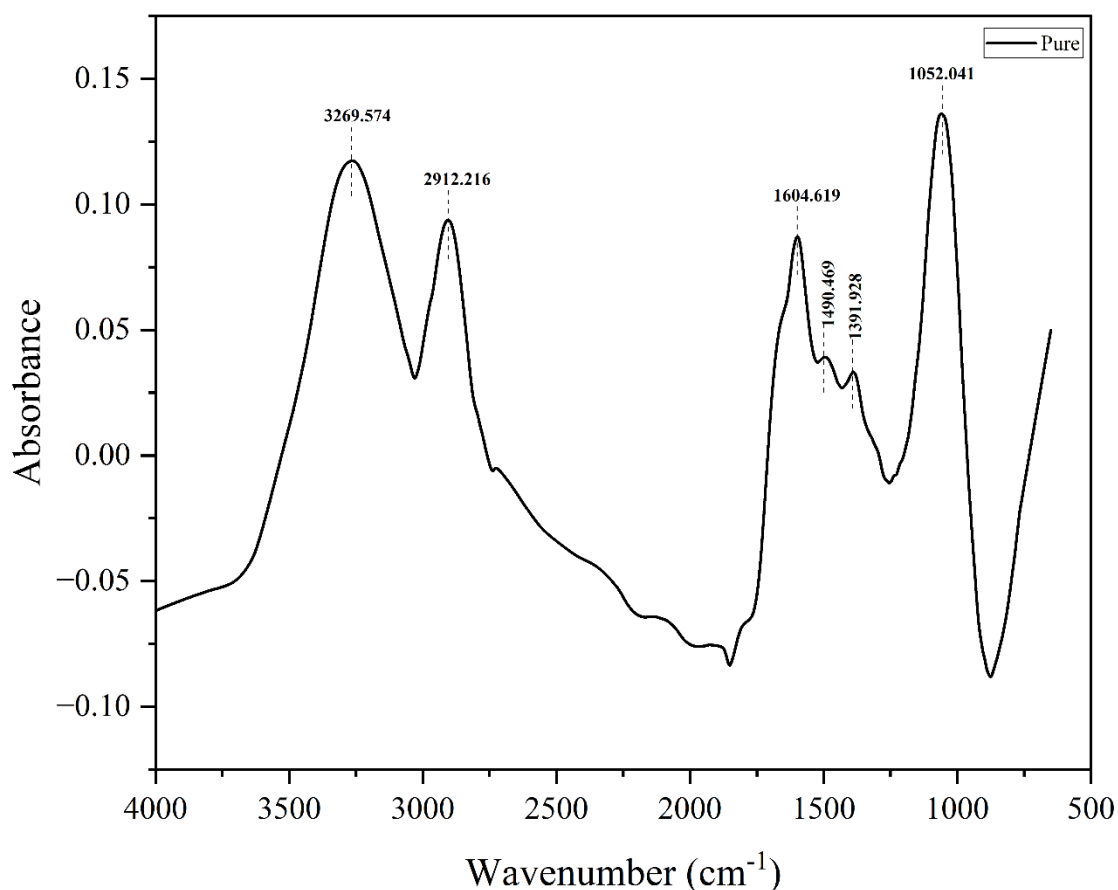


Figura 3. Espectro médio no infravermelho médio de amostras de leite liofilizado puro de vacas da raça Jersey.

O espectro médio das amostras de leite liofilizado puro, obtido por FTIR apresenta picos característicos com diferentes intensidades de absorção, o que possibilita associação destes picos com os grupos funcionais presentes nas amostras. É possível observar, na região de 3500 a 3000 cm^{-1} vibrações associadas ao grupo OH. A forte ligação de hidrogênio entre as moléculas de água favorece a formação de banda larga no pico 3270 cm^{-1} (SANTOS et al., 2026).

O segundo pico marcante, em 2912 cm^{-1} , está associado ao estiramento C–H das cadeias alifáticas dos lipídios e proteínas, principalmente das metilas e metilenos. Essa banda reflete a presença de cadeias longas de ácidos graxos, sendo característica da fração lipídica do leite (DINÇ et al., 2023).

Na região de 1600–1500 cm^{-1} , destacam-se as bandas em 1604 e 1490 cm^{-1} , atribuídas às vibrações Amida primária (C=O) e Amida secundária (N–H e C–N), respectivamente. Essas bandas são marcadores estruturais importantes das ligações peptídicas das proteínas, principalmente da caseína. A presença dessas bandas confirma a integridade das proteínas após

o processo de liofilização, sugerindo que a liofilização não provocou desnaturação significativa (PEREIRA et al., 2020).

Em 1391 cm^{-1} , observa-se banda menos intensa, atribuída à vibração de deformação C–H provenientes de aminoácidos ou ácidos orgânicos. Essa região também pode refletir pequenas contribuições da lactose e do fosfato de cálcio, comum no leite. Por fim, o pico em 1052 cm^{-1} corresponde ao estiramento C–O característico dos açúcares, banda típica de carboidratos e indica a presença expressiva de lactose no leite liofilizado (PEREIRA et al., 2020).

O espectro MIR do leite liofilizado exhibe as bandas vibracionais esperadas para uma matriz rica em proteínas, lipídios e carboidratos, confirmando que a estrutura química fundamental do leite foi preservada após a liofilização. A ausência de deslocamentos ou perdas significativas nas bandas principais indica que o processo de secagem não alterou substancialmente as interações moleculares, mantendo a organização físico-química característica da amostra original.

3.2.2. Padrão de Ciprofloxacina

O estudo do espectro no infravermelho médio do padrão de Ciprofloxacina, abrangendo a faixa de 4000 a 650 cm^{-1} (Figura 4). Este perfil espectral mostra múltiplas bandas bem definidas, atribuídas às vibrações características dos grupos funcionais presentes na estrutura da molécula, composta por anéis aromáticos, grupo carboxílico, função cetona, amina secundária e o grupo piperazínico (KHWAZA, MLALA e ADERIBIGBE, 2024).

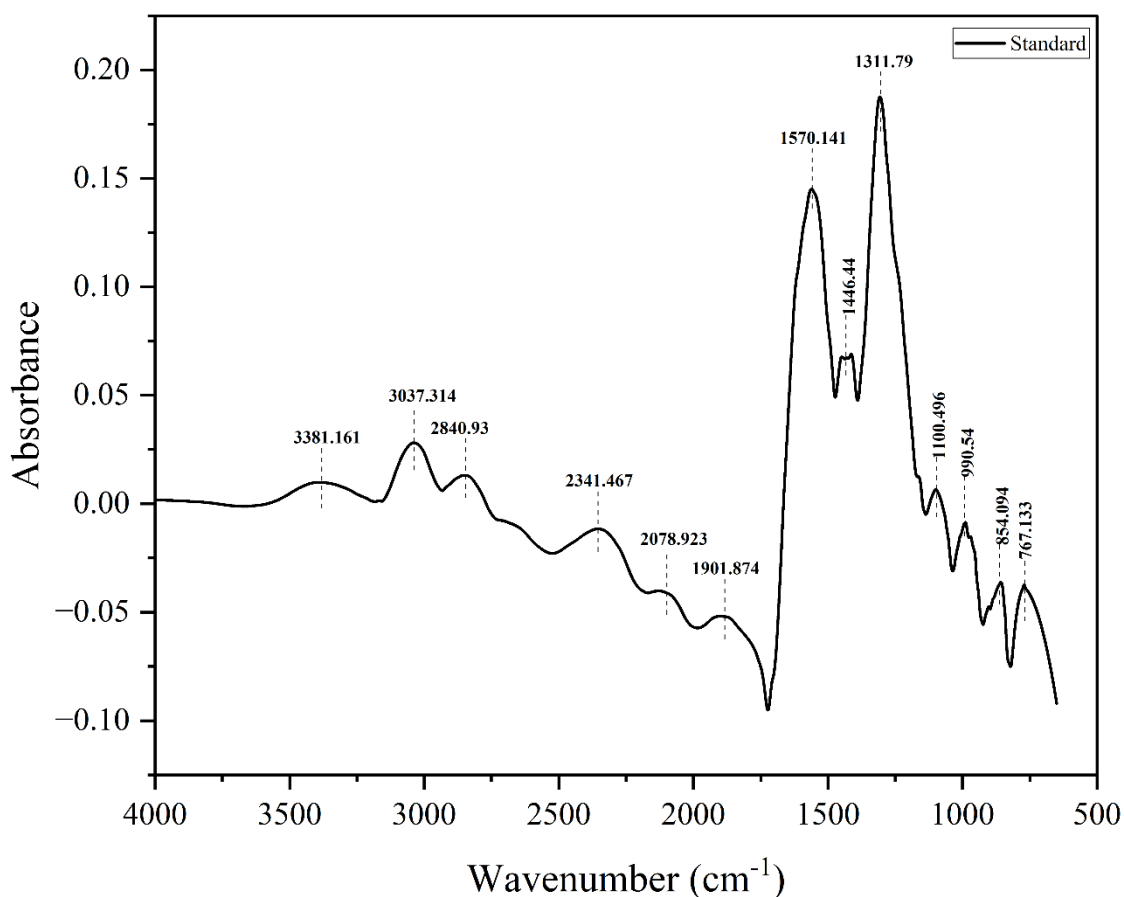


Figura 4. Espectro no infravermelho médio do padrão de Ciprofloxacina.

A banda larga observada na região de 3500–3200 cm^{-1} é atribuída ao estiramento O–H do grupo carboxílico (–COOH) e às vibrações de N–H do grupo amina da piperazina. Na região de 3100–3000 cm^{-1} , podem ser observadas bandas menos intensas associadas aos estiramentos C–H aromáticos, provenientes do anel quinolônico da Ciprofloxacina (TAN et al., 2012). Já as bandas entre 2950 e 2840 cm^{-1} correspondem aos estiramentos C–H alifáticos da cadeia piperazínica (TUREL et al., 1997; SAHOO et al., 2011; RAJALAKSHMI, GUNASEKARAN, e KUMARESAN, 2014; NITHYA E SUNDARAM, 2015).

Entre 1570 e 1440 cm^{-1} , são observadas bandas relacionadas às vibrações C=C e C–N aromáticas, representando as oscilações do anel benzênico e do núcleo heterocíclico da estrutura. Essas bandas confirmam a presença do sistema aromático estendido característico da Ciprofloxacina (SANTOS et al., 2026).

Na região de 1380–1300 cm^{-1} , aparecem bandas atribuídas às deformações C–H e às vibrações C–N da piperazina, indicando a integridade do anel saturado. Por fim, entre 1100 e 1000 cm^{-1} , observam-se bandas de estiramento C–F, C–N e C–O–C, também atribuídas à piperazina e a contribuições do anel quinolônico (SAHOO et al., 2011; SANTOS et al., 2026).

A análise espectroscópica vibracional MIR do padrão de Ciprofloxacina apresenta todas as bandas esperadas para uma fluoroquinolona pura, confirmando sua identidade estrutural e integridade química. As bandas de C=O (carboxila e cetona), N–H, C–F e C–N da piperazina são marcadores vibracionais fundamentais e podem ser utilizadas para monitorar a presença do padrão em matrizes complexas (SANTOS et al., 2026).

3.2.3. Amostras originais fortificadas e sintéticas

A Figura 5 apresenta os espectros médios MIR do leite liofilizado fortificado com diferentes concentrações do padrão de Ciprofloxacina (A) e das amostras sintéticas correspondentes geradas via PCA (B), juntamente com os espectros do leite puro e do padrão do antibiótico.

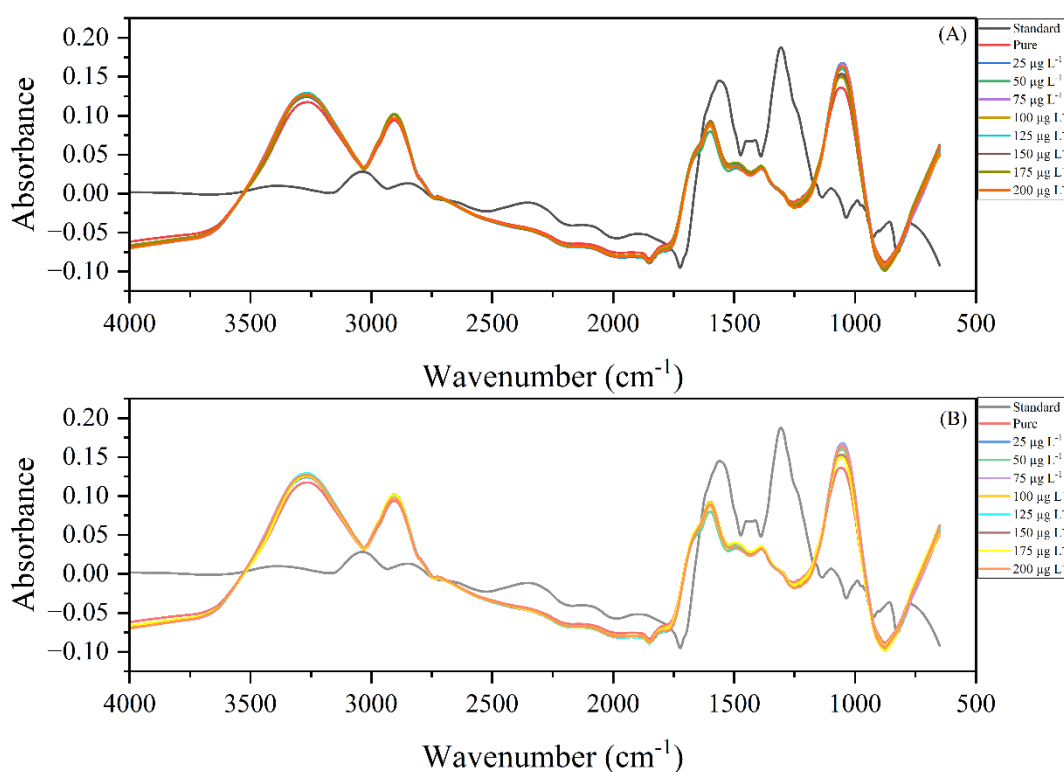


Figura 5. Espectros médios no infravermelho médio de amostras de leite liofilizado puro e fortificadas com diferentes concentrações do padrão de Ciprofloxacina, considerando dados originais (A) e sintéticos (B).

No espectro (A), observa-se o comportamento característico das amostras analisadas, composto pelos perfis espectrais do leite liofilizado puro e do padrão de Ciprofloxacina, bem como pelas variações correspondentes às amostras fortificadas. O leite puro apresenta as bandas típicas de uma matriz rica em proteínas, lipídios e carboidratos, enquanto o padrão do antibiótico exhibe sinais vibracionais mais definidos, atribuídos principalmente aos estiramentos

N–H e O–H ($\approx 3380\text{ cm}^{-1}$), C=O ($\approx 1570\text{--}1600\text{ cm}^{-1}$) e C–F/C–N ($\approx 1100\text{--}1000\text{ cm}^{-1}$) discutidos anteriormente (SAHOO et al., 2011; TAN et al., 2012; SANTOS et al., 2026).

As amostras fortificadas, por sua vez, revelam espectros intermediários, nos quais o perfil do leite permanece predominante, porém com alterações sutis e sistemáticas à medida que aumenta a concentração do padrão de Ciprofloxacina. As principais modificações ocorrem nas regiões de $1600\text{--}1500\text{ cm}^{-1}$ e $1100\text{--}1000\text{ cm}^{-1}$, que apresentam intensificação gradual das bandas e pequenas mudanças na forma dos picos. Tais variações não indicam necessariamente as interações químicas entre o padrão e os constituintes do leite, mas refletem efeitos físicos de adição e sobreposição espectral, resultantes da combinação das absorções características do padrão de Ciprofloxacina com as vibrações naturais da matriz. Na faixa de $1600\text{--}1500\text{ cm}^{-1}$, observam-se contribuições associadas à flexão dos grupos N–H e O–H e ao estiramento do grupo C=O do padrão do antibiótico, os quais se sobrepõem às vibrações da amida I e II das proteínas do leite. Já na região de $1100\text{--}1000\text{ cm}^{-1}$, identificam-se modificações e sobreposições entre as vibrações C–O e C–F do padrão e as absorções C–O dos carboidratos do leite (DINÇ et al., 2023; PRZYBYLSKA et al., 2021; PEREIRA et al., 2020).

Os espectros sintéticos gerados por PCA reproduzem adequadamente o comportamento observado experimentalmente, refletindo a tendência gradual de modificação das bandas com o aumento da concentração e preservando as regiões espectrais mais sensíveis à presença do padrão de Ciprofloxacina (Figura 5B). A geração sintética, baseada nas componentes principais dos dados experimentais, permite avaliar a robustez e a variabilidade do modelo, além de demonstrar que as diferenças espectrais identificadas no conjunto original são reais (FOLLI et al., 2024).

A similaridade entre os espectros reais e sintéticos confirma a consistência do modelo PCA na representação das variações associadas à fortificação, reforçando o potencial da abordagem multivariada para identificar o padrão do antibiótico em matrizes complexas como o leite. Dessa forma, os resultados apresentados na Figura 5B validam o comportamento espectral observado em Figura 5A, evidenciando que o método de geração sintética é eficaz para simular a distribuição das amostras no espaço multivariado (FOLLI et al., 2024).

3.3. Características espectrais obtidas por NIR

3.3.1. Leite liofilizado puro

O espectro médio das cinco amostras de leite liofilizado puro apresenta picos característicos com diferentes intensidades de absorção, o que possibilita associação destes

picos com os grupos funcionais presentes nas amostras de leite liofilizado (Figura 6). O espectro de infravermelho próximo em comprimentos de onda varia de 1100 a 2500 nm. O leite é uma matriz complexa composta principalmente por água, lipídios, proteínas, lactose e minerais. No NIR, a absorção é dominada por combinações harmônicas e sobretons dos grupos O–H, C–H e N–H (ZHANG et al., 2022).

