



UEPB

**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I - CAMPINA GRANDE
CENTRO DE CIÊNCIA E TECNOLOGIA
DEPARTAMENTO DE QUÍMICA INDUSTRIAL
CURSO DE GRADUAÇÃO EM QUÍMICA INDUSTRIAL**

MAYARA MACEDO DA MATA

**AVALIAÇÃO DE ALGORITMOS DE SELEÇÃO DE VARIÁVEIS PARA A
CONSTRUÇÃO DE MODELOS PARA QUANTIFICAÇÃO DE DISSULFETO DE
CAPTORPIL EM COMPRIMIDOS**

**CAMPINA GRANDE
2022**

MAYARA MACEDO DA MATA

**AVALIAÇÃO DE ALGORITMOS DE SELEÇÃO DE VARIÁVEIS PARA A
CONSTRUÇÃO DE MODELOS PARA QUANTIFICAÇÃO DE DISSULFETO DE
CAPTORPIL EM COMPRIMIDOS**

Trabalho de Conclusão de Curso apresentada ao Departamento de Química da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Graduada em Química Industrial.

Orientadora: Prof.^a Dra. Welma Thaise Silva Vilar

Coorientador: Prof.^a Dra. Simone da Silva Simões

**CAMPINA GRANDE
2022**

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

M425a Mata, Mayara Macedo da.

Avaliação de algoritmos de seleção de variáveis para a construção de modelos para quantificação de dissulfeto de captopril em comprimidos [manuscrito] / Mayara Macedo da Mata. - 2022.

47 p. : il. colorido.

Digitado.

Trabalho de Conclusão de Curso (Química Industrial - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2023.

"Orientação : Profa. Dra. Welma Thaise Silva Vilar , Coordenação do Curso de Química Industrial - CCT."

"Coorientação: Profa. Dra. Simone da Silva Simões , Departamento de Química - CCT."

1. Quimiometria. 2. Degradação de medicamentos. 3. Controle de qualidade. 4. Farmacologia. I. Título

21. ed. CDD 615.1

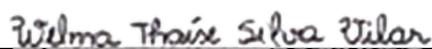
MAYARA MACEDO DA MATA

**AVALIAÇÃO DE ALGORITMOS DE SELEÇÃO DE VARIÁVEIS PARA A
CONSTRUÇÃO DE MODELOS PARA QUANTIFICAÇÃO DE DISSULFETO DE
CAPTORPIL EM COMPRIMIDOS**

Trabalho de Conclusão de Curso
apresentada ao Departamento de
Química da Universidade Estadual
da Paraíba, como requisito parcial à
obtenção do título de Graduada em
Química Industrial.

Aprovada em: 13/12/2022 .

BANCA EXAMINADORA



Prof.^a Dra. Welma Thaise Silva Vilar
(Orientadora)
Universidade Estadual da Paraíba (UEPB)



Prof.^a Dra. Adriana da Silva Lima
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Gessenildo Pereira Rodrigues
Faculdade Reboças (FRCG)

AGRADECIMENTOS

Agradeço a Deus por ser essencial e luz na minha vida, por toda trajetória conquistada até aqui, por ter me dado força e coragem para enfrentar as diversidades da vida e por me dar clareza de pensamentos para aprender dia após dia.

A minha mãe Marquesa e a minha irmã Maynara, por me motivarem sempre em busca dos meus sonhos. Amo muito vocês.

A minha família pelo apoio, incentivo e por fazerem parte desta história de conquistas.

A meu namorado Wylliam Eduardo por todo companheirismo, afeto e amor. Obrigada por aguentar minhas crises de estresse e ansiedade e por estar presente nos momentos mais importantes para mim. Amo você!

As minhas amigas Samara, Louyse e Diego pelo seu carinho, estímulo, amizade, companheirismo e pelas palavras de motivação que me fizeram ter forças para seguir em frente.

A Prof^a. Dr^a Welma Vilar por toda orientação e pelos ensinamentos que irei levar para vida profissional e pessoal e que foram fundamentais para a realização deste sonho. Obrigada por todo apoio.

A todos os professores e coordenação de curso de Química Industrial da UEPB que tiveram sua cota de contribuição na finalização nesta etapa da minha vida.