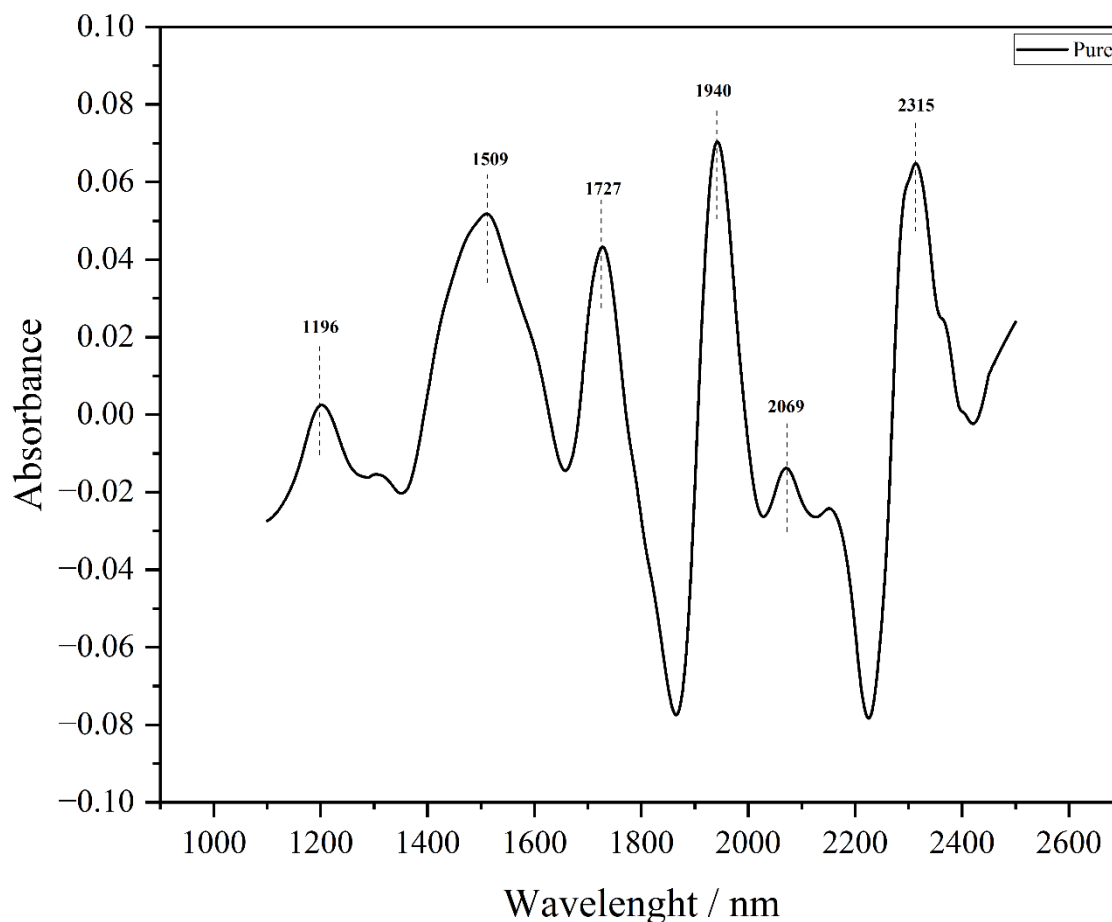


Figura 6. Espectro médio no infravermelho próximo de amostras de leite liofilizado puro de vacas da raça Jersey.

A banda em 1196 nm está relacionada a sobretons de ligações C–H e O–H de ácidos graxos e açúcares, enquanto o sinal em 1509 nm reflete o primeiro overtone de O–H e N–H, atribuído à presença de lactose e proteínas. A absorção em 1727 nm está associada ao primeiro overtone de C–H, indicativo da fração lipídica, e a intensa banda em 1940 nm corresponde a combinações de O–H de água ligada e grupos hidroxila de carboidratos. Em 2069 nm, observa-se contribuição de bandas de N–H e C=O relacionadas às ligações amídicas das proteínas, enquanto a região de 2315 nm resulta de combinações C–H e C–O, associadas principalmente a lipídios e carboidratos (DAOUD et al., 2020; BEĆ, GRABSKA e HUCK, 2020).

3.3.2. Padrão de Ciprofloxacina

O perfil espectral do padrão da Ciprofloxacina no NIR apresenta bandas características, bem definidas e com variações sistemáticas de absorbância (Figura 7). Estas variações são atribuídas principalmente a sobretons e bandas de combinação de vibrações fundamentais, envolvendo ligações do tipo C–H, O–H e N–H, típicas da espectroscopia no infravermelho próximo (BELLON-MAUREL et al., 2011).

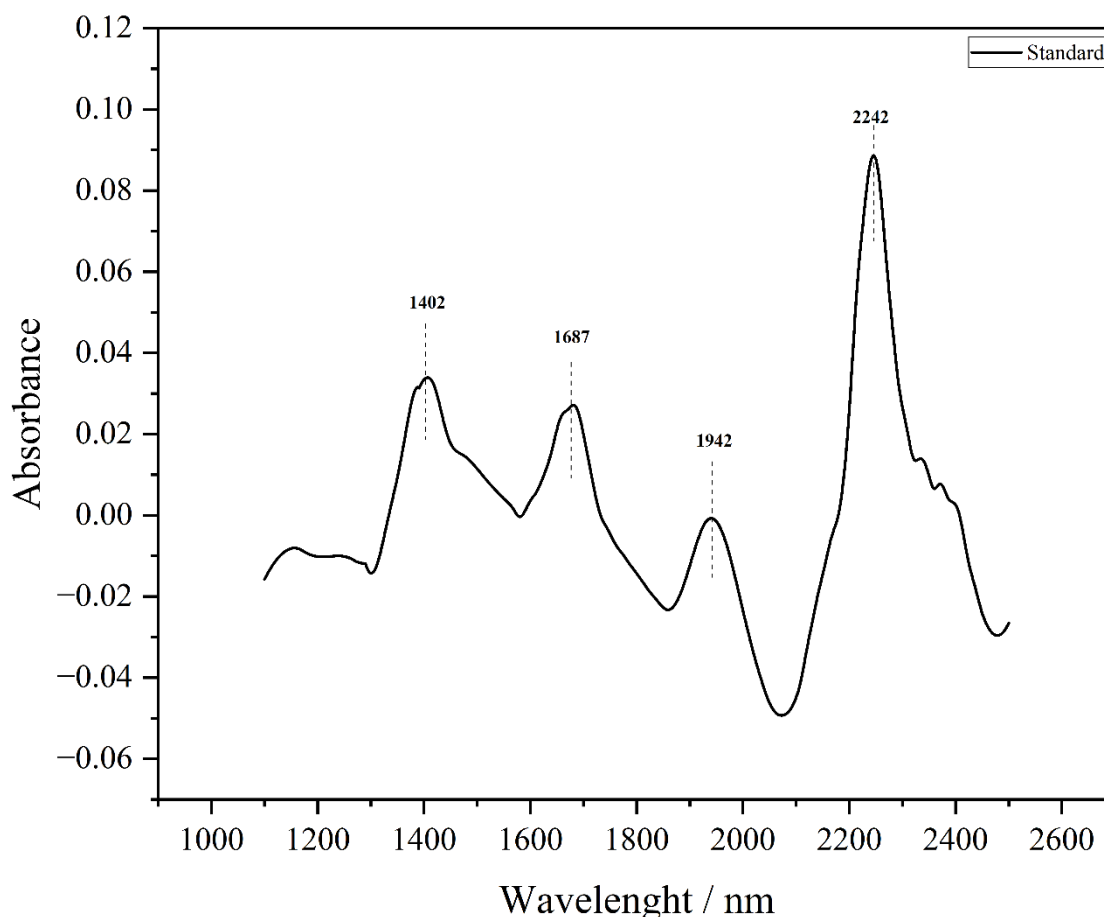


Figura 7. Espectro no infravermelho próximo do padrão de Ciprofloxacina.

Dentre as bandas observadas, destaca-se a absorção em 1402 nm, associada predominantemente ao primeiro sobreton do estiramento O–H, com possível contribuição do estiramento N–H (FATEMI, A., SINGH, V., KAMRUZZAMAN, M., 2022). Essa banda está relacionada à presença do grupo ácido carboxílico (–COOH) e das aminas presentes no anel piperazínico da Ciprofloxacina (TAN et al., 2012).

A banda observada em 1687 nm é atribuída ao primeiro sobreton do estiramento C–H, envolvendo tanto grupos alifáticos quanto aromáticos (QIU et al., 2018; SCHUMER et al., 2025). Na estrutura da Ciprofloxacina, essa absorção está relacionada às ligações C–H do anel

quinolônico e do anel piperazínico, além dos grupos metilênicos da molécula (TAN et al., 2012).

Em 1942 nm, observa-se uma banda característica de combinação envolvendo estiramento e deformação angular O–H (SCHUMER et al., 2025). No contexto do padrão de Ciprofloxacina, essa absorção pode estar associada ao grupo carboxílico (SANTOS et al., 2026).

A absorção mais intensa do espectro ocorre em 2242 nm, sendo atribuída a bandas de combinação envolvendo estiramentos e deformações C–H, com possíveis contribuições de combinações envolvendo C=O e N–H (QIU et al., 2018). A elevada intensidade dessa banda indica forte contribuição da parte orgânica da Ciprofloxacina, especialmente dos anéis aromáticos e das cadeias alifáticas (SANTOS et al., 2026).

É importante destacar que, apesar da Ciprofloxacina apresentar em sua estrutura química o grupo C–F, característico das fluoroquinolonas e considerado essencial para sua atividade farmacológica, não foi observada banda específica atribuída a esse grupo no espectro NIR. Essa ausência é esperada e está relacionada às limitações intrínsecas da espectroscopia no infravermelho próximo, cuja região espectral é dominada por sobretons e bandas de combinação de vibrações fundamentais envolvendo ligações do tipo X–H (X = C, N, O) (SCHUMER et al., 2025). A ligação C–F apresenta elevada energia vibracional e baixa polarizabilidade, resultando em modos fundamentais que absorvem predominantemente na região do infravermelho médio (tipicamente abaixo de 1300 cm^{-1}), sem gerar sobretons ou combinações com intensidade suficiente para serem detectados no NIR (SAHOO et al., 2011).

De modo geral, os resultados demonstram que o espectro NIR do padrão de Ciprofloxacina apresenta bandas coerentes com os principais grupos funcionais da molécula. As regiões espectrais destacadas fornecem informações químicas relevantes e indicam potencial para aplicação em métodos quimiométricos voltados à detecção e quantificação de Ciprofloxacina em matrizes complexas.

3.3.3. Amostras originais fortificadas e sintéticas

São apresentados os espectros médios no infravermelho próximo do padrão de Ciprofloxacina, do leite liofilizado puro e das amostras de leite liofilizado fortificadas com diferentes concentrações de Ciprofloxacina, obtidos a partir de dados originais (8A) e sintéticos (8B), na faixa espectral aproximada de 1100 a 2500 nm. De modo geral, observa-se que os perfis espectrais são predominantemente influenciados pela matriz láctea, com variações

sistemáticas de absorvência associadas ao aumento da concentração do analito (Figura 8).

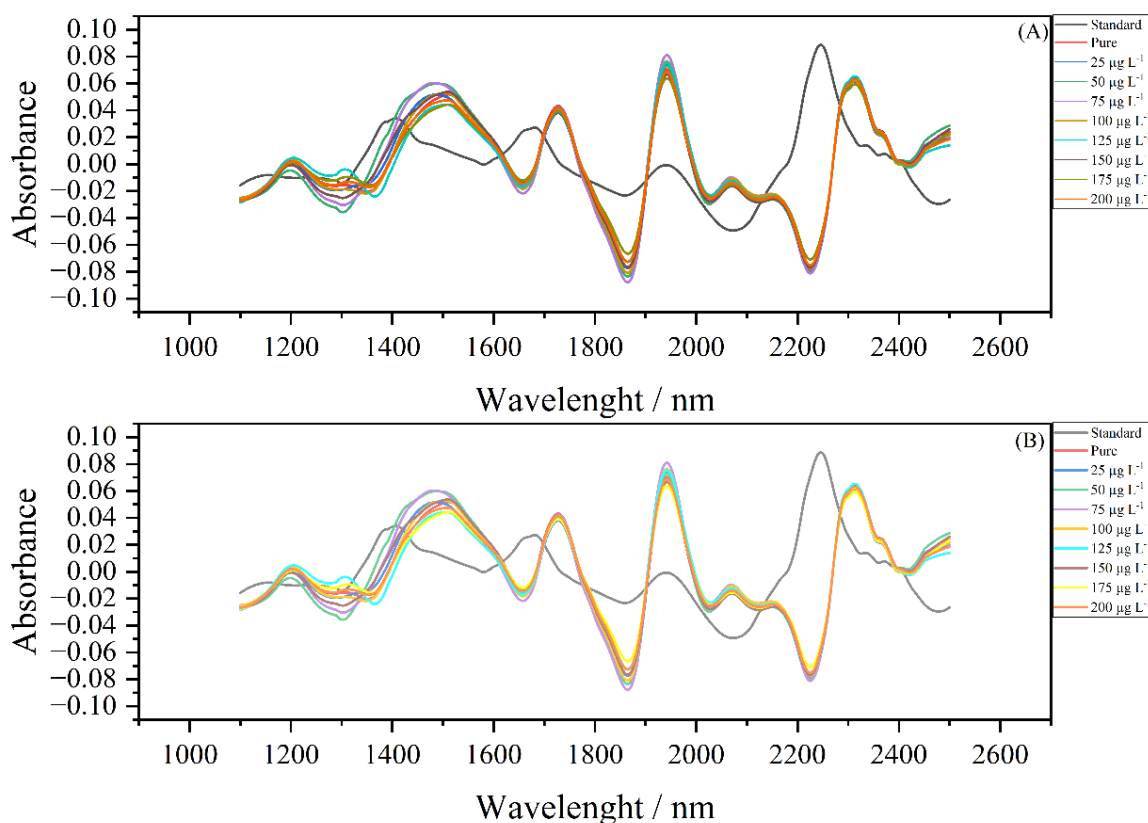


Figura 8. Espectro médio no infravermelho próximo das amostras de leite liofilizado de vacas da raça Jersey fortificadas com diferentes concentrações do padrão de Ciprofloxacina, considerando dados originais (A) e sintéticos (B).

Ao comparar os espectros das amostras fortificadas com Ciprofloxacina com o leite liofilizado puro (Figura 8A). Observa-se que o aumento da concentração do analito não resultou no aparecimento de novas bandas espectrais bem definidas. Podemos observar sim variações graduais de intensidade e forma das bandas já existentes, especialmente nas regiões associadas a estiramentos C–H, N–H e O–H (SCHUMER et al., 2025).

Os espectros do padrão de Ciprofloxacina apresentam bandas características nas regiões de 1400–1450 nm, 1680–1700 nm, 1940 nm e 2200–2250 nm, associadas principalmente aos grupos O–H e N–H (ácido carboxílico e aminas), bem como aos estiramentos C–H do anel quinolônico e do anel piperazínico (SCHUMER et al., 2025; SANTOS et al., 2026). Contudo, devido à elevada complexidade da matriz láctea e à sobreposição espectral típica da região do NIR, essas contribuições do analito manifestam-se nos espectros das amostras fortificadas, devido à alterações sutis no perfil espectral global (DAOUD et al., 2020; BEĆ, GRABSKA e HUCK, 2020).

A comparação entre os espectros originais (Figura 8A) e sintéticos (Figura 8B) revela que os dados sintéticos gerados por PCA, preservam adequadamente as principais características espectrais das amostras reais. A posição das bandas e o comportamento sistemático associado ao aumento da concentração de Ciprofloxacina foram mantidos. As diferenças entre as concentrações foram consistentes em ambos os conjuntos de dados, indicando que o procedimento de geração de amostras sintéticas por PCA reproduziu de forma realista a variabilidade espectral associada ao analito, sem introduzir distorções artificiais (Folli et al., 2024).

3.4. Análise multivariadas

3.4.1. Regressão de Mínimos Quadrados Parciais – PLS

A Regressão de Mínimos Quadrados Parciais (PLS) é um modelo direcionado, e foi utilizado para prever as diferentes concentrações de Ciprofloxacina nas amostras de leite liofilizado, reais e nos dados combinados (sintéticas e reais), com o intuito de reduzir a dispersão entre os dados observados e previstos.

São apresentados os desempenhos dos modelos de regressão por mínimos quadrados parciais (PLS) desenvolvidos a partir de espectros MIR e NIR, considerando diferentes estratégias de pré-processamento espectral, a composição do conjunto de dados (amostras originais e a combinação de dados sintéticos com reais) e o número de variáveis latentes (Tabela 1).

São reportadas as métricas de desempenho em calibração e validação, incluindo coeficiente de correlação (r), erros de calibração e predição (RMSEC e RMSEP), bem como indicadores de capacidade preditiva (RPD e RER) (LIMA et al., 2020; COQUEIRO et al., 2024; FREITAS et al., 2025). Estes dados permitem comparação direta entre os modelos quanto à robustez, generalização e complexidade, fornecendo informações para a seleção do modelo mais adequado para aplicações analíticas (Tabela1).

Tabela 1. Parâmetros de calibração dos modelos de regressão por mínimos quadrados parciais (PLS) obtidos a partir de espectros no infravermelho médio (MIR) e no infravermelho próximo (NIR) para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado de vacas da raça Jersey.

Pré-processamento	Modelo	Dados	Calibração				Validação				
			LV's	r	RMSEC ($\mu\text{g L}^{-1}$)	RPD	RER	r	RMSEP ($\mu\text{g L}^{-1}$)	RPD	RER
1ªDerivada	MIR	Reais	5	0,971	15,464	4,254	12,933	0,104	68,416	0,971	2,923
		Combinados	5	0,963	17,180	3,769	11,641	0,962	17,609	3,683	11,358
	NIR	Reais	5	0,627	50,231	1,310	3,982	0,207	91,657	0,725	2,182
		Combinados	15	0,880	30,577	2,118	6,541	0,874	31,403	2,065	6,369
SNV	MIR	Reais	5	0,998	2,682	24,526	74,570	0,031	66,533	0,998	3,006
		Combinados	5	0,981	12,405	5,220	16,123	0,975	14,249	4,551	14,036
	NIR	Reais	8	0,804	38,375	1,714	5,212	0,089	90,531	0,734	2,209
		Combinados	15	0,974	14,498	4,466	13,795	0,970	15,541	4,173	12,869
Correção de baseline	MIR	Reais	5	0,998	2,855	23,037	70,044	0,063	65,422	1,015	3,057
		Combinados	5	0,977	13,536	4,784	14,776	0,970	15,812	4,101	12,649
	NIR	Reais	8	0,867	32,173	2,045	6,216	0,031	71,554	0,928	2,795
		Combinados	15	0,985	10,938	5,920	18,284	0,982	12,074	5,371	16,564
MSC	MIR	Reais	5	0,997	4,353	15,112	45,948	0,063	66,566	0,998	3,005

Pré-processamento	Modelo	Dados	Calibração				Validação				
			LV's	r	RMSEC (µg L ⁻¹)	RPD	RER	r	RMSEP (µg L ⁻¹)	RPD	RER
		Combinados	5	0,970	15,558	4,162	12,855	0,966	16,587	3,910	12,058
	NIR	Reais	10	0,917	25,639	2,566	7,800	0,044	92,505	0,718	2,162
		Combinados	15	0,989	9,458	6,846	21,146	0,987	10,249	6,328	19,515

Combinados: dados originais + sintéticos

Os resultados obtidos demonstram diferenças marcantes no desempenho dos modelos PLS em função do tipo de pré-processamento espectral, do modelo (MIR e NIR) e do conjunto de dados (reais e dados combinados).

Para os modelos construídos exclusivamente com amostras reais, observou-se a ocorrência de sobreajuste em todos os casos avaliados, evidenciada por desempenhos excelentes na etapa de calibração, mas com queda significativa na capacidade preditiva durante a validação. Destaca-se que o aumento acentuado do erro de predição reforça a presença de *overfitting*, indicando que os modelos capturaram variações específicas do conjunto de calibração, sem adequada capacidade de generalização. Esse comportamento pode ser principalmente atribuído ao número reduzido de amostras reais, que limita a representatividade do espaço químico modelado (PATE et al., 2020).