E agradeço a todos que de forma direta ou indiretamente contribuíram e fizeram parte deste ciclo que se encerra, o meu muito obrigado.

“Algumas pessoas marcam a nossa vida para sempre, umas porque nos vão ajudando na construção, outras porque nos apresentam projetos de sonho e outras ainda porque nos desafiam a construí-los”.

“Mentalize um futuro para sua vida e não pare até chegar lá”. (Rafael Magalhães)

RESUMO

As doenças cardiovasculares são as principais causas de morte no Brasil, representando uma grande sobrecarga para a saúde pública e impedindo gravemente o desenvolvimento socioeconômico. Um dos medicamentos mais utilizados pelos brasileiros no tratamento da hipertensão é o captopril, considerado um dos fármacos básicos para a população, sendo disponibilizado pelo Sistema único de Saúde (SUS). O dissulfeto de captopril é um produto de degradação do captopril que apresenta alta sensibilidade quando exposto a alta umidade e temperatura, tornando-se prejudicial ao tratamento terapêutico. Para determinar concentrações dos princípios ativos na indústria farmacêutica, métodos espectroscópicos estão sendo usados em conjunto com análises quimiométricas. Devido ao avanço tecnológico a quimiometria surge como uma ferramenta a fim de adquirir o máximo de informação sobre o sistema, tornando-se crucial para a química analítica. Neste sentido o objetivo deste trabalho foi utilizar as seleções de variáveis Algoritmo Genético (GA, “Genetic Algorithm”), Algoritmo das Projeções Sucessivas (SPA) e Stepwise (SW) associados ao modelo quimiométrico MLR para avaliar a concentração do produto de degradação do captopril em comprimidos por espectroscopia NIR. Foram utilizados 23 lotes, sendo que 400 comprimidos pertencentes ao mesmo lote e 220 comprimidos pertencentes aos 22 lotes distintos (10 comprimidos de cada lote). Para o estudo da degradação os comprimidos de Captopril foram preservados dentro de sua embalagem primária e 200 comprimidos foram distribuídos em placa de petri (sem tampa) para ficarem expostos às condições de degradação. As medidas espectrais foram adquiridas por meio de espectrofotômetros FTIR. Os espectros NIR de todas as amostras foram registradas e modelos de classificação foram construídos para determinar a concentração de captopril. Foi possível perceber que diferentes pré-processamentos apresentaram resultados satisfatórios para os modelos empregados, com baixos valores de RMSECV. Sendo assim, o pré-processamento utilizando a 2ª derivada de Savitzky-Golay com janela de 11 pontos foi escolhido por apresentar menores valores de RMSECV e maiores valores de r_{cv} . Contudo, a metodologia proposta e desenvolvida neste trabalho apresentaram bons desempenhos para quantificar produto de degradação em captopril. Tais característica é relevante para que se realize esta metodologia, por órgãos de vigilância, o controle da qualidade de captopril no mercado.

Palavras-Chave: captopril; degradação; infravermelho próximo; quimiometria; seleção de variáveis.

ABSTRACT

Cardiovascular diseases are the main cause of death in Brazil, representing a great burden for public health and seriously impeding socioeconomic development. One of the drugs most used by Brazilians in the treatment of hypertension is captopril, considered one of the basic drugs for the population, being made available by the Unified Health System (SUS). Captopril disulfide is a captopril degradation product that is highly sensitive when exposed to high humidity and temperature, making it harmful to therapeutic treatment. To determine the intentions of active principles in the pharmaceutical industry, spectroscopic methods are being used in conjunction with chemometric analyses. Due to technological advances, chemometrics emerges as a tool to expand classification issues, becoming crucial for analytical chemistry. In this sense the objective of this work was to use the Genetic Algorithm (GA), Successive Projection Algorithm (SPA) and Stepwise (SW) variable selections associated with the MLR chemometric model to evaluate the concentration of the captopril degradation product in tablets by NIR spectroscopy. 23 batches were used, with 400 pills belonging to the same batch and 220 pills belonging to 22 different batches (10 pills from each batch). In order to study the degradation of the Captopril tablets, the tablets were preserved inside their primary packaging and 200 tablets were distributed in a petri dish (without a lid) to be exposed to the degradation conditions. Spectral measurements were acquired using FTIR spectrophotometers. Nir spectra of all samples were recorded and classification models were constructed to determine captopril concentration. In the results it is possible to perceive that different pre-processings presented satisfactory results for the employed models, with low values of RMSECV. Therefore, pre-processing using the 2nd derivative of Savitzky-Golay with an 11-point window was chosen because it presents lower RMSECV values and higher r_{cv} values. However, the methodology proposed and developed in this work showed good performance to quantify the degradation product in captopril. Such characteristics are relevant for this methodology to be carried out, by surveillance agencies, to control the quality of captopril on the market.

Keywords: captopril; degradation; near infrared; chemometrics; Selection of variables.

LISTA DE FIGURAS

Figura 1 –	Estrutura do captopril	17
Figura 2 –	Reação de degradação do captopril com a formação de dissulfeto de captopril.....	18
Figura 3 –	Bomem Modelo FTLA 2000-160 com acessório para refletância difusa	
Figura 4 –	Powder Sampl/R	32
Figura 5 –	Espectros médio NIR das amostras de captopril.....	35
Figura 6 –	Intervalos de variáveis selecionadas pelo Ispa/PLS.....	37
Figura 7 –	Espectros NIR médios do conjunto de dados com indicação das variáveis selecionadas pelos algoritmos SPA (●), SW (■) e GA (▲), para captopril, a,b e c respectivamente.....	38
Figura 8 –	Elipse de confiança para o captopril obtidas para os modelos iSPA/PLS (linha azul), MLR/SPA (linha verde), MLR/SW (linha vermelho) e MLR/GA (linha laranja).....	41
Figura 9 –	Gráfico dos valores preditos versus os valores de referências obtidos pelos modelos iSPA/PLS (a), MLR/SPA (b), MLR/SW (c) e MLR/GA (d).....	42

LISTA DE TABELAS

Tabela 1 –	Resultados obtidos para os conjuntos de calibração na determinação de captopril para os diferentes pré-processamentos.....	36
Tabela 2 –	Resultados obtidos pelos modelos iSPA/PLS, MLR/SPA, MLR/SW e MLR/GA para o conjunto de predição em termos de RMSEP, valores de correlação e de determinação. N é o número de amostras de predição.....	40

LISTA DE ABREVIATURAS E SIGLAS

ECA	Enzima de Conversão da Angiotensina
GA	Algoritmo Genético (<i>Genetic Algorithm</i>)
IECA	Inibidores da Enzima Conversora de Angiotensina
LAFEPE	Laboratório Farmacêutico de Pernambuco
LDA	Análise Discriminante Linear
MLR	Regressão Linear Múltipla
NIR	Infravermelho Próximo (<i>Near Infrared</i>)
PAD	Pressão Arterial Diastólica
PAS	Pressão Arterial Sistólica
PCA	Análise de Componentes Principais (<i>Principal Componente Analysis</i>)
PCR	Reação em Cadeia da Polimerase (<i>Polymerase Chain Reaction</i>)
PLS	Mínimos Quadrados Parciais (<i>Patial Least Squares</i>)
RMSEC	Raiz Quadrada do Erro Médio Quadrático da Calibração (<i>Root Mean Square Error of Calibration</i>)
RMSEP	Raiz Quadrada do Erro Médio Quadrático de Previsão (<i>Root Mean Squares Error of Prediction</i>)
RMSE	Raiz Quadrada do Erro Médio
SPA	Algoritmo das Projeções Sucessivas
SQR	Substância Química de Referência
SW	<i>Stepwise</i>