A incorporação de dados sintéticos combinados aos dados reais resultou em mudança significativa no comportamento dos modelos, tanto para MIR quanto para NIR. Observou-se redução consistente da diferença entre RMSEC e RMSEP, além de maior proximidade entre os valores de coeficiente de correlação (r) obtidos em calibração e validação, indicando melhor capacidade de generalização (FAN et al., 2019; FOWLER et al., 2020).

Assim como o estudo de Santos et al. (2021) o número de variáveis latentes mostrou-se um critério decisivo na avaliação da qualidade dos modelos. Os modelos que utilizaram 5 variáveis latentes (LVs), especialmente no MIR, atingiram o melhor equilíbrio entre complexidade e desempenho. O número maior de LVs, como nos modelos NIR (até 15), não necessariamente implicou melhora significativa do r , mas sim aumento do risco de sobreajuste. De acordo com Freitas et al. (2025), a utilização de número reduzido de variáveis latentes contribui para a construção de modelos mais simples, estáveis e com maior capacidade de generalização.

Entre os pré-processamentos avaliados, o *Standard Normal Variate* (SNV) destacou-se por promover correção eficiente dos efeitos de espalhamento multiplicativo e variações de caminho óptico, resultando em espectros mais comparáveis e melhor alinhados quimicamente (WANG et al., 2022). Quando aplicado aos conjuntos de dados combinados, o SNV proporcionou elevados coeficientes de correlação e redução dos erros de calibração, evidenciando sua capacidade de realçar a informação química relevante.

O modelo PLS baseado em espectros MIR, utilizando dados sintéticos combinados aos dados reais e pré-processamento por SNV, demonstrou elevada robustez e confiabilidade para

a predição da concentração de Ciprofloxacina em amostras de leite liofilizado. Foram necessárias apenas cinco variáveis latentes para a construção do modelo (Tabela 1). Os coeficientes de correlação superiores a 0,97, tanto na calibração quanto na validação, indicam forte concordância entre os valores observados e preditos, com 95% de confiança (Figura 9). O erro de predição inferior a $15 \mu\text{g L}^{-1}$ (menos de 8% da faixa total) evidenciam a precisão do método para a quantificação e detecção de Ciprofloxacina, mesmo em baixas concentrações ($25 \mu\text{g L}^{-1}$). Este resultado é importante, pois esta concentração é menor que o limite máximo de resíduo (LMR) de $100 \mu\text{g L}^{-1}$ estabelecido para a Ciprofloxacina (UNIÃO EUROPEIA, 2010). Como resultado, fica demonstrada a sensibilidade do método e sua adequação para aplicações analíticas e de controle de qualidade em níveis regulatórios L^{-1} (JESUS et al., 2020; COQUEIRO et al., 2024; UNIÃO EUROPEIA, 2010).

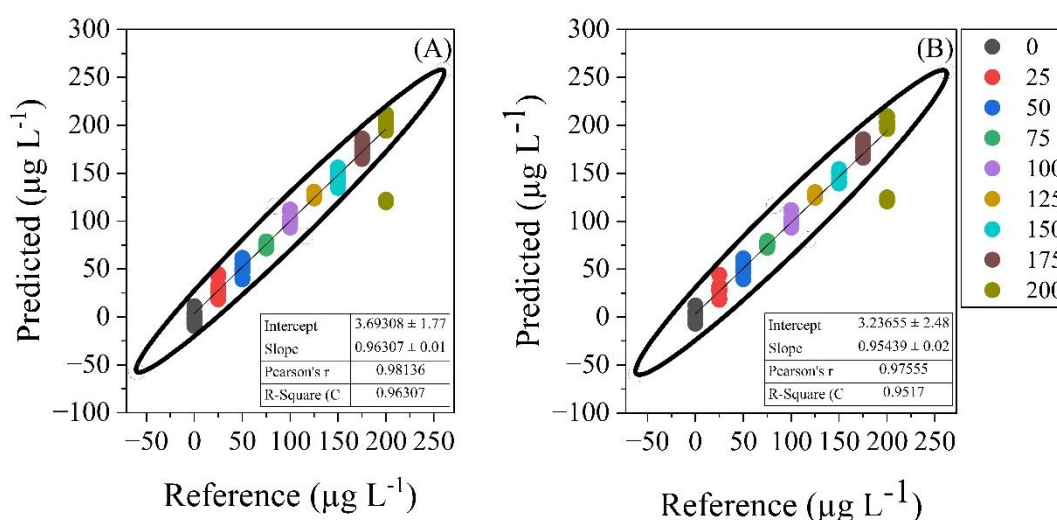


Figura 9. Gráfico de correlação entre os valores reais e previstos pelo modelo PLS-MIR utilizando dados combinados e pré-processamento por SNV, para conjuntos de calibração (A) e validação (B).

Valores de RPD superiores a 2 e RER superiores a 10 são amplamente reconhecidos como indicativos de modelos confiáveis e com bom poder preditivo (COQUEIRO et al., 2024; FREITAS et al., 2025). Conforme apresentado (Tabela 1), o modelo PLS baseado em espectros MIR, utilizando dados sintéticos combinados aos dados reais e pré-processamento por SNV, apresentou valores de RPD superiores a 4 e RER superiores a 14 tanto nas etapas de calibração quanto de validação, os quais são classificados como excelentes para o desempenho de modelos preditivos, evidenciando a elevada robustez e capacidade de generalização do modelo desenvolvido (JESUS et al., 2020; SANTOS et al., 2021).

Observa-se que a maioria das amostras se distribui ao longo da linha de identidade, enquanto algumas se posicionam fora da elipse de confiança (Figura 9), especialmente em regiões extremas de concentração ($200 \mu\text{g L}^{-1}$). Embora diagnósticos formais de *outliers* tenham identificado algumas amostras como discrepantes do ponto de vista estatístico, sua exclusão resultou em piora do desempenho preditivo do modelo no conjunto de validação, com o aumento do RMSEP e diminuição do r (Tabela 2).

Tabela 2. Métricas de desempenho dos dados de validação com e sem outliers.

Com Outliers		
r	RMSEP (%)	VIÉS
0,975	14,249	-1,324
Sem Outliers		
r	RMSEP (%)	VIÉS
0,917	27,513	-7,212

A significância do viés foi avaliada pelo teste t de Student, considerando o conjunto de validação com 107 amostras (graus de liberdade = 107) e nível de confiança de 95% (t crítico = 1,98). Para o modelo com *outliers*, o t calculado (0,97) foi inferior ao t crítico, indicando ausência de viés ($-1,324 \mu\text{g L}^{-1}$) estatisticamente significativo. Em contraste, no modelo sem outliers, o t calculado (2,82) superou o valor crítico, evidenciando a presença de erro sistemático ($-7,212 \mu\text{g L}^{-1}$) (LLETÍ et al., 2005; BOTELHO, MENDES e SENA, 2013). Esses resultados demonstram que a exclusão dos *outliers* reduziu a exatidão do modelo, indicando que tais amostras representam variabilidade química relevante e contribuem para a robustez da calibração PLS (DAYANANDA E COZZOLINO, 2022).

Diante da seleção do modelo mais robusto, torna-se fundamental avaliar a interpretação quimiométrica das variáveis latentes que sustentam seu desempenho. Nesse contexto, os *loadings* das cinco variáveis latentes (LVs) do modelo PLS baseado em espectros MIR pré-processados SNV, permitem relacionar o comportamento estatístico do modelo com as informações químicas extraídas do espectro (Figura 10).

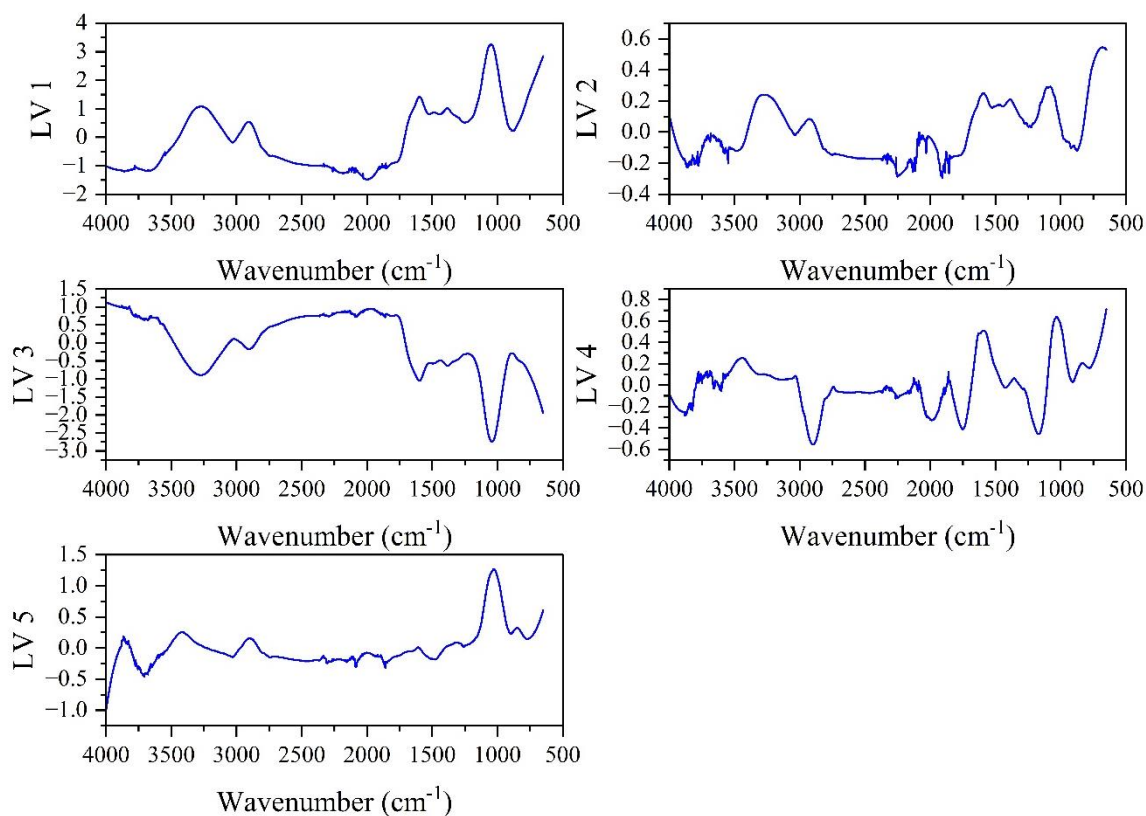


Figura 10. *Loadings* das variáveis latentes (LV1–LV5) do modelo de regressão por mínimos quadrados parciais (PLS–MIR), desenvolvido a partir de espectros pré-processados por SNV, para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado de vacas da raça Jersey

A LV1, responsável pela maior fração da variância explicada, apresenta contribuições relevantes nas regiões de 3600–3000 cm^{-1} , associadas aos estiramentos O–H e N–H, e entre 1700–1500 cm^{-1} , atribuídas às vibrações de C=O e C=C aromático, bandas características da Ciprofloxacina e da matriz láctea (PEREIRA et al., 2020; TAN et al., 2012). As LV2 e LV3 exibem sinais mais específicos nas regiões de 1600–1500 cm^{-1} e 1200–900 cm^{-1} , relacionadas aos estiramentos C–N, C–F e deformações do anel aromático, diretamente associados à estrutura química da Ciprofloxacina. Este resultado evidencia seu papel na discriminação e quantificação do padrão do antibiótico (SAHOO et al., 2011; SANTOS et al., 2026). As LV4 e LV5 apresentam contribuições menos intensas e mais localizadas, associadas a ajustes residuais e variações secundárias da matriz, sem comprometer a estabilidade do modelo. O uso de apenas cinco variáveis latentes, aliado à coerência química observada nos *loadings*, reforça a robustez e capacidade de generalização do modelo selecionado, em concordância com os excelentes resultados preditivos discutidos anteriormente.

Complementarmente à análise das variáveis latentes (Figura 11) o resultado do coeficiente de regressão (β), do modelo PLS baseado em espectros MIR pré-processados por

SNV, evidenciam diretamente as regiões espectrais com maior contribuição para a predição da concentração de Ciprofloxacina.

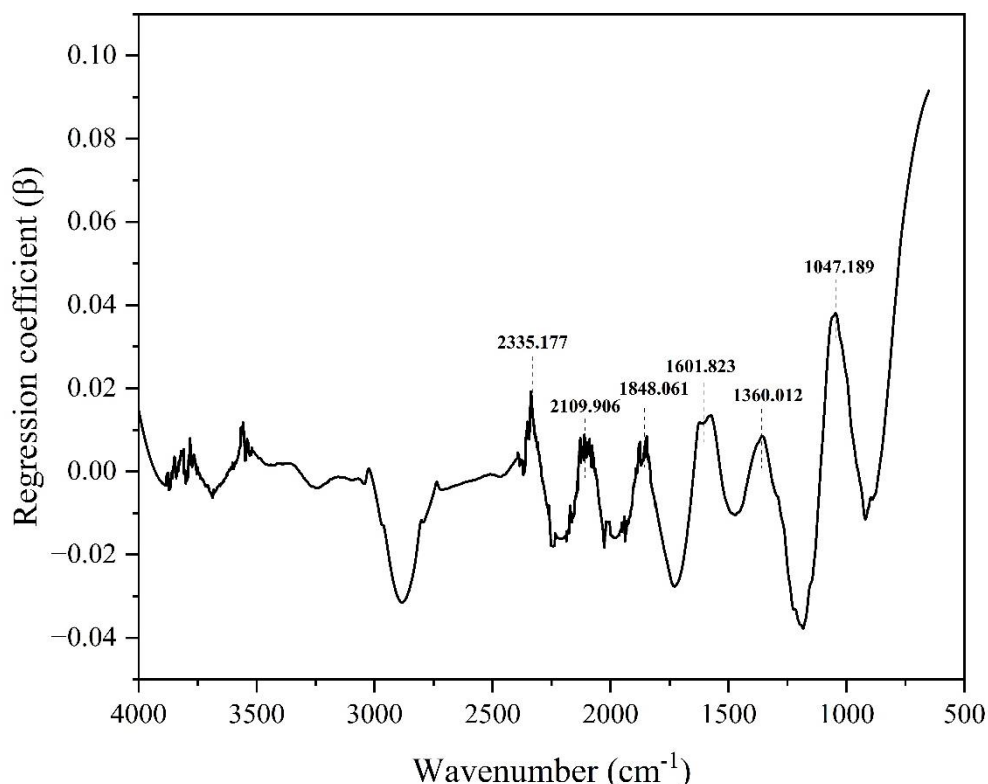


Figura 11. Coeficientes de regressão do modelo de regressão por mínimos quadrados parciais (PLS-MIR), com pré-processamento por SNV, para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado de vacas da raça Jersey.

Observam-se coeficientes de maior magnitude, positivos e negativos, principalmente nas regiões em torno de 1601 cm^{-1} , 1488–1360 cm^{-1} e 1047 cm^{-1} , atribuídas, respectivamente, às vibrações de C=C aromático, C–N e C–F, bem como aos estiramentos C–O, bandas características da estrutura molecular da Ciprofloxacina (SAHOO et al., 2011; SANTOS et al., 2026). Adicionalmente, sinais relevantes na região de 2335–2109 cm^{-1} podem estar associados a combinações vibracionais da matriz (PEREIRA et al., 2020). A predominância dessas regiões confirma que o modelo baseia sua capacidade preditiva em informações espectrais quimicamente coerentes com o analito de interesse, e não em variações aleatórias da matriz. Assim, a distribuição dos coeficientes de regressão reforça a robustez, seletividade e interpretação química do modelo PLS selecionado, em concordância com o desempenho preditivo.

3.4.1.1. Avaliação da robustez do modelo PLS por validação Monte Carlo

Com o objetivo de avaliar a robustez estatística do modelo selecionado, foi realizada uma validação externa por simulação Monte Carlo com 50 iterações (Tabela 3), aplicada ao modelo PLS–MIR com pré-processamento por SNV e dados combinados com cinco variáveis latentes (EZENARRO e SCHORN-GARCÍA, 2025).

Tabela 3. Desempenho do modelo PLS-MIR-SNV com dados combinados avaliado por simulação de Monte Carlo (média $\pm$ DP).

Calibração			
r	RMSEC (%)	RPD	RER
0.974 $\pm$ 0.011	14.334 $\pm$ 1.601	4.576 $\pm$ 0.543	14.134 $\pm$ 1.678
Validação			
r	RMSEP (%)	RPD	RER
0.960 $\pm$ 0.036	17.852 $\pm$ 3.952	3.796 $\pm$ 0.783	11.706 $\pm$ 2.416

Os resultados evidenciaram estabilidade consistente das métricas de desempenho, com valores médios de r superiores a 0,97 na calibração e 0,96 na validação, acompanhados de baixa variabilidade, conforme indicado pelos desvios padrão reduzidos. A proximidade entre os valores de RMSEC (14,33 $\pm$ 1,60) e RMSEP (17,85 $\pm$ 3,95) reforça a ausência de sobreajuste e confirma a adequada capacidade de generalização do modelo. Adicionalmente, os elevados valores médios de RPD ($> 3,7$) e RER (> 11) obtidos no conjunto de validação confirmam o excelente poder preditivo e a confiabilidade do modelo ao longo das diferentes partições aleatórias dos dados (JESUS et al., 2020; COQUEIRO et al., 2024).

Ao longo das 50 iterações do teste de Monte Carlo, observa-se nas Figuras 12A e 12B, que o conjunto de calibração apresenta desempenho consistente.

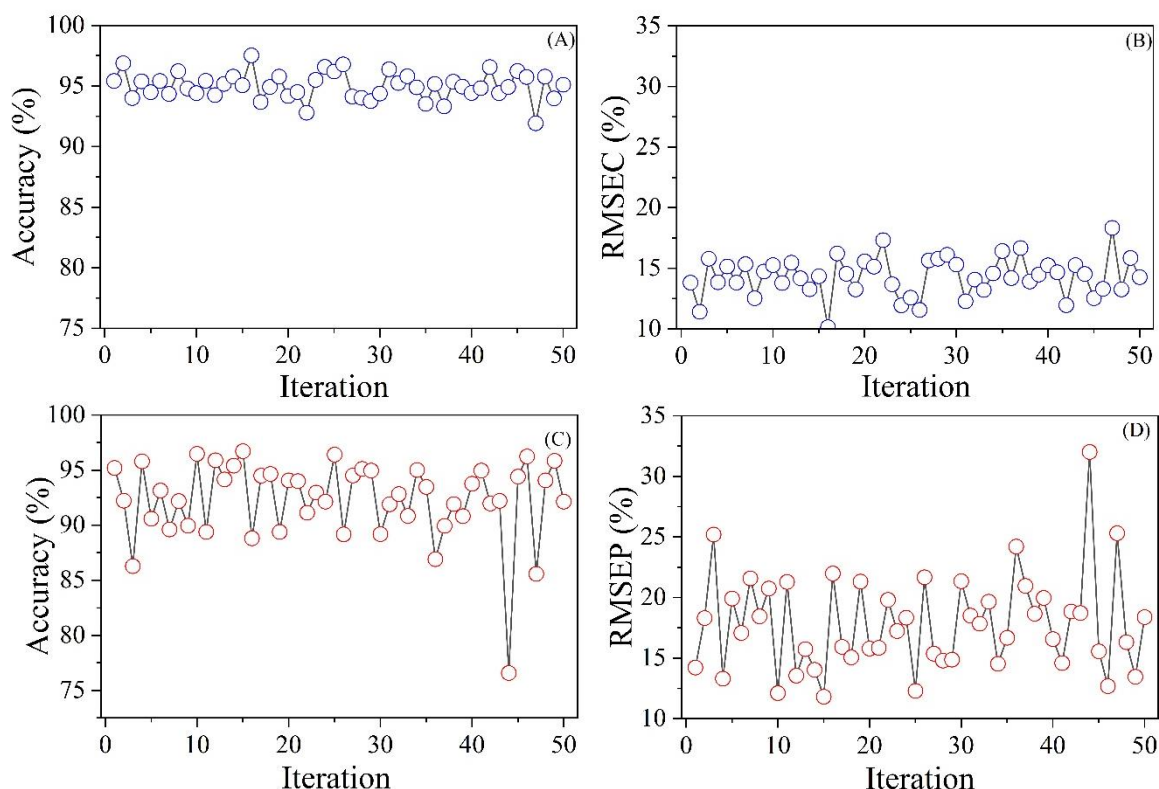


Figura 12. Variação das métricas de desempenho obtidas nas 50 iterações do teste de Monte Carlo, para o conjunto de calibração (A e B) e para o conjunto de validação (C e D).