SUMÁRIO

1	INTRODUÇÃO	11
2	OBJETIVOS	13
2.1	Objetivos Gerais	14
2.2	Objetivos Específicos	14
3	REVISÃO DA LITERATURA	16
3.1	Captopril	16
3.1.1	<i>Dissulfeto de Captopril</i>	17
3.2	Região Espectral do Infravermelho Próximo	18
3.3	Quimiometria	18
3.3.1	<i>Calibração Multivariada</i>	19
<i>3.3.1.1</i>	<i>Regressão Parcial de mínimos quadrados (PLS)</i>	20
<i>3.3.1.2</i>	<i>Regressão Linear Múltipla (MLR)</i>	21
3.3.2	<i>Seleção de Variáveis</i>	22
<i>3.3.2.1</i>	<i>Algoritmo Genético (GA)</i>	23
<i>3.3.2.2</i>	<i>Algoritmo das projeções sucessivas (SPA)</i>	24
<i>3.3.2.3</i>	<i>Stepwise (SW)</i>	26
4	MATERIAL E MÉTODOS	28
4.1	Amostras	28
4.2	Degradação dos Comprimidos de Captopril	28
4.3	Instrumentação Analítica	29
4.4	Solução Padrão de Dissulfeto de Captopril	29
4.5	Solução Padrão de Captopril	30
4.6	Tratamento de Dados	30
5	RESULTADOS E DISCUSSÃO	31
5.1	Espectros Nir	31
5.2	Modelos de Calibração	32
<i>5.2.1</i>	<i>Avaliação dos Modelos de calibração Desenvolvidos</i>	34
<i>5.2.1.1</i>	<i>iSPA/PLS</i>	34
<i>5.2.1.2</i>	<i>Seleção de Variáveis</i>	35
5.2.2	Análise dos modelos de predição	36
6	CONCLUSÃO	43
	REFERÊNCIA	44

1 INTRODUÇÃO

Desde os primórdios o homem vem enfrentando desafios para controlar, reduzir ou eliminar os sofrimentos causados pelas enfermidades. É evidente a importância do cuidado à saúde e a contribuição da terapia medicamentosa. O cuidado farmacêutico é uma execução de saúde pública, sendo determinante para a eficiência das ações na saúde. Assim, formando um elo entre a saúde pública e o usuário do medicamento (CONASS, 2007).

As principais causas de mortes no Brasil são causadas por doenças cardiovasculares que representa uma taxa de 30,8% dos óbitos, atingindo cerca de 30 milhões de brasileiros, ocasionando um alto custo na saúde pública, afetando rigorosamente o desenvolvimento socioeconômico (SANTOS, 2017).

A hipertensão arterial (pressão alta) é caracterizada pela pressão sanguínea diastólica permanentemente aumentada acima de 90 mmHg e com a sistólica acima de 140 mmHg (RAFAEL; FARIA, 2013). Dentre a terapia medicamentosa podemos encontrar o captopril que está no *ranking* entre os fármacos mais usado pelos brasileiros para o tratamento da hipertensão, que corresponde a, aproximadamente, 11, 2% de todos os fármacos declarados (MENGUE et al., 2016). Este fármaco é disponibilizado pelo Sistema único de Saúde (SUS) como um medicamento considerado essencial para o acesso à população. Contudo, o captopril apresenta considerável instabilidade, sendo estritamente passível à degradação oxidativa causada por elevadas temperaturas e umidade, ou ainda pela mistura a excipientes higroscópicos (RAFAEL; FARIA, 2013).

O Tratamento pode ser efetuado de duas maneiras: através de uma dieta com pouco sal e gordura, além da mudança de hábitos, e através do uso de medicamentos. A primeira opção é indicada para casos mais simples, onde se consegue manter a pressão controlada apenas com a mudança de hábitos saudáveis. Contando que, a segunda opção é quando há um maior comprometimento dos órgãos-alvo (coração), assim sendo necessário o uso de medicamentos para o controle da hipertensão arterial (ARAÚJO, 2004).

O tratamento através do uso de medicamentos deve ser eficaz para manter os limites de normalidade para a pressão arterial, evitando eventos mórbidos cardiovasculares fatais e não fatais, que visa, sobretudo, manter os valores tensionais abaixo de 140/90 mmHg (SOCIEDADE BRASILEIRA DE HIPERTENSÃO, 2006). Dentre a terapia medicamentosa podemos encontrar o captopril que está no *ranking* entre os fármacos mais usado pelos brasileiros para o tratamento da hipertensão, que corresponde a, aproximadamente, 11, 2% de todos os fármacos declarados (MENGUE et al., 2016).

A degradação do captopril é danosa ao tratamento terapêutico, pois uma elevada concentração de dissulfeto de captopril implica na diminuição do teor ativo. Além disso, quando os valores de dissulfeto de captopril estiverem acima do permitido pela legislação causará reações não satisfatórias no organismo, sendo o limite de 3% em relação a quantidade do ativo (NÓBREGA, 2006).