A acurácia (Figura 12A), variou discretamente em torno de 95%, e o RMSEC (Figura B) apresentou-se relativamente baixo e estável, indicando boa capacidade de ajuste do modelo às amostras de calibração e ausência de sobreajuste. Já nas Figuras 12C e 12D, correspondentes ao conjunto de validação, nota-se maior variabilidade tanto na acurácia quanto no RMSEP. Na verdade, isso é o reflexo da aleatoriedade na divisão dos dados e a resposta do modelo frente a amostras não utilizadas na calibração. Apesar dessas oscilações, os valores médios mantêm-se próximos aos observados na calibração, o que demonstra robustez e reprodutibilidade do modelo nas diferentes iterações.

A validação por simulações Monte Carlo mostrou-se eficaz para avaliar a robustez, estabilidade e capacidade de generalização do modelo PLS–MIR desenvolvido, uma vez que permite estimar o desempenho preditivo a partir de múltiplas partições aleatórias dos conjuntos de calibração e validação (KVALHEIM, GRUN e RAJALAHTI, 2019; EZENARRO e SCHORN-GARCÍA, 2025). Diferentemente de uma única divisão calibração/validação, esse procedimento fornece avaliação mais realista da variabilidade das métricas do modelo, reduzindo o risco de conclusões otimistas e reforçando a confiabilidade dos resultados obtidos para a predição da concentração de Ciprofloxacina em amostras de leite liofilizado.

3.4.2. Regressão por Componentes Principais-PCR

De modo geral, os modelos desenvolvidos por Regressão por Componentes Principais (PCR) não apresentaram desempenho satisfatório para a quantificação do analito (Tabela 4), evidenciando limitações significativas de robustez e capacidade de generalização. Esses resultados podem estar associados a uma limitação conhecida da PCR, na qual os componentes principais são selecionados com base apenas na variância das variáveis espectrais (X), sem considerar diretamente sua relação com a resposta (y). Isso pode resultar em componentes subótimos para predição, contribuindo para baixo desempenho preditivo em conjuntos de validação (GAUTAM et al., 2015; PATE et al., 2020).

Tabela 4. Parâmetros de calibração por regressão em componentes principais (PCR) obtidos a partir de espectros MIR e NIR para a predição da concentração de Ciprofloxacina em leite liofilizado de vacas Jersey.

Pré-processamento	Modelo	Dados	Calibração					Validação			
			PCs	r	RMSEC	RPD	RER	r	RMSEP	RPD	RER
1ªDerivada	MIR	Reais	15	0.861	32.850	2.002	6.088	0.220	63.737	1.032	3.138
		Combinados	15	0.997	4.689	13.809	42.654	0.843	36.039	1.795	5.550
	NIR	Reais	2	0.438	64.755	1.075	3.220	0.374	62.821	1.015	3.258
		Combinados	2	0.267	62.386	1.038	3.206	0.298	62.289	1.038	3.211
SNV	MIR	Reais	15	0.767	41.433	1.588	4.827	0.205	64.046	1.027	3.123
		Combinados	15	0.999	3.031	21.362	65.984	0.613	52.383	1.235	3.818
	NIR	Reais	1	0.339	60.730	1.083	3.293	0.199	63.719	1.032	3.139
		Combinados	2	0.643	49.416	1.310	4.047	0.274	68.700	0.941	2.911
Correção de linha de base	MIR	Reais	13	0.861	32.850	2.002	6.088	0.220	63.737	1.032	3.138
		Combinados	15	0.991	8.792	7.364	22.747	0.563	53.536	1.208	3.736
	NIR	Reais	3	0.338	60.746	1.083	3.292	0.292	61.821	1.064	3.235
		Combinados	2	0.508	55.592	1.165	3.598	0.196	67.924	0.952	2.944
MSC	MIR	Reais	9	0.711	45.360	1.450	4.409	0.178	64.204	1.025	3.115
		Combinados	15	0.999	3.047	21.247	65.628	0.420	60.535	1.068	3.304
	NIR	Reais	1	0.341	60.680	1.084	3.296	0.199	63.728	1.032	3.138
		Combinados	2	0.653	48.911	1.324	4.089	0.290	67.902	0.953	2.945

Entre as abordagens testadas, o melhor desempenho foi obtido para os dados MIR combinados (amostras reais e sintéticas) submetidos à primeira derivada de Savitzky–Golay. Essa configuração apresentou melhorias consistentes nas métricas de calibração, indicando que a combinação da maior especificidade química do MIR, da ampliação do conjunto de dados por amostras sintéticas e da aplicação da derivada foi eficaz em realçar variações espectrais associadas à concentração do analito.

No entanto, apesar de representar o melhor resultado entre os modelos avaliados, essa abordagem ainda não pode ser considerada adequada do ponto de vista preditivo. Os erros de validação permaneceram elevados, e a discrepância entre calibração e validação indica que o modelo continua sensível ao sobreajuste, limitando sua aplicação prática. Assim, os resultados demonstram que, embora o uso de dados combinados MIR com derivada de Savitzky–Golay tenha melhorado o desempenho do PCR, o método, nas condições avaliadas, não se mostrou suficientemente robusto para a quantificação confiável do analito, reforçando a necessidade de explorar abordagens quimiométricas alternativas como PLS (GAUTAM et al., 2015).

4. Conclusão

A combinação de dados sintéticos e reais, aliada ao pré-processamento por Standard Normal Variate (SNV), foi fundamental para reduzir o sobreajuste e melhorar a capacidade de generalização do modelo. O modelo PLS–MIR–SNV, validado por simulações Monte Carlo, apresentou elevada robustez, sensibilidade e confiabilidade, evidenciando seu potencial para aplicações analíticas e de controle de qualidade. Os resultados demonstraram que a espectroscopia MIR, associada à regressão por mínimos quadrados parciais (PLS), apresentou desempenho superior em relação à NIR para a quantificação de Ciprofloxacina em amostras de leite liofilizado, permitindo a construção de modelos robustos e confiáveis.

5. Referencias

- ARSAND, J. B., JANK, L., MARTINS, M. T., HOFF, R. B., BARRETO, F., PIZZOLATO, T. M., SIRTORI, C. Determination of aminoglycoside residues in milk and muscle based on a simple and fast extraction procedure followed by liquid chromatography coupled to tandem mass spectrometry and time of flight mass spectrometry. **Talanta**, v. 154, p. 38-45,
- BEĆ, K. B., GRABSKA, J., HUCK, C. W. Near-infrared spectroscopy in bio-applications. **Molecules**, v. 25, n. 12, p. 2948, 2020.

BELLON-MAUREL, Véronique; MCBRATNEY, Alex. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils—Critical review and research perspectives. **Soil Biology and Biochemistry**, v. 43, n. 7, p. 1398-1410, 2011.

BERNIER, G.V., VISSCHERS, V. H., BODMER, M., MEYLAN, M. Dairy farmers' perceptions and practices related to the management of milk containing antibiotic residues. **Frontiers in Veterinary Science**, v. 8, p. 787828, 2022.

BOTELHO, B. G., MENDES, B. A., SENA, M. M. Development and analytical validation of robust near-infrared multivariate calibration models for the quality inspection control of mozzarella cheese. **Food Analytical Methods**, v. 6, n. 3, p. 881-891, 2013.

BRITO, I. C. P., DA SILVA SANTOS, H. L., SANTOS, K. E. O., DE ALBUQUERQUE FERNANDES, S. A. Modification and validation of miniaturized QuEChERS method for multi-residual determination of antibiotics in milk from a tropical region. **Microchemical Journal**, v. 207, p. 112025, 2024.

CAIS-SOKOLIŃSKA, D., TEICHERT, J., GAWAŁEK, J. JOLANTA. Formação de espuma e outras propriedades funcionais do leite de égua liofilizado. **Alimentos**, v. 12, n. 11, pág. 2274, 2023.

CANDEIAS, D. N. C., SILVA, K. M., PEREIRA, H. S., BEZERRA, L. P., DA SILVA, J. D. S., DE SOUSA FERNANDES, D. D., DINIZ, P. H. G. D. Geographical origin authentication of instant coffee from southern Bahia using MIR and NIR spectroscopy coupled with DD-SIMCA. **Food Chemistry**, v. 479, p. 143698, 2025.

CHRISTOPHE, O. S., GRELET, C., SIMON, H., LEBLOIS, J., BAETEN, V., DEHARENG, F. Stress detection/assessment in dairy cows by the determination of cortisol in hair by near-infrared (NIR), mid-infrared (MIR) and Raman spectroscopy. **Scientific Reports**, v. 15, n. 1, p. 27083, 2025.

COQUEIRO, J. S., LIMA, A. B. S., JESUS, J. C., SILVA, R. R., FERRÃO, S. P. B., SANTOS, L. S. Ensuring authenticity of cinnamon powder: Detection of adulteration with coffee husk and corn meal using NIR, MIR spectroscopy and chemometrics. **Food Control**, v. 166, p. 110681, 2024.

DAOUD, S., BOU-MAROUN, E., WASCHATKO, G., HOREMANS, B., MESTDAGH, R., BILLECKE, N., & CAYOT, P. Detection of lipid oxidation in infant formulas: Application of infrared spectroscopy to complex food systems. **Foods**, v. 9, n. 10, p. 1432, 2020.

DAYANANDA, B., COZZOLINO, D. Beyond the Black Box—Practical considerations on the use of chemometrics combined with sensing technologies in food science applications. **Chemosensors**, v. 10, n. 8, p. 323, 2022.

DINÇ, B., ÜSTÜNSOY, R., ERTAŞ, T., ŞEN, E. A comparative analysis of ten milk samples with differential scanning calorimetry and Fourier transform infrared spectroscopy. **Food and Health**, v. 9, n. 3, p. 184-192, 2023.

DONG, Y., LI, Y., ZHENG, H., WANG, R., XU, M. A new dynamic model and transfer learning based intelligent fault diagnosis framework for rolling element bearings race faults: Solving the small sample problem. **ISA transactions**, v. 121, p. 327-348, 2022.

EUROPEAN UNION. *Commission Regulation (EU) No 37/2010 of 22 December 2009* on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin. Official Journal of the European Union, L 15, p. 1–72, 20 Jan. 2010. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010R0037>. Accessed on: 4 June 2025.

EZENARRO, J., SCHORN-GARCÍA, D. How are chemometric models validated? A systematic review of linear regression models for NIRS data in food analysis. **Journal of Chemometrics**, v. 39, n. 6, p. e70036, 2025.

FAN, Y., CUI, X., HAN, H., LU, H. Chiller fault diagnosis with field sensors using the technology of imbalanced data. **Applied Thermal Engineering**, v. 159, p. 113933, 2019.

FATEMI, A., SINGH, V., KAMRUZZAMAN, M. Identification of informative spectral ranges for predicting major chemical constituents in corn using NIR spectroscopy. **Food chemistry**, v. 383, p. 132442, 2022.

FIGUEIRA, A., VAZ, B. Survey on synthetic data generation, evaluation methods and GANs. **Mathematics**, v. 10, n. 15, p. 2733, 2022.

FOLLI, G. S., NASCIMENTO, M. H., LOVATTI, B. P., ROMÃO, W., FILGUEIRAS, P. R. A generation of synthetic samples and artificial outliers via principal component analysis and evaluation of predictive capability in binary classification models. **Chemometrics and Intelligent Laboratory Systems**, v. 251, p. 105154, 2024.

FOWLER, E. E., BERGLUND, A., SCHELL, M. J., SELLERS, T. A., ESCHRICH, S., HEINE, J. Empirically-derived synthetic populations to mitigate small sample sizes. **Journal of biomedical informatics**, v. 105, p. 103408, 2020.

FREITAS OLIVEIRA, T., SANTOS, G. R., JESUS, J. C., COQUEIRO, J. S., SILVA, R. R., FERRÃO, S. P. B., SANTOS, L. S. Assessing the authenticity of the sport supplement creatine using spectroscopic techniques and multivariate analysis. **Food Chemistry**, v. 484, p. 144282, 2025.

FREITAS, A. G. M., MINHO, L. A. C., DE MAGALHÃES, B. E. A., DOS SANTOS, W. N. L., SANTOS, L. S., & DE ALBUQUERQUE FERNANDES, S. A. Infrared spectroscopy combined with random forest to determine tylosin residues in powdered milk. **Food Chemistry**, v. 365, p. 130477, 2021.

GAUTAM, R., VANGA, S., ARIESE, F., UMAPATHY, S. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. **EPJ Techniques and Instrumentation**, v. 2, n. 1, p. 8, 2015.

HAO, Y., LI, X., ZHANG, C. Improving prediction model robustness with virtual sample construction for near-infrared spectra analysis. **Analytica Chimica Acta**, v. 1279, p. 341763, 2023.

HEINE, J., FOWLER, E. E., BERGLUND, A., SCHELL, M. J., ESCHRICH, S. Techniques to produce and evaluate realistic multivariate synthetic data. **Scientific Reports**, v. 13, n. 1, p. 12266, 2023.

HELLEQUIN, L. V., BORRÀS, V. J., ROMANN, P., VISHWANATHAN, N., SOUQUET, J., VILLIGER, T. K. Synthetic spectral libraries for Raman model calibration. **Analytical and Bioanalytical Chemistry**, v. 417, p. 5675–5689, 2025.

HU, J., XU, Z., LI, M., HE, Y., LIU, Y. Discriminant analysis and quantitative study of antibiotics in infant milk powder based on hyperspectral detection. **Vibrational Spectroscopy**, v. 114, p. 103244, 2021.

JESUS, J. C. D., SILVA, G. J., GONÇALVES, B. H. R., SOUZA, M. R. D., SANTOS, L. S., FERRÃO, S. P. Quick identification of the time of maturation of artisanal minas cheese by FTIR-ATR spectroscopy and multivariate techniques. **Journal of the Brazilian Chemical Society**, v. 31, n. 10, p. 2000-2011, 2020.

KALUNKE, R. M., GRASSO, G., D'OVIDIO, R., DRAGONE, R., FRAZZOLI, C. Detection of ciprofloxacin residues in cow milk: a novel and rapid optical β -galactosidase-based screening assay. **Microchemical Journal**, v. 136, p. 128-132, 2018.

KHWAZA, V., MLALA, S., & ADERIBIGBE, B. A. Advancements in synthetic strategies and biological effects of ciprofloxacin derivatives: A Review. **International Journal of Molecular Sciences**, v. 25, n. 9, p. 4919, 2024.

KULKO, R. D., HANUS, A., ELSEER, B. Generative Artificial Intelligence for Synthetic Spectral Data Augmentation in Sensor-Based Plastic Recycling. **Sensors**, v. 25, n. 13, p. 4114, 2025.

KVALHEIM, O. M., GRUNG, B., & RAJALAHTI, T. Number of components and prediction error in partial least squares regression determined by Monte Carlo resampling strategies. **Chemometrics and Intelligent Laboratory Systems**, v. 188, p. 79-86, 2019.

LIMA, A. B. S., BATISTA, A. S., JESUS, J. C., JESUS SILVA, J., DE ARAÚJO, A. C. M., SANTOS, L. S. Fast quantitative detection of black pepper and cumin adulterations by near-infrared spectroscopy and multivariate modeling. **Food Control**, v. 107, p. 106802, 2020.

LLETÍ, R., MELÉNDEZ, E., ORTIZ, M. C., SARABIA, L. A., SÁNCHEZ, M. S. Outliers in partial least squares regression: Application to calibration of wine grade with mean infrared data. **Analytica Chimica Acta**, v. 544, n. 1-2, p. 60-70, 2005.

MOUAZEN, A. M., KUANG, B., DE BAERDEMAEKER, J., RAMON, H. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. **Geoderma**, v. 158, n. 1-2, p. 23-31, 2010.

NITHYA, R., SUNDARAM, N. M. Biodegradation and cytotoxicity of ciprofloxacin- loaded hydroxyapatite-polycaprolactone nanocomposite film for sustainable bone implants. **International Journal of Nanomedicine**, v. 10, p. 119–127, 2015.

PATE, A., EMSLEY, R., SPERRIN, M., MARTIN, G. P., VAN STAA, T. Impact of sample size on the stability of risk scores from clinical prediction models: a case study in cardiovascular disease. **Diagnostic and prognostic research**, v. 4, p. 1-12, 2020.

PEREIRA, C., LUIZ, L. C., BELL, M. J. V., ANJOS, V. Near and mid infrared spectroscopy to assess milk products quality: a review of recent applications. **J. Dairy Res. Technol**, v. 3, n. 1, p. 1-10, 2020.

PEREIRA, E. S., CRUZ-TIRADO, J. P., CRIPPA, B. L., MORASI, R. M., DE ALMEIDA, J. M., BARBIN, D. F., SILVA, N. C. C. Portable near infrared (NIR) spectrometer coupled with machine learning to classify milk with subclinical mastitis. **Food Control**, v. 163, p. 110527,

2024.

PRIYADARSINI, S., LAXMIKARTHIKA, V. S., RANA, S. S., AHMAD, F., CHAUDHARY, Y., PANDA, M., & KUMAR, Y. Comparative physicochemical analysis of freeze-dried milk powders from camel, goat and donkey. **Applied Food Research**, p. 101356, 2025.

PRZYBYLSKA, N., ŚLIWIŃSKA-BARTKOWIAK, M., KOŚCIŃSKI, M., ROTNICKI, K., BARTKOWIAK, M., JURGA, S. Confined effect of water solution of ciprofloxacin in carbon nanotubes studied by Raman and Fourier Transform Infrared Spectroscopy methods. **Journal of Molecular Liquids**, v. 336, p. 115938, 2021.

QIAO, M., YING, G. G., SINGER, A. C., ZHU, Y. G. Review of antibiotic resistance in China and its environment. **Environment international**, v. 110, p. 160-172, 2018.

QIU, G., LÜ, E., LU, H., XU, S., ZENG, F., SHUI, Q. Single-kernel FT-NIR spectroscopy for detecting supersweet corn (*Zea mays* L. *saccharata* sturt) seed viability with multivariate data analysis. **Sensors**, v. 18, n. 4, p. 1010, 2018.

RAJALAKSHMI, K., GUNASEKARAN, S., KUMARESAN, S. Vibrational spectra, electronic and quantum mechanical investigations on ciprofloxacin. **Indian Journal of Physics**, v. 88(7), p. 733–744, 2014.

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.

RODRÍGUEZ-HERNÁNDEZ, P., NÚÑEZ-SÁNCHEZ, N., MOLINA-GAY, S., RODRÍGUEZ-ESTÉVEZ, V., CARDOSO-TOSET, F. Near infrared spectroscopy analysis as a screening tool to classify milk from bovine subclinical mastitis and promote pathogen-based therapy. **Applied Food Research**, v. 5, n. 1, p. 100651, 2025.

SAHOO, S. U. B. H. A. S. H. R. E. E., CHAKRABORTI, C. K., MISHRA, S. C., NANDA, U. N., NAIK, S. FTIR and XRD investigations of some fluoroquinolones. **International Journal of Pharmacy and Pharmaceutical Sciences**, v. 3, n. 3, p. 165-170, 2011.

SANTANA, V. R., MAGALHÃES, B. E. A. D., MINHO, L. A. C., FREITAS, A. G. M. D., SOUZA, R. D., BATISTA, A. S., ... FERNANDES, S. A. D. A. Random forest and infrared spectroscopy to detect contamination of freeze-dried milk by ivermectin residues. **Ciência Rural**, v. 55, n. 12, p. e20240536, 2025.

- SANTOS, C. P. D. S., BARBOSA, R. S., SOUZA, R. D., FERNANDES, S. A. A. Mid-infrared spectroscopy combined with chemometrics for the detection of ciprofloxacin and Bronopol® residues in milk. **Food Chemistry**, v. 504, p. 147826, 2026.
- SANTOS, I. A., CONCEIÇÃO, D. G., VIANA, M. B., DE JESUS SILVA, G., SANTOS, L. S., FERRÃO, S. P. B. NIR and MIR spectroscopy for quick detection of the adulteration of cocoa content in chocolates. **Food Chemistry**, v. 349, p. 129095, 2021.
- SCHUMER, N. G., AHMED, M. W., RAUSCH, K., SINGH, V., KAMRUZZAMAN, M. Chemometric-based approach for economically motivated fraud detection in organic spices via NIR spectroscopy. **Journal of Food Composition and Analysis**, v. 142, p. 107538, 2025.
- TAN, Z., TAN, F., ZHAO, L., LI, J. The synthesis, characterization and application of ciprofloxacin complexes and its coordination with copper, manganese and zirconium ions. **Journal of Crystallization Process and Technology**, v. 2, n.2, p. 55-63, 2012.
- TUREL, I., BUKOVEC, P., QUIRÓS, M. Crystal structure of ciprofloxacin hexahydrate and its characterization. **International Journal of Pharmaceutics** v. 152, p. 59–65, 1997.
- WANG, J., LIAO, C., ZHAO, J., GAO, W. Vis–NIR Hyperspectral Dimensionality Reduction for Nondestructive Identification of China Northeast Rice. In: **The International Conference on Image, Vision and Intelligent Systems (ICIVIS 2021)**. Singapore: Springer Nature Singapore, 2022. p. 913-922, 2022.
- WEMETTE, M.; SAFI, A. G.; BEAUVAIS, W.; CERES, K.; SHAPIRO, M.; MORONI, P.; WELCOME, F. L.; IVANEK, R. New York State dairy farmers' perceptions of antibiotic use and resistance: A qualitative interview study. **PLOS ONE**, v. 15, n. 5, e0232937, 2020.
- XIONG, Y., ZHANG, D., YE, C., WANG, Y., DENG, X., DENG, D., HUANG, G. Ultra-sensitive detection of ciprofloxacin hydrochloride in milk by molecularly imprinted electrochemical sensor based on S-CoFe-MOFs/AuNPs. **Journal of Food Composition and Analysis**, v. 122, p. 105439, 2023.
- ZHANG, W., KASUN, LC, WANG, QJ, ZHENG, Y., LIN, Z. A review of machine learning for near-infrared spectroscopy. **Sensors**, v. 22, n. 24, p. 9764, 2022.
- ZHOU, S., ZHANG, X., ZHANG, J., ZANG, C., FAN, R., WANG, J., YANG, Y. Differences in physicochemical properties and proteomics analysis of spray-and freeze-dried milk powders from bovine, goat, and horse sources. **Journal of Dairy Science**, v. 108, n. 2, p. 1367-1379, 2025.

CAPÍTULO IV: ARTIGO II - USO DE ESPECTROSCOPIAS VIBRACIONAIS E QUIMIOMETRIA PARA TRIAGEM RÁPIDA DE CIPROFLOXACINA EM LEITE EM PÓ OBTIDO POR LIOFILIZAÇÃO

1. Introdução

O leite em pó é amplamente utilizado na indústria de alimentos devido à sua estabilidade, conveniência e versatilidade tecnológica (WEDEL et al., 2022; SCHONS et al., 2025). Entre os métodos empregados para sua obtenção, o *spray drying* é o mais comum na escala industrial, por apresentar menor custo operacional e maior eficiência produtiva (WEDEL et al., 2022). Outra técnica utilizada para obtenção do leite em pó é a liofilização, embora tenha menor viabilidade industrial, tem a vantagem de preservar de forma mais eficaz os compostos bioativos e as propriedades físico-químicas do leite, uma vez que ocorre sob baixas temperaturas e pressão reduzida (ZHANG et al., 2022; ALIBEKO et al., 2025). Essa característica torna o leite liofilizado particularmente adequado para análises laboratoriais, nas quais se busca manter a integridade química da amostra (DONEVA et al., 2021; ALIBEKO et al., 2025; ZHOU et al., 2025).

A produção leiteira está frequentemente associada ao uso de medicamentos veterinários em rebanhos, empregados com fins terapêuticos ou profiláticos em animais de produção (ALVES et al., 2019). No entanto, o uso inadequado ou indiscriminado de antibióticos pode resultar na presença de resíduos no leite, comprometendo sua qualidade e segurança para o consumo humano (BRITO et al., 2024).

Entre os contaminantes de maior relevância, destacam-se os antibióticos da classe das fluoroquinolonas, especialmente a Ciprofloxacina, amplamente utilizada na medicina veterinária para o tratamento de infecções bacterianas em rebanhos leiteiros (MAHMOOD et al., 2016). A presença residual desse medicamento em produtos lácteos representa um risco à saúde pública, podendo contribuir para o desenvolvimento de resistência antimicrobiana e provocar reações adversas em consumidores sensíveis (TASCI, CANBAY e DOGANTURK, 2021). Por esse motivo, a legislação vigente estabelece um limite máximo de resíduo (LMR) de 100 µg L⁻¹ de Ciprofloxacina em leite (UNIÃO EUROPEIA, 2010), valor que demanda métodos de detecção de alta sensibilidade.

Nas fazendas e laticínios, a triagem inicial de resíduos antimicrobianos é geralmente realizada por meio de testes rápidos, devido à sua praticidade e baixo custo (BRITO et al., 2024). Dentre os mais utilizados, destacam-se os ensaios baseados em inibição microbiana, como o Delvotest® (BION et al., 2016; JAGESAR e KERKHOF, 2023), Charm MRL® (SHI

et al., 2021; HASONOVÁ et al., 2025) e Snap Beta-Lactam® (DUREL, GALLINA E PELLET, 2019), que detectam a presença de antibióticos a partir da inibição do crescimento bacteriano ou pela interação com enzimas específicas. Esses métodos apresentam vantagens operacionais, como rapidez e facilidade de uso, porém possuem limitações quanto à especificidade, sensibilidade e faixa de antibióticos detectáveis, podendo gerar falsos positivos ou negativos dependendo da matriz e das condições de amostragem (DUREL, GALLINA E PELLET, 2019; BRITO et al., 2024).

Nesse contexto, as técnicas espectroscópicas, como o infravermelho médio (MIR) e o infravermelho próximo (NIR), têm se consolidado como alternativas promissoras para o monitoramento de resíduos de antibióticos em produtos de origem animal (FENG et al., 2024; LUIZ et al., 2018; TEIXEIRA et al., 2020; QIN, XIE E YING, 2016). Essas abordagens são não destrutivas, rápidas e ambientalmente sustentáveis, permitindo a obtenção de informações estruturais detalhadas sobre a amostra e possibilitando análises simultâneas de múltiplos componentes (RINGSTED, SIESLER e ENGELSEN, 2017; NÍ FHUARÁIN et al., 2024). Quando combinadas com ferramentas quimiométricas, as técnicas espectroscópicas tornam-se capazes de detectar diferenças químicas sutis que não são evidentes na inspeção visual dos espectros (ZHU et al., 2019; AMIRVARESI et al., 2021; MOGHADDAM et al., 2022).

Dentre os métodos quimiométricos mais empregados, a Análise de Componentes Principais (PCA) permite a exploração da variabilidade espectral e a identificação de padrões e agrupamentos entre amostras, enquanto a Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) possibilita a construção de modelos supervisionados para classificação de amostras em diferentes categorias de concentração (AMIRVARESI et al., 2021; YANG et al., 2021). Tais abordagens são particularmente úteis quando aplicadas a matrizes complexas, como o leite em pó liofilizado, nas quais os efeitos da matriz podem sobrepor-se às bandas do analito (ZHU et al., 2020; CHATZIMICHAIL et al., 2025).

Considerando que a robustez dos modelos multivariados depende diretamente da representatividade e variabilidade do conjunto de dados, foram empregadas neste estudo amostras sintéticas geradas via Análise de Componentes Principais (PCA). A utilização de amostras sintéticas tem como finalidade ampliar a base amostral de forma estatisticamente coerente, reduzindo possíveis vieses decorrentes de conjuntos experimentais limitados (FLANAGAN e GLAVIN, 2023). A geração de amostras via PCA, cria novas observações que preservam a estrutura de correlação multivariada dos espectros originais, uma vez que as amostras são sintetizadas no espaço dos scores e posteriormente reconstruídas a partir das componentes principais (PADINJAREKKARA, ALECCI e POPA, 2024). Dessa forma, é

possível aumentar a densidade do espaço amostral sem introduzir ruído artificial, garantindo maior estabilidade, generalização e capacidade preditiva aos modelos quimiométricos (PADINJAREKKARA, ALECCI e POPA, 2024; CALDEIRA et al., 2024).

Dessa forma, o presente trabalho tem como objetivo avaliar o desempenho de métodos quimiométricos aplicados à análise exploratória e de classificação (PCA e PLS-DA) de espectros MIR e NIR de leite em pó liofilizado fortificado com diferentes concentrações de Ciprofloxacina, utilizando dados reais e dados combinados (reais combinados com sintéticos) submetidos a distintos pré-processamentos espectrais. A proposta busca contribuir para o desenvolvimento de estratégias de triagem rápida, confiável e multivariada, que possam complementar ou substituir testes convencionais no monitoramento de resíduos de antibióticos em leite.

2. Material e Métodos

2.1. Amostras

Foram coletadas do tanque resfriador amostras de leite cru do rebanho de vacas da raça Jersey que não receberam tratamento com antibióticos nos últimos 60 dias, pertencentes ao rebanho da Embrapa Clima Temperado, localizada na cidade de Pelotas – Rio Grande do Sul.

2.2. Delineamento experimental

O padrão do antibiótico Ciprofloxacina de grau analítico foi utilizado na forma de pó. Inicialmente, foi preparada uma solução estoque de Ciprofloxacina a $10.000 \mu\text{g L}^{-1}$ em leite. A solução estoque foi preparada imediatamente antes do uso, diluindo-as até uma concentração equivalente a duas vezes o LMR (Limite Máximo de Resíduos) deste antibiótico para obtenção das concentrações: 25, 50, 75, 100, 125, 150, 175 e $200 \mu\text{g L}^{-1}$.

A amostra controle (0), denominada amostra “pura” não recebeu adição do antibiótico. A impregnação foi realizada em cinco dias consecutivos para cada concentração, homogeneizando-se cuidadosamente as amostras por agitação manual por 1 minuto, e encaminhadas para liofilização no Laboratório de Análise de Alimentos, da Embrapa Clima Temperado e posteriores análises no Laboratório de no Laboratório do Centro de Estudos em Leite (CEL), do Departamento de Tecnologia Rural e Animal (UESB).

2.3. Liofilização

As amostras de leite, tanto puras quanto fortificadas com o padrão de Ciprofloxacina, foram transferidas para microtubos de 2 mL (Eppendorf®), congeladas a $-20\text{ }^{\circ}\text{C}$ e posteriormente submetidas à liofilização a $-48\text{ }^{\circ}\text{C} \pm 2\text{ }^{\circ}\text{C}$ e 0,100 mBar durante 48 h, utilizando liofilizador L108 (LioTop, São Paulo, Brasil).

2.4. Espectroscopia no infravermelho médio (FTIR-MIR)

As análises espectroscópicas foram realizadas em amostras de leite liofilizado (com e sem adição de Ciprofloxacina) utilizando espectrômetro Agilent Cary 630 (Santa Clara, EUA), operando na faixa de $4000\text{--}650\text{ cm}^{-1}$. Os espectros foram obtidos no modo de absorbância, empregando a técnica de reflectância total atenuada (ATR) com cristal de diamante. As leituras foram registradas e processadas com os softwares Microlab e Resolution Pro (Agilent Technologies). Para cada amostra, uma pequena porção do pó liofilizado foi aplicada diretamente sobre o cristal, garantindo contato uniforme com a superfície de leitura.

2.5. Espectroscopia no infravermelho próximo (NIR)

Os espectros NIR foram obtidos com espectrômetro SpectraStar 2500XL (Unity Scientific, Brookfield, CT, EUA), equipado com fonte de halogênio de tungstênio e detector InGaAs. As medidas foram realizadas no modo de refletância (%R), convertidas em absorbância ($\log 1/R$). Os espectros foram coletados na faixa de $1100\text{--}2500\text{ nm}$, com resolução de 1 nm, totalizando 1400 variáveis espectrais, resultantes da média de 64 varreduras por amostra.

2.6. Pré-processamento dos dados espectrais

Para cada conjunto espectral (MIR e NIR), foram avaliados quatro diferentes pré-processamentos com o intuito de reduzir ruído instrumental, efeitos de espalhamento e variações de linha de base. Os métodos testados incluíram: (i) primeira derivada de Savitzky–Golay (DSG) com janela de 15 pontos e polinômio de segunda ordem, (ii) correção de linha de base, (iii) Correção Multiplicativa de Espalhamento - *Multiplicative Scatter Correction* (MSC) e (iv) Variável Normal Padrão - *Standard Normal Variate* (SNV).

2.7. Amostras sintéticas

As amostras sintéticas utilizadas neste estudo são as mesmas geradas e descritas detalhadamente no Artigo 1, por meio de Análise de Componentes Principais (PCA), conforme metodologia de Folli et al. (2024). Assim, não foi realizada nova etapa de geração sintética.

Para o presente trabalho, os espectros sintéticos e reais foram reorganizados em três classes de concentração de Ciprofloxacina, de forma a refletir o comportamento químico do analito em faixas distintas:

- Classe 0: amostras sem adição de antibiótico;
- Classe 1: amostras com concentrações entre 25 e 100 $\mu\text{g L}^{-1}$ (até o LMR da Ciprofloxacina) (UNIÃO EUROPEIA, 2010);
- Classe 2: amostras entre 125 e 200 $\mu\text{g L}^{-1}$ (acima do LMR).

Essa reclassificação permitiu avaliar o desempenho dos modelos quimiométricos considerando o agrupamento por faixas de concentração, garantindo consistência estatística e comparabilidade com o conjunto de dados original.

2.8. Análises quimiométricas

As análises quimiométricas foram realizadas com o objetivo de explorar a variabilidade espectral e avaliar o potencial discriminante entre as amostras de leite em pó liofilizado contendo diferentes concentrações de Ciprofloxacina. Para tanto, aplicaram-se as técnicas de Análise de Componentes Principais (PCA) e Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA), utilizando os conjuntos espectrais MIR e NIR previamente obtidos.

A PCA foi empregada como ferramenta exploratória não supervisionada para identificar tendências, agrupamentos e possíveis outliers no conjunto de dados. Essa análise foi conduzida considerando as três classes de amostras definidas experimentalmente: classe 0 (sem antibiótico), classe 1 (25 a 100 $\mu\text{g L}^{-1}$) e classe 2 (125 a 200 $\mu\text{g L}^{-1}$). A seguir, foi aplicada a PLS-DA, técnica supervisionada amplamente empregada para classificação de dados espectrais. Essa abordagem buscou maximizar a separação entre as três classes de concentração de Ciprofloxacina, modelando a relação entre as variáveis espectrais e as categorias definidas.

Todas as etapas de modelagem quimiométrica foram realizadas no ambiente MATLAB® R2025b (MathWorks, Natick, MA, EUA).

2.8.1. Análise de Componentes Principais – PCA

A Análise de Componentes Principais (PCA) foi utilizada para reduzir a dimensionalidade de um conjunto de variáveis correlacionadas, facilitando a interpretação dos dados e a identificação de padrões. Os dados foram organizados em matriz $n \times p$, linhas representavam as amostras de diferentes concentrações de Ciprofloxacina e as colunas foram

preenchidas com os dados de absorbância dos espectros MIR e NIR, que foram usados como variáveis independentes.

Foram extraídos dois componentes principais (CP's), representando combinações lineares entre variáveis originais que explicam mais de 95% da variância total dos dados. A variável concentração foi usada como identificador para cada observação.

2.8.2. Análise de Mínimos Quadrados Parciais - PLS-DA

A análise discriminante por mínimos quadrados parciais (PLS-DA) foi aplicada aos espectros MIR e NIR com o objetivo de classificar as amostras de acordo com os níveis de antibiótico. Inicialmente, os dados espectrais foram organizados em uma matriz **X**, contendo as intensidades espectrais como variáveis independentes, e um vetor **y**, representando as classes categóricas associadas às concentrações do analito. A divisão dos dados em conjuntos de calibração (60%) e validação externa (40%) foi realizada pelo algoritmo de Kennard–Stone (FOLLI et al., 2024).

As métricas de exatidão (R^2), especificidade, precisão e sensibilidade foram empregadas como parâmetros de avaliação do desempenho dos modelos desenvolvidos (FOLLI et al., 2024).

3. Resultados e Discussão

3.1. Investigação espectral e procedimentos de pré-processamento

Os espectros MIR do leite liofilizado puro (0) e do leite liofilizado fortificado com Ciprofloxacina em níveis até o LMR (1) e acima do LMR (2) apresentaram bandas características da matriz láctea, com variações sutis de intensidade e forma associadas à presença do antibiótico (Figura 1).

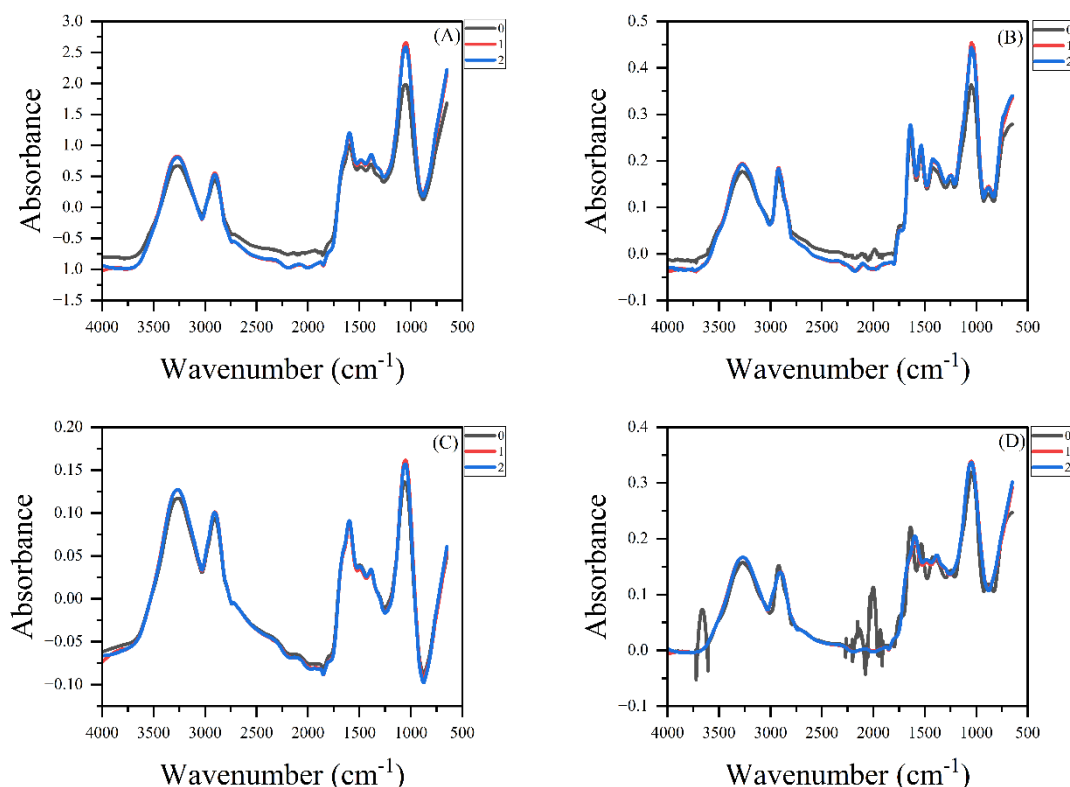


Figura 1. Espectro MIR processados por SNV (A), MSC (B), correção de linha de base (C) e DSG (D).

É possível observar que os pré-processamentos por SNV (Figura 1A), MSC (Figura 1B) e correção de linha de base (Figura 1C) apresentaram comportamento semelhante na redução de variações não químicas dos espectros MIR, principalmente associadas a efeitos físicos da matriz liofilizada, como espalhamento e variações de intensidade global. Embora visualmente próximos, esses métodos atuam de forma distinta: a correção de linha de base remove componentes aditivos, enquanto MSC e SNV minimizam efeitos multiplicativos e de espalhamento (YAN, 2025; ROSALES-MARTÍNEZ et al., 2025). Nestes três processamentos observa-se picos sobrepostos que não realçaram com clareza a influência química do padrão do antibiótico.

A aplicação da primeira derivada de Savitzky–Golay (Figura 1D) apresentou comportamento distinto dos métodos de normalização e correção de espalhamento, ao atuar principalmente no realce de variações locais do espectro e na resolução de bandas sobrepostas (ROSALES-MARTÍNEZ et al., 2025). No leite liofilizado, regiões dominadas por bandas largas nas regiões 3600–3000 cm⁻¹ (O–H e N–H) e 3000–2800 cm⁻¹ (C–H de lipídios) (MOHAMED et al., 2021), apresentaram redução do efeito de sobreposição, permitindo melhor distinção das contribuições relativas da matriz. Já na região de 1700–1400 cm⁻¹, correspondente

principalmente às bandas de amida I e II das proteínas (DANILOSKI et al., 2022; SANTANA et al., 2025), a derivada evidenciou mudanças no perfil espectral associadas à fortificação com Ciprofloxacina. Essas alterações estão relacionadas à presença de grupos carbonila (C=O) e amina (N-H) do antibiótico (RATHORE et al., 2019), cujas absorções se sobrepõem às vibrações proteicas do leite, tornando-se mais perceptíveis após a derivação.

Na região de $1300\text{--}900\text{ cm}^{-1}$, atribuída a vibrações C-F e C-N, (anel quinolônico do antibiótico) (BRUCKMANN et al., 2024), C-O e C-O-C (relacionado a lactose da matriz láctea) (SANTANA et al., 2025), a primeira derivada destacou sobreposições entre as classes 0, 1 e 2, sugerindo que o aumento da concentração de Ciprofloxacina influencia a resposta espectral nessas bandas. Entretanto, observa-se também um aumento do ruído, especialmente em regiões de menor intensidade espectral ($2500\text{--}2000\text{ cm}^{-1}$), o que é um efeito esperado da derivação (ROSALES-MARTÍNEZ et al., 2025).

Os mesmos procedimentos foram aplicados aos espectros NIR e podem ser observados na Figura 2.

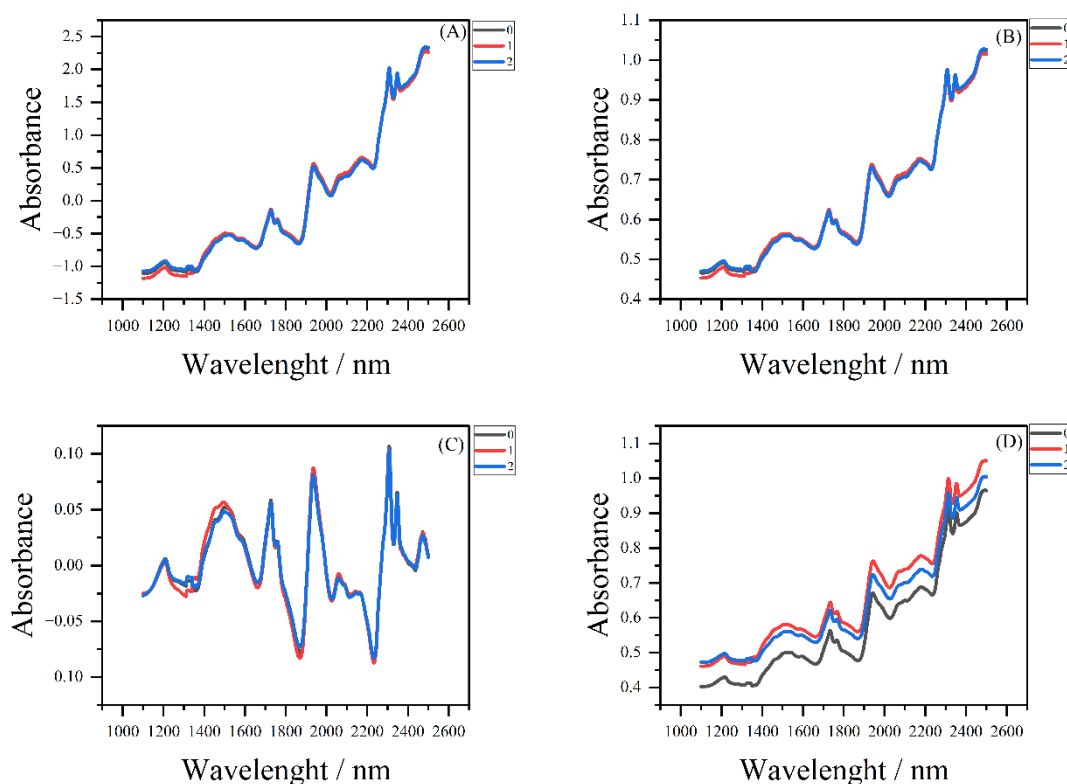


Figura 2. Espectro NIR processados por SNV (A), MSC (B), correção de linha de base (C) e DSG (D).

Os espectros NIR (Figura 2) das amostras de leite liofilizado puro (classe 0) e fortificadas com Ciprofloxacina (classes 1 e 2) apresentaram perfis típicos de matriz láctea, com bandas predominantes associadas à O–H (~1450 e ~1800 nm) de lactose, C–H de lipídios (~1720–1760 nm) e combinações envolvendo C–H, O–H e N–H (~2100–2300 nm) de proteínas, carboidratos e lipídios (MOHAMED et al., 2021). Devido à baixa concentração do fármaco, as diferenças entre as classes nos espectros brutos foram sutis e mascaradas pela forte contribuição da matriz.

A correção de linha de base (Figura 2C) reduziu deslocamentos verticais e melhorou o alinhamento espectral, porém não promoveu separação evidente entre as classes. As variações permaneceram discretas, principalmente na região de 2100–2300 nm, onde há possível contribuição dos grupos carbonílicos e amínicos da Ciprofloxacina (QIU et al., 2018; YAN, 2025). Os pré-processamentos MSC (Figura 2B) e SNV (Figura 2A) minimizaram os efeitos de espalhamento típicos de amostras sólidas, reduzindo a variabilidade física entre espectros. Após esses tratamentos, as diferenças químicas tornaram-se relativamente mais evidentes, especialmente nas regiões possivelmente associadas a C=O, C–N e N–H (~2100–2300 nm), sugerindo influência da concentração crescente do fármaco (classe 2 > classe 1 > classe 0) (QIU et al., 2018; YAN, 2025; ROSALES-MARTÍNEZ et al., 2025).

A aplicação da primeira derivada Savitzky–Golay (Figura 2D) intensificou as diferenças sutis entre as classes, realçando possíveis variações locais associadas principalmente às regiões de ~1600–1700 nm (C=C aromático e N–H da Ciprofloxacina) e ~2100–2300 nm (QIU et al., 2018; SCHUMER et al., 2025; ROSALES-MARTÍNEZ et al., 2025). Entretanto, ainda se observa leve tendência global no espectro derivado, indicando que a derivada isoladamente não eliminou completamente efeitos residuais de linha de base.

3.2. Análises quimiométricas

3.2.1. Análise de Componentes Principais

A análise exploratória por PCA foi aplicada aos espectros MIR (4000–650 cm^{-1}) e NIR (2500–1000 nm) do leite liofilizado fortificado com diferentes concentrações de Ciprofloxacina, com o objetivo de avaliar o potencial discriminante entre as classes sem antibiótico (classe 0), até o LMR (classe 1) e acima do LMR (classe 2). Foram testados diferentes pré-processamentos espectrais, em que a primeira derivada Savitzky–Golay apresentou o melhor desempenho, por realçar variações químicas sutis e reduzir efeitos de espalhamento. No caso do MIR, o recorte espectral entre 1800 e 800 cm^{-1} mostrou-se o mais adequado, por abranger as principais bandas características da Ciprofloxacina e minimizar a sobreposição das classes, resultando em um modelo mais estável e interpretável.

As duas primeiras componentes principais (PC1 e PC2) explicaram conjuntamente mais de 98% da variância total nos espectros MIR (Figura 3) e acima de 99% nos espectros NIR (Figura 4), evidenciando boa representatividade dos dados.

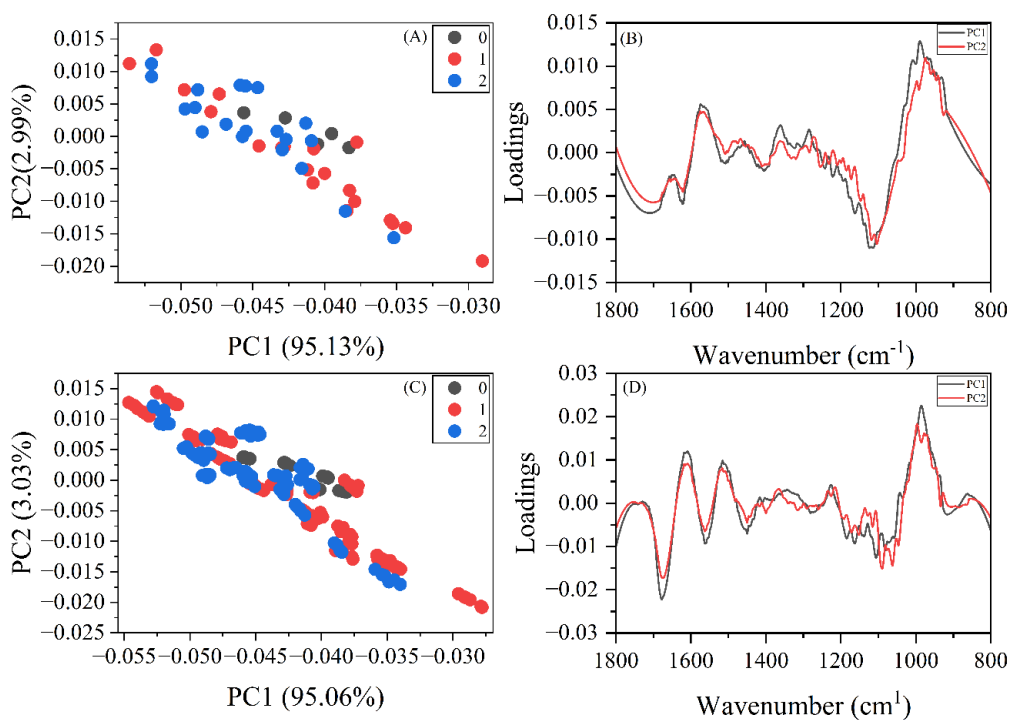


Figura 3. PCA dos espectros MIR (A, C) gráficos de *scores* das amostras originais e dos dados combinados, respectivamente e (B, D) *loadings* correspondentes às mesmas análises.

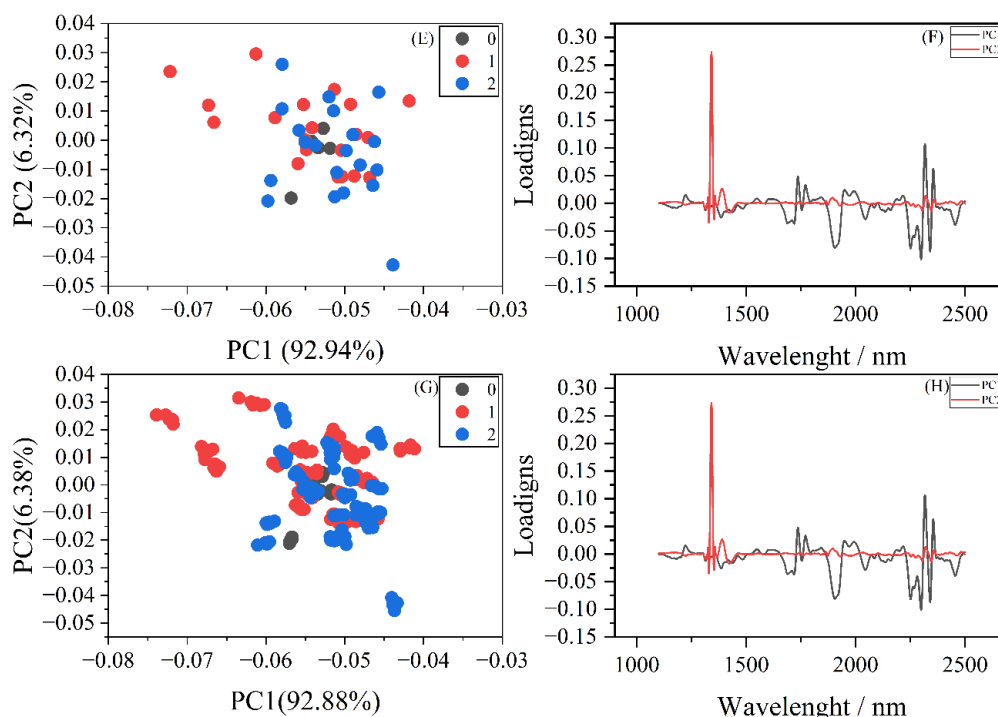


Figura 4. PCA dos espectros NIR (E, G) gráficos de *scores* das amostras originais e dos dados combinados, respectivamente e (F, H) *loadings* correspondentes às mesmas análises.

Em ambos os conjuntos, observa-se acentuada sobreposição entre as classes no espaço de *scores*, indicando que as principais fontes de variação estão associadas à composição natural do leite e não diretamente à concentração de Ciprofloxacina. Os *loadings* mostram, no MIR (Figura 3), as maiores contribuições da PC1 concentram-se nas regiões em torno de $\sim 1650\text{ cm}^{-1}$ (amida I), $\sim 1540\text{ cm}^{-1}$ (amida II) e $\sim 1740\text{ cm}^{-1}$ (C=O de lipídios), bandas típicas da matriz proteica e lipídica do leite, mas também presentes no espectro do antibiótico (CHATZIMICHAIL et al., 2025). Apesar da sobreposição entre as classes, as regiões dos *loadings* entre $\sim 1300\text{--}1200\text{ cm}^{-1}$ e $\sim 1450\text{ cm}^{-1}$ sugerem contribuição espectral atribuída à Ciprofloxacina, principalmente na PC2, associadas às vibrações C–F, C–N, C=O e modos de deformação do anel quinolônico (BRUCKMANN et al., 2024). Essas bandas indicam que, embora a variação global seja dominada pela matriz, há indícios de que o antibiótico exerce influência mensurável sobre a estrutura multivariada dos espectros, refletindo sua presença nos dados e contribuindo parcialmente para a variância capturada pelas componentes principais.

Os *loadings* referentes aos espectros NIR (Figura 4) indicam que a PC1 apresenta contribuições marcantes próximas de $\sim 2300\text{ nm}$ e $\sim 1700\text{ nm}$, regiões tipicamente associadas a combinações e sobretons envolvendo ligações C–H, N–H e modos de combinação de O–H característicos de proteínas, carboidratos e lipídeos, o que corrobora a predominância da composição do leite na explicação da variância (BEĆ, GRABSKA e HUCK, 2020). Já a PC2

apresenta picos notáveis entre ~1300 nm e ~1400 nm, região onde aparecem sobretons e combinações relacionadas a grupos aromáticos (C–H, N–H, C–N), compatíveis com modos vibracionais esperados para estruturas como as da Ciprofloxacina (TAN et al., 2012). Assim, apesar da sobreposição das classes, os *loadings* indicam que algumas regiões NIR carregam sinal atribuível ao antibiótico e influenciaram parcialmente as componentes principais.

Com a inclusão das amostras sintéticas, observa-se apenas um aumento na densidade de pontos no espaço das componentes principais, sem alteração significativa na distribuição das classes. No entanto, a proximidade entre as amostras originais e sintéticas sugere que o processo de geração via PCA foi estatisticamente consistente, reproduzindo adequadamente a estrutura multivariada dos dados e preservando suas relações espectrais. Ainda assim, a separação entre classes não se aprimorou, indicando que a variabilidade introduzida não foi suficiente para capturar diferenças espectrais relacionadas à concentração de Ciprofloxacina.

Em síntese, a aplicação da primeira derivada Savitzky–Golay permitiu melhorar a qualidade espectral e explorar a coerência estrutural dos dados, mas não foi suficiente para promover a discriminação entre as classes. Esses resultados, observados de forma semelhante para as regiões MIR e NIR, demonstram que as diferenças espectrais associadas à Ciprofloxacina estão abaixo do limite de sensibilidade da PCA, reforçando a necessidade de empregar métodos supervisionados, como PLS-DA, para uma separação mais robusta e interpretável.

3.2.2. Análise de Mínimos Quadrados Parciais - PLS-DA

A análise discriminante por mínimos quadrados parciais (PLS-DA) foi aplicada com o objetivo de classificar amostras de leite em pó liofilizado em três categorias distintas: sem antibiótico (classe 0), baixa concentração de Ciprofloxacina (25 a 100 $\mu\text{g L}^{-1}$, classe 1) e alta concentração de Ciprofloxacina (125 a 200 $\mu\text{g L}^{-1}$, classe 2). Diferentes estratégias de pré-processamento espectral foram avaliadas, considerando cenários com e sem a inclusão de amostras sintéticas, bem como a variação do número de variáveis latentes (LVs) (DUARTE et al., 2022; THEAKSTONE et al., 2021; HIRRI et al., 2016). A Tabela 1 apresenta os resultados obtidos para os modelos PLS-DA em espectros MIR e NIR.

Tabela 1. Parâmetros de desempenho para os modelos PLS-DA usando espectroscopias MIR e NIR.

Tratamento	Modelo	Dados	Calibração						Validação			
			LV	Classe	Exatidão	Precisão	Sens.	Especif.	Exatidão	Precisão	Sens.	Especif.
DSG	MIR	Originais	3	0	0,963	1	0,667	1	1	1	1	1
				1	0,963	0,923	1	0,933	1	1	1	1
				2	1	1	1	1	1	1	1	1
		DC	4	0	0,994	1	0,944	1	0,991	1	0,917	1
				1	0,994	0,986	1	0,989	0,991	0,979	1	0,983
				2	1	1	1	1	1	1	1	1
	NIR	Originais	5	0	0,963	0,75	1	0,958	0,944	0,667	1	0,937
				1	0,963	1	0,916	1	0,889	1	0,750	1
				2	1	1	1	1	0,833	0,778	0,875	0,80
		DC	4	0	0,876	0,46875	0,833	0,881	0,842	0,413	1	0,823
				1	0,827	0,893	0,694	0,933	0,759	0,893	0,521	0,950
				2	0,802	0,770	0,792	0,811	0,824	0,784	0,833	0,817
SNV	MIR	Originais	3	0	1	1	1	1	0,889	-	0	1
				1	1	1	1	1	0,611	0,538	0,875	0,40
				2	1	1	1	1	0,611	0,6	0,375	0,8

Tratamento	Modelo	Dados	LV	Classe	Calibração				Validação			
					Exatidão	Precisão	Sens.	Especif.	Exatidão	Precisão	Sens.	Especif.
MSC	DC	Originais	4	1	0,667	0,615	0,667	0,667	0,667	0,667	0,50	0,80
				2	0,741	0,727	0,667	0,80	0,555	0,50	0,625	0,50
				0	0,858	0,419	0,722	0,875	0,805	0,304	0,583	0,833
				1	0,802	0,794	0,750	0,844	0,750	0,744	0,667	0,817
				2	0,759	0,761	0,667	0,833	0,704	0,690	0,604	0,783
				0	0,926	1	0,333	1	0,889	-	0	1
				1	0,926	0,857	1	0,867	0,611	0,538	0,875	0,40
				2	1	1	1	1	0,611	0,60	0,375	0,80
	DC	Originais	3	0	0,914	0,7	0,388	0,979	0,926	0,70	0,583	0,969
				1	0,969	0,935	1	0,944	0,935	0,872	1	0,883
				2	0,883	0,853	0,889	0,878	0,861	0,884	0,792	0,912
				0	0,889	-	0	1	0,833	0	0	0,937
	NIR	Originais	4	1	0,778	0,714	0,833	0,733	0,611	0,545	0,750	0,5
				2	0,741	0,692	0,75	0,733	0,555	0,5	0,375	0,7
				0	0,833	0,378	0,778	0,840	0,851	0,409	0,750	0,864
				1	0,753	0,767	0,639	0,844	0,741	0,727	0,667	0,80

Tratamento	Modelo	Dados	Calibração				Validação					
			LV	Classe	Exatidão	Precisão	Sens.	Especif.	Exatidão	Precisão	Sens.	Especif.
			2		0,747	0,734	0,667	0,811	0,703	0,690	0,604	0,783

DSG: primeira derivada Savitzky–Golay; DC: dados combinados; Sens: Sensibilidade; Especif: Especificidade.

Os resultados demonstraram que a espectroscopia MIR aliada ao pré-processamento primeira derivada Savitzky–Golay, à inclusão de amostras sintéticas e ao uso de quatro variáveis latentes, apresentou o melhor desempenho, conciliando alta exatidão, precisão, sensibilidade e especificidade tanto nas etapas de calibração quanto de validação (Tabela 1). Essa configuração alcançou uma exatidão média de 0,994 no conjunto de calibração e 0,991 no conjunto de validação, com valores de sensibilidade e precisão superiores a 0,9 em todas as classes, evidenciando a robustez do modelo. Esses resultados corroboram os resultados reportados por Hu et al. (2021), que obtiveram desempenho semelhante na diferenciação de três antibióticos em leite em pó, utilizando PLS-DA.

O uso da primeira derivada Savitzky–Golay mostrou-se particularmente eficiente por eliminar variações de linha de base e normalizar diferenças de intensidade (HIRRI et al., 2016), preservando, contudo, as características espectrais relacionadas à concentração do antibiótico. Esse comportamento favoreceu a separação entre classes adjacentes (0–1 e 1–2), o que é crucial em análises onde pequenas variações de concentração possuem relevância analítica. Aliada ao pré-processamento a inclusão de amostras sintéticas contribuiu de maneira expressiva para o aumento da capacidade classificatória do modelo (CALDEIRA et al., 2024), principalmente nas classes 0 (sem antibiótico) e 1 (baixa concentração), que apresentavam desempenhos insatisfatórios no conjunto original.

A Figura 5 apresenta os gráficos de valores preditos obtidos a partir do modelo PLS-DA MIR, associado ao pré-processamento por primeira derivada de Savitzky–Golay e ao uso de dados combinados, para a classificação de amostras de leite em pó liofilizado quanto à presença do padrão de Ciprofloxacina. As classes avaliadas foram leite puro (classe 0), baixa concentração de Ciprofloxacina ($\leq 100 \mu\text{g L}^{-1}$, correspondente ao limite máximo de resíduo – LMR) (UNIÃO EUROPEIA, 2010) e alta concentração de Ciprofloxacina (> 100 até $200 \mu\text{g L}^{-1}$).

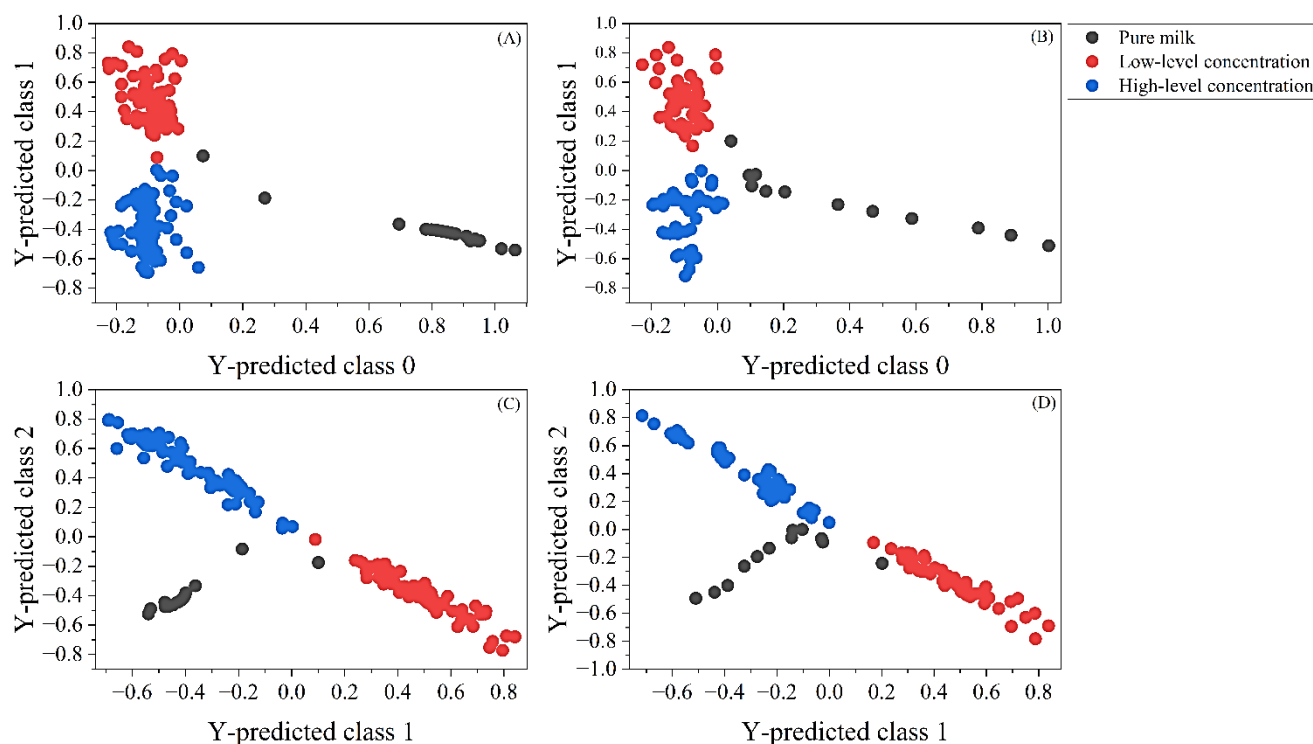


Figura 5. Gráficos de dispersão dos valores Y preditos pelo modelo PLS-DA para as diferentes classes; A e C correspondem ao conjunto de calibração, enquanto B e D referem-se ao conjunto de validação externa.

Nos gráficos, observa-se uma clara separação entre o leite puro (classe 0) e as amostras fortificadas, evidenciando a capacidade do modelo em discriminar amostras conformes e não conformes do ponto de vista regulatório. As amostras classificadas como baixa concentração (até o LMR) apresentam valores preditos intermediários, enquanto as amostras com alta concentração de Ciprofloxacina exibem respostas bem distintas, afastadas do grupo do leite puro. De acordo com Gaugain et al. (2019), essa separação é fundamental para aplicações em controle de qualidade, uma vez que permite identificar de forma confiável amostras que excedem o limite permitido pela legislação.

Do ponto de vista do controle de qualidade do leite, os resultados indicam que a abordagem baseada em PLS-DA aplicada a dados espectrais é promissora como ferramenta de triagem rápida. O método permite identificar amostras livres de Ciprofloxacina, amostras dentro do LMR e, principalmente, amostras acima do limite máximo permitido, que representam risco à saúde pública e devem ser rejeitadas (HU et al., 2021). Assim, o modelo apresenta potencial para ser utilizado como método complementar ou de pré-seleção em programas de monitoramento de resíduos de antibióticos em leite.

A separação observada nos gráficos de valores preditos do modelo PLS-DA é confirmada quantitativamente pelas matrizes de confusão apresentadas na Tabela 2, referentes aos conjuntos de calibração e validação.

Tabela 2. Matriz de confusão para amostras de calibração e validação com base no modelo PLS-DA, considerando dados MIR.

	Calibração			Validação		
Classe Real	Classe 0	Classe 1	Classe 2	Classe 0	Classe 1	Classe 2
	(%)	(%)	(%)	(%)	(%)	(%)
Classe 0	94,4	6,6	0,0	91,7	8,3	0,0
Classe 1	0,0	100	0,0	0,0	100	0,0
Classe 2	0,0	0,0	100	0,0	0,0	100

As matrizes de confusão obtidas para os conjuntos de calibração e validação externa evidenciam o elevado desempenho do modelo PLS-DA na classificação das amostras de leite em pó liofilizado quanto à presença de Ciprofloxacina. Observa-se que a maioria das amostras foi corretamente classificada ao longo da diagonal principal, com taxas de erro reduzidas e concentradas principalmente entre as classes leite puro e baixa concentração, próximas ao limite máximo de resíduo (LMR). Destaca-se que nenhuma amostra com concentração acima do LMR foi classificada como leite puro, o que demonstra o caráter conservador e a adequação do modelo para aplicações em controle de qualidade e segurança alimentar.

Nesta mesma abordagem, a Figura 6 apresenta os gráficos de *scores* (A) e *loadings* (B) do modelo PLS-DA MIR, pré-processado com a primeira derivada de Savitzky–Golay e com o uso de dados combinados. O modelo foi construído com quatro variáveis latentes (LVs), as quais fornecem uma interpretação multivariada consistente e coerente com os resultados de classificação previamente discutidos. A utilização de quatro LVs mostrou-se adequada por permitir a captura da maior parte da variância relevante dos dados espectrais, assegurando simultaneamente boa capacidade discriminante entre as classes e robustez do modelo, conforme evidenciado pelos elevados valores de acurácia, sensibilidade e especificidade obtidos tanto na calibração quanto na validação externa.

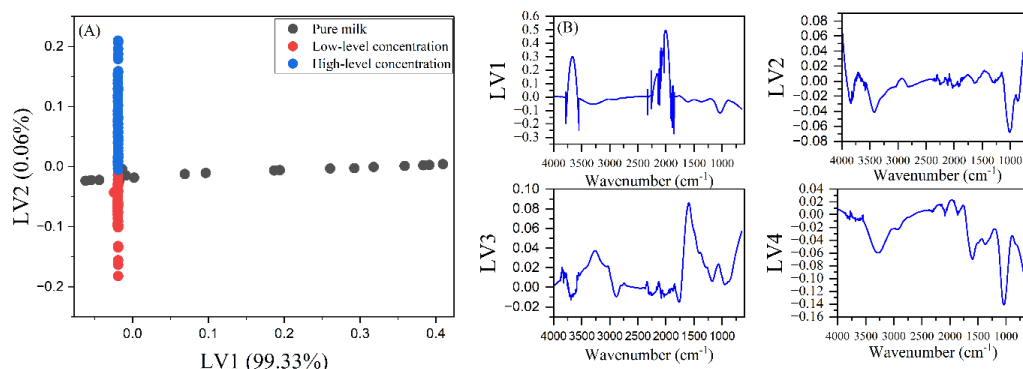


Figura 6. Gráficos de *scores* (A) e *loadings* (B) do modelo PLS-DA.

O gráfico de *scores* ($LV1 \times LV2$) evidencia uma separação clara entre as amostras de leite puro, baixa concentração e alta concentração de Ciprofloxacina. Observa-se que a LV1, responsável por aproximadamente 99,33% da variância explicada, exerce papel dominante na discriminação entre o leite sem antibiótico e as amostras fortificadas. Essa separação indica que a principal fonte de variação capturada pelo modelo está associada às características intrínsecas da matriz do leite.

A análise dos *loadings* da LV1 confirma essa interpretação, uma vez que são observadas contribuições espectrais significativas nas regiões características dos constituintes majoritários do leite. Destacam-se bandas associadas às vibrações de estiramento O–H e N–H na região de aproximadamente 3600–3200 cm^{-1} , relacionadas principalmente à água e às proteínas, bem como bandas de estiramento C–H de grupos alifáticos em torno de 2950–2850 cm^{-1} , típicas de lipídios (BALAN et al., 2020). Além disso, contribuições nas regiões de 1700–1500 cm^{-1} , atribuídas às bandas amida I e amida II das proteínas, e na região de 1200–900 cm^{-1} , associada a carboidratos (CHATZIMICHAIL et al., 2025), reforçam que a LV1 descreve predominantemente a variabilidade estrutural da matriz láctea.

Em contraste, as variáveis latentes subsequentes (LV2, LV3 e LV4), embora expliquem frações menores da variância total, estão diretamente relacionadas às variações espectrais associadas à presença da Ciprofloxacina. Os *loadings* dessas LVs apresentam contribuições relevantes em regiões espectrais compatíveis com o padrão do antibiótico, indicando que essas variáveis latentes capturam diferenças químicas mais sutis relacionadas à sua incorporação na matriz do leite.

Especificamente, observam-se contribuições em regiões associadas a vibrações de grupos funcionais característicos da Ciprofloxacina, como bandas na região de 1700–1600 cm^{-1} , atribuídas a vibrações de estiramento C=O e C=C do anel quinolônico, além de contribuições entre 1500–1400 cm^{-1} relacionadas a deformações N–H e vibrações aromáticas (ALQARNI et

al., 2025). Regiões adicionais entre 1300–1000 cm^{-1} , associadas a estiramentos C–N e C–F (BRUCKMANN et al., 2024), também se destacam nos *loadings* das LVs superiores, reforçando a associação dessas variáveis latentes com o perfil espectral do antibiótico.

A contribuição conjunta das LVs 2, 3 e 4 é fundamental para a discriminação entre as amostras de baixa e alta concentração de Ciprofloxacina, especialmente em torno do limite máximo de resíduo (LMR). Embora essas LVs expliquem parcelas menores da variância global, sua inclusão melhora a capacidade do modelo em capturar variações químicas sutis relacionadas à concentração do analito, refletindo-se no bom desempenho observado no conjunto de validação externa.

De forma geral, a coerência entre os gráficos de *scores*, os perfis de *loadings*, a atribuição química das regiões espectrais e os elevados indicadores de desempenho confirma que o modelo PLS-DA com quatro variáveis latentes é adequado para a classificação de amostras de leite em pó liofilizado quanto à presença e ao nível de Ciprofloxacina. Esses resultados corroboram a discussão anterior e evidenciam que a estratégia de pré-processamento adotada, aliada ao uso de dados combinados, foi eficaz para realçar tanto as características da matriz do leite quanto as assinaturas espectrais do padrão do antibiótico.

3.2.2.1. Avaliação da robustez do modelo PLS-DA por validação Monte Carlo

Com o objetivo de avaliar a robustez estatística do modelo selecionado, foi realizada uma validação externa por simulação Monte Carlo com 50 iterações, aplicada ao modelo PLS-DA – MIR com pré-processamento por DSG e dados combinados com quatro variáveis latentes. A Tabela 3, apresenta a média e o desvio padrão das métricas avaliadas no modelo.

Tabela 3. Desempenho do modelo PLS-DA-MIR-DSG com dados combinados avaliado por simulação de Monte Carlo (média $\pm$ DP).

Calibração				
Classe	Exatidão	Precisão	Sensibilidade	Especificidade
0	1.0 $\pm$ 0.0	0.982 $\pm$ 0.027	1.0 $\pm$ 0.0	0.998 $\pm$ 0.004
1	0.995 $\pm$ 0.007	0.998 $\pm$ 0.006	0.995 $\pm$ 0.007	0.998 $\pm$ 0.005
2	0.998 $\pm$ 0.006	0.999 $\pm$ 0.003	0.998 $\pm$ 0.006	1.0 $\pm$ 0.002
Validação				
Classe	Exatidão	Precisão	Sensibilidade	Especificidade
0	0.988 $\pm$ 0.029	0.969 $\pm$ 0.038	0.988 $\pm$ 0.029	0.996 $\pm$ 0.005
1	0.989 $\pm$ 0.013	0.992 $\pm$ 0.012	0.989 $\pm$ 0.013	0.994 $\pm$ 0.009
2	0.995 $\pm$ 0.010	0.998 $\pm$ 0.008	0.995 $\pm$ 0.010	0.998 $\pm$ 0.006

Os resultados obtidos a partir de 50 iterações de validação Monte Carlo demonstraram elevada robustez e estabilidade do modelo PLS-DA-MIR-DSG com dados combinados. Na calibração, as métricas apresentaram valores próximos de 1 (exatidão entre 0,995 e 1,000), com desvios padrão muito baixos ($\leq 0,007$), indicando excelente capacidade discriminante e baixa variabilidade entre as reamostragens. Na validação, os resultados permaneceram igualmente elevados (exatidão entre 0,988 e 0,995), com pequenas variações ($DP \leq 0,029$), evidenciando boa generalização do modelo para dados não utilizados no treinamento. As altas sensibilidades e especificidades em todas as classes confirmam desempenho equilibrado, sem viés para nenhuma categoria específica. De forma geral, os baixos desvios padrão associados às métricas médias indicam estabilidade estatística e consistência preditiva, reforçando a confiabilidade do modelo sob diferentes divisões aleatórias dos dados.

4. Conclusão

Neste estudo, demonstramos que a espectroscopia no infravermelho médio, combinada com pré-processamento por primeira derivada de Savitzky–Golay e modelagem por PLS-DA, é capaz de classificar amostras de leite em pó liofilizado quanto à presença e níveis de Ciprofloxacina com alto desempenho nas condições experimentais testadas. A inclusão de amostras sintéticas geradas via PCA aumentou a densidade amostral e melhorou a capacidade classificatória, especialmente na distinção entre amostras livres de antibiótico e aquelas próximas ao limite máximo de resíduo (LMR). Em suma, a abordagem espectroscópica-quimiométrica proposta constitui uma ferramenta potencialmente eficiente para triagem inicial de resíduos de Ciprofloxacina em leite em pó, complementando métodos confirmatórios e contribuindo para programas de monitoramento de segurança alimentar.

5. Referencias

- ALIBEKOV, R. S., AITBAYEVA, A. Z., KHAMITOVA, B. M., AZIMOV, A. M., SULAIMAN, A., MUSTAPA KAMAL, S. M., TAIP, F. S. Effects of maltodextrin in freeze drying on the physical and functional properties of different type of milk powder. **Cogent Food & Agriculture**, v. 11, n. 1, p. 2473540, 2025.
- ALQARNI, L. S., ALHARBI, A. A., AISSA, M. A. B., MODWI, A. Fabrication of green ternary $CeO_2@Co_3O_4@g-C_3N_4$ nanocomposite as an efficient and durable sorbent for the ciprofloxacin removal. **Journal of the Indian Chemical Society**, p. 102129, 2025.

ALVES, J. F., PAULA, G. H., SILVA, R. C., LEÃO, P. V., LEÃO, K. M., NICOLAU, E. S., SILVA, M. A. Residues of antibiotics in milk: persistence and quality interference. **Canadian Journal of Animal Science**, v. 100, n. 1, p. 93-101, 2019.

AMIRVARESI, A., NIKOUNEZHAD, N., AMIRAHMADI, M., DARAEI, B., PARASTAR, H. Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection. **Food Chemistry**, v. 344, p. 128647, 2021.

BALAN, B., DHAULANIYA, A. S., JAMWAL, R., SODHI, K. K., KELLY, S., CANNAVAN, A., SINGH, D. K. Application of Attenuated Total Reflectance-Fourier Transform Infrared (ATR-FTIR) spectroscopy coupled with chemometrics for detection and quantification of formalin in cow milk. **Vibrational Spectroscopy**, v. 107, p. 103033, 2020.

BEĆ, K. B., GRABSKA, J., HUCK, C. W. Near-infrared spectroscopy in bio-applications. **Molecules**, v. 25, n. 12, p. 2948, 2020.

BION, C., BECK HENZELIN, A., QU, Y., PIZZOCRI, G., BOLZONI, G., BUFFOLI, E. Analysis of 27 antibiotic residues in raw cow's milk and milk-based products—validation of Delvotest® T. **Food Additives & Contaminants: Part A**, v. 33, n. 1, p. 54-59, 2016.

BRITO, I. C. P., DA SILVA SANTOS, H. L., SANTOS, K. E. O., DE ALBUQUERQUE FERNANDES, S. A. Modification and validation of miniaturized QuEChERS method for multi-residual determination of antibiotics in milk from a tropical region. **Microchemical Journal**, v. 207, p. 112025, 2024.

BRUCKMANN, F. S., FUHR, A. C. F. P., ZIBETTI, L., BENDER, C. R., SILVA, L. F. O MARTINELLO, K. B., DOTTO, G. L. Adsorption of ciprofloxacin from aqueous solution and fresh synthetic urine by graphene oxide: Conventional and statistical physics modeling approaches. **Chemical Engineering Journal**, v. 487, p. 150484, 2024.

CALDEIRA, G. R., COSTA, T. O., NASCIMENTO, M. H., CORRADINI, P. G., FILGUEIRAS, P. R., FERREIRA, D. C., SOUZA, M. D. O. Authentication of Specialty Coffees from the Fluminense Northwest and Caparaó Regions (Brazil) Using UV-Vis Spectroscopy and Synthetic Samples Partial Least Square Discriminant Analysis (SS-PLS-DA). **Journal of the Brazilian Chemical Society**, v. 35, p. e-20230144, 2024.

CALDEIRA, G. R., COSTA, T. O., NASCIMENTO, M. H., CORRADINI, P. G., FILGUEIRAS, P. R., FERREIRA, D. C., SOUZA, M. D. O. Authentication of Specialty Coffees from the Fluminense Northwest and Caparaó Regions (Brazil) Using UV-Vis

Spectroscopy and Synthetic Samples Partial Least Square Discriminant Analysis (SS-PLS-DA). **Journal of the Brazilian Chemical Society**, v. 35, p. e-20230144, 2024.

CHATZIMICHAIL, K., LADIKA, G., CHRISTODOULOU, P., BARTZIS, V., KONTELES, S. J., LAZOU, A. E., SINANOGLU, V. J. Comparative Evaluation of Cow and Goat Milk Samples Utilizing Non-Destructive Techniques and Chemometric Approaches. **Applied Sciences**, v. 15, n. 20, p. 10883, 2025.

CHATZIMICHAIL, K., LADIKA, G., CHRISTODOULOU, P., BARTZIS, V., KONTELES, S. J., LAZOU, A. E., SINANOGLU, V. J. Comparative Evaluation of Cow and Goat Milk Samples Utilizing Non-Destructive Techniques and Chemometric Approaches. **Applied Sciences**, v. 15, n. 20, p. 10883, 2025.

DANILOSKI, D., MCCARTHY, N. A., O'CALLAGHAN, T. F., VASILJEVIC, T. Authentication of β -casein milk phenotypes using FTIR spectroscopy. **International Dairy Journal**, v. 129, p. 105350, 2022.

DONEVA, M. D., DYANKOVA, S. M., MITEVA, D. P., NACHEVA, I. B., METODIEVA, P. M. Cryobiological studies and freeze drying of cow's milk and curd. **J. Chem. Technol. Meta**, v. 56, p. 932-937, 2021.

DUARTE, J. M., SALES, N. G. S., BRAGA, J. W. B., BRIDGE, C., MARIC, M., SOUSA, M. H., & DE ANDRADE GOMES, J. Discrimination of white automotive paint samples using ATR-FTIR and PLS-DA for forensic purposes. **Talanta**, v. 240, p. 123154, 2022.

DUREL, Luc; GALLINA, Guglielmo; PELLET, Terence. Assessment of ceftiofur residues in cow milk using commercial screening test kits. **Veterinary Record Open**, v. 6, n. 1, p. e000329, 2019.

EUROPEAN UNION. *Commission Regulation (EU) No 37/2010 of 22 December 2009 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin.* Official Journal of the European Union, L 15, p. 1–72, 20 Jan. 2010. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010R0037>. Accessed on: 4 June 2025.

FENG, Y., DONG, F., CHEN, Y., LI, H., RODAS-GONZÁLEZ, A., WANG, S. Combining Vis-NIR and NIR hyperspectral imaging techniques with a data fusion strategy for prediction of norfloxacin residues in mutton. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 322, p. 124844, 2024.

FLANAGAN, A. R., GLAVIN, F. G. A comparative analysis of data synthesis techniques to improve classification accuracy of raman spectroscopy data. **Journal of Chemical Information and Modeling**, v. 64, n. 7, p. 2311-2322, 2023.

FOLLI, G. S., NASCIMENTO, M. H., LOVATTI, B. P., ROMÃO, W., FILGUEIRAS, P. R. A generation of synthetic samples and artificial outliers via principal component analysis and evaluation of predictive capability in binary classification models. **Chemometrics and Intelligent Laboratory Systems**, v. 251, p. 105154, 2024.

GAUGAIN, M., MOMPELAT, S., FOURMOND, M. P., MANCEAU, J., ROLLAND, J. G., LAURENTIE, M., HURTAUD-PESSEL, D. A non-targeted LC-HRMS approach for detecting exposure to illegal veterinary treatments: The case of cephalosporins in commercial laying Hens. **Journal of Chromatography A**, v. 1599, p. 161-171, 2019.

HASOŇOVÁ, L., SAMKOVA, E., BALDIKOVA, E., REINDL, K., NEJESCHLEBOVA, H., HONESOVÁ, S. J., HANUŠ, O. Occurrence and seasonal variation of aflatoxin M1 in raw cow milk in the Czech Republic, Central Europe. **Food Control**, v. 170, p. 111028, 2025.

HIRRI, A., BASSBASI, M., PLATIKANOV, S., TAULER, R., OUSSAMA, A. FTIR spectroscopy and PLS-DA classification and prediction of four commercial grade virgin olive oils from Morocco. **Food Analytical Methods**, v. 9, n. 4, p. 974-981, 2016.

HU, J., XU, Z., LI, M., HE, Y., LIU, Y. Discriminant analysis and quantitative study of antibiotics in infant milk powder based on hyperspectral detection. **Vibrational Spectroscopy**, v. 114, p. 103244, 2021.

JAGESAR, D. C., KERKHOF, J. Validation of Delvotest® Fast BT for Detection of β -Lactams and Tetracyclines in Raw Cow's Milk: Performance Tested Method SM 102104. **Journal of AOAC International**, v. 106, n. 1, p. 88-98, 2023.

LUIZ, L. D. C., BELL, M. J. V., ROCHA, R. A. D., LEAL, N. L., ANJOS, V. D. C. D. Detection of veterinary antimicrobial residues in milk through near-infrared absorption spectroscopy. **Journal of Spectroscopy**, v. 2018, n. 1, p. 5152832, 2018.

MAHMOOD, T., ABBAS, M., ILYAS, S., AFZAL, N., NAWAZ, R. Quantification of fluoroquinolone (enrofloxacin, norfloxacin and ciprofloxacin) residues in cow milk. **Ijcsbs**, v. 10, p. 10-15, 2016.

MOGHADDAM, H. N., TAMIJI, Z., LAKEH, M. A., KHOSHAYAND, M. R., MAHMOODI, M. H. Multivariate analysis of food fraud: A review of NIR based instruments in tandem with chemometrics. **Journal of Food Composition and Analysis**, v. 107, p. 104343, 2022.

MOHAMED, H., NAGY, P., AGBABA, J., & KAMAL-ELDIN, A. Use of near and mid infrared spectroscopy for analysis of protein, fat, lactose and total solids in raw cow and camel milk. **Food Chemistry**, v. 334, p. 127436, 2021.

NÍ FHUARÁIN, A. M., O'DONNELL, C. P., LUO, J., GOWEN, A. A Review on MIR, NIR, Fluorescence and Raman Spectroscopy Combined with Chemometric Modeling to Predict the Functional Properties of Raw Bovine Milk. **ACS Food Science & Technology**, v. 4, n. 10, p. 2258-2271, 2024.

PADINJAREKKARA, S. K., ALECCI, J., POPA, M. XPCA Gen: Extended PCA Based Tabular Data Generation Model with Regularisation. In: **International Conference on Pattern Recognition Applications and Methods**. Cham: Springer Nature Switzerland, 2024. p. 46-62.

QIN, J., XIE, L., YING, Y. Rapid determination of tetracyclines hydrochloride using ATR FT-MIR spectroscopy. **Food Analytical Methods**, v. 9, n. 10, p. 2880-2886, 2016.

QIU, G., LÜ, E., LU, H., XU, S., ZENG, F., SHUI, Q. Single-kernel FT-NIR spectroscopy for detecting supersweet corn (*Zea mays* L. *saccharata* sturt) seed viability with multivariate data analysis. **Sensors**, v. 18, n. 4, p. 1010, 2018.

RATHORE, H. S., SENTHILVELAN, T., VASANTHARAJA, R., ABRAHAM, L. S., PRAKASH, D., SIVAGNANAM, U. T., GUPTA, S. Fabrication and characterization of chitosan film impregnated ciprofloxacin drug: A comparative study. **Biocatalysis and Agricultural Biotechnology**, v. 18, p. 101078, 2019.

RINGSTED, T., SIESLER, H. W., ENGELSEN, S. B. Monitoring the staling of wheat bread using 2D MIR-NIR correlation spectroscopy. **Journal of Cereal Science**, v. 75, p. 92-99, 2017.

ROSALES-MARTÍNEZ, O., GRANDA-GUTIÉRREZ, E. E., GARCÍA-HERNÁNDEZ, R. A., ALEJO-ELEUTERIO, R., FLORES-FUENTES, A. A. Spectral Derivatives Improve FTIR-Based Machine Learning Classification of Plastic Polymers. **Modelling**, v. 6, n. 4, p. 115, 2025.

RUSSO, R. E., FATTOBENE, M., ZAMPONI, S., CONTI, P., HERRERO, A., BERRETTONI, M. Exploring Environmental Element Monitoring Data Using Chemometric Techniques: A Practical Case Study from the Tremiti Islands (Italy). **Molecules**, v. 31, n. 2, p. 232, 2026.

- SANTANA, V. R., MAGALHÃES, B. E. A. D., MINHO, L. A. C., FREITAS, A. G. M. D., SOUZA, R. D., BATISTA, A. S., FERNANDES, S. A. D. A. Random forest and infrared spectroscopy to detect contamination of freeze-dried milk by ivermectin residues. **Ciência Rural**, v. 55, n. 12, p. e20240536, 2025.
- SCHONS, P. F., COLET, R., PINTO, S. S., VERRUCK, S. Powdered Milk. In: **Dairy Foods Processing**. New York, NY: Springer US, 2024. p. 179-192.
- SCHUMER, N. G., AHMED, M. W., RAUSCH, K., SINGH, V., KAMRUZZAMAN, M. Chemometric-based approach for economically motivated fraud detection in organic spices via NIR spectroscopy. **Journal of Food Composition and Analysis**, v. 142, p. 107538, 2025.
- SHI, R., YU, Z., HO, H., WU, W., WANG, J., TIAN, X., HAN, R. Survey of aflatoxin M1 in commercial liquid milk products in China. **Journal of Food Protection**, v. 84, n. 2, p. 200-203, 2021.
- TAN, Z., TAN, F., ZHAO, L., LI, J. The synthesis, characterization and application of ciprofloxacin complexes and its coordination with copper, manganese and zirconium ions. **Journal of Crystallization Process and Technology**, v. 2, n.2, p. 55-63, 2012.
- TASCI, F., CANBAY, H. S., DOGANTURK, M. Determination of antibiotics and their metabolites in milk by liquid chromatography-tandem mass spectrometry method. **Food Control**, v. 127, p. 108147, 2021.
- TEIXEIRA, R. C., LUIZ, L. C., JUNQUEIRA, G. M. A., BELL, M. J. V., ANJOS, V. C. Detection of antibiotic residues in Cow's milk: A theoretical and experimental vibrational study. **Journal of Molecular Structure**, v. 1215, p. 128221, 2020.
- THEAKSTONE, A. G., BRENNAN, P. M., JENKINSON, M. D., MILLS, S. J., SYED, K., RINALDI, C., BAKER, M. J. Rapid spectroscopic liquid biopsy for the universal detection of brain tumours. **Cancers**, v. 13, n. 15, p. 3851, 2021.
- WEDEL, C., ATAMER, Z., DETTLING, A., WENNING, M., SCHERER, S., HINRICHS, J. Towards low-spore milk powders: A review on microbiological challenges of dairy powder production with focus on aerobic mesophilic and thermophilic spores. **International Dairy Journal**, v. 126, p. 105252, 2022.
- YAN, C. A review on spectral data preprocessing techniques for machine learning and quantitative analysis. **iScience**, 2025.

YANG, C., GUANG, P., LI, L., SONG, H., HUANG, F., LI, Y., HU, J. Early rapid diagnosis of Alzheimer's disease based on fusion of near-and mid-infrared spectral features combined with PLS-DA. **Optik**, v. 241, p. 166485, 2021.

ZHANG, Y., ZHENG, Z., LIU, C., TAN, C. P., XIE, K., LIU, Y. A comparative study between freeze-dried and spray-dried goat milk on lipid profiling and digestibility. **Food Chemistry**, v. 387, p. 132844, 2022.

ZHENG-FENG, L. I., GUANG-JIN, X. U., JIA-JUN, W. A. N. G., GUO-RONG, D. U., WEN-SHENG, C. A. I., XUE-GUANG, S. H. A. O. Outlier detection for multivariate calibration in near infrared spectroscopic analysis by model diagnostics. **Chinese Journal of Analytical Chemistry**, v. 44, n. 2, p. 305-309, 2016.

ZHOU, S., ZHANG, X., ZHANG, J., ZANG, C., FAN, R., WANG, J., YANG, Y. Differences in physicochemical properties and proteomics analysis of spray-and freeze-dried milk powders from bovine, goat, and horse sources. **Journal of Dairy Science**, n. 2, v. 108, p. 1367-1379, 2025.

ZHU, D., KEBEDE, B., CHEN, G., MCCOMB, K., FREW, R. Impact of freeze-drying and subsequent storage on milk metabolites based on ¹H NMR and UHPLC-QToF/MS. **Food Control**, v. 116, p. 107017, 2020.

ZHU, M. Z., WEN, B., WU, H., LI, J., LIN, H., LI, Q., LIU, Z. The Quality Control of Tea by Near-Infrared Reflectance (NIR) Spectroscopy and Chemometrics. **Journal of Spectroscopy**, v. 2019, n. 1, p. 8129648, 2019.