Diante do exposto, fica clara a necessidade da determinação e quantificação do dissulfeto de captopril em comprimidos. E para isto, metodologias quimiométricas, dando ênfase a seleção de variáveis, foram utilizadas neste trabalho.

A seleção de variáveis é bastante expandida em diversos campos da quimiometria. A técnica é classificada conforme o subconjunto de variáveis selecionado (seleção de variáveis individuais discretas e intervalos) e a presença de variáveis aleatória (randômicas ou determinísticas) (GOMES et al., 2013).

O uso desta técnica permite a construção de modelos robustos e de fácil interpretação com o objetivo de diminuir a quantidade de variáveis que apresentam comportamento semelhante a ruído, ou caso seja de pertinente, selecionar as variáveis mais relevantes para o modelo, uma vez que a escolha de regiões específicas pode minimizar os erros destes. Portanto, muitos algoritmos de seleção de variáveis vêm sendo desenvolvidas para distinguir as variáveis que tem importância para o modelo e eliminar aquelas variáveis desnecessárias (MORO et al., 2021).

A seleção de variáveis é utilizada em diversos algoritmos de calibração multivariada, dentre eles temos: Regressão Linear Múltipla (MLR, do inglês *Multiple Linear Regression*) e Regressão por Mínimos Quadrados Parciais (PLS, do inglês *Partial Least Squares Regression*). Contudo, como no PLS as regiões espectrais podem conter ruídos heterocedásticos (dispersão dos dados em torno de uma reta), bem como a inclusão de variáveis redundantes que afetam a preditividade do modelo quimiométricos, nestes casos a técnica pode ser implementada. E para o MLR a seleção é uma alternativa necessária à construção do modelo devido a multicolinearidade (BRERETON, 2003).

Dessa forma, objetiva-se utilizar os algoritmos de seleção de variáveis GA, SW e SPA para quantificar o produto de degradação do captopril em comprimidos, empregando espectroscopia NIR.

2 OBJETIVOS

2.1 Objetivo geral

Utilizar as seleções de variáveis Algoritmo Genético (GA, “Genetic Algorithm”), Algoritmo das Projeções Sucessivas (SPA), Stepwise (SW) associados ao modelo quimiométricos MLR, e regressão pelos mínimos quadrados parciais por intervalos (iPLS) para avaliar a concentração do produto de degradação do captopril em comprimidos por espectroscopia NIR.

2.2 Objetivos específicos

- ✓ Construir modelos de calibração e validação multivariada para quantificar o produto de degradação do captopril em comprimidos;
- ✓ Avaliar o modelo obtido por meio de alguns parâmetros de desempenho (RMSECV, RMSEP e R^2);
- ✓ Selecionar variáveis mais relevantes para o modelo;
- ✓ Validar o modelo aplicando a calibração multivariada e a espectroscopia NIR.

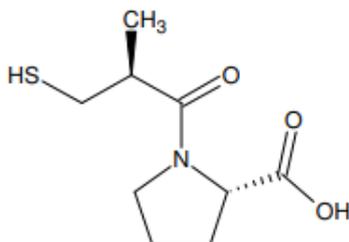
3 REVISÃO DA LITERATURA

3.1 Captopril

A hipertensão arterial é uma doença crônica, no qual uma pressão é exercida pelo sangue nas artérias do corpo, caracterizada pelos níveis pressóricos elevados. A aparição da elevação da pressão arterial acontece quando a Pressão Arterial Sistólica (PAS) e a Pressão Arterial Diastólica (PAD) são encontradas com valores iguais ou maiores que 140 mmHg e 90 mmHg (BARROSO, et al., 2021).

O captopril (figura 1) é um fármaco, anti-hipertensivo, que pertence a classe de inibidores da enzima conversora de angiotensina (IECA), que exerce influência na contração dos vasos sanguíneos, controlando a hipertensão arterial. Captopril se apresenta na forma de um pó cristalino branco ou quase branco, prontamente solúvel em metanol, cloreto de metileno e água e soluções diluídas de hidróxidos alcalinos. Apresenta faixa de fusão de 105° C a 108° C e massa molecular de 217,29 g/mol. Sua fórmula molecular é C₉H₁₅NO₃S e possui valores de pKa de 3,7 (carboxila) e 9,8 (tiol) (ANVISA, 2009).

Figura 1. Estrutura do captopril



Fonte: Lamim (2006).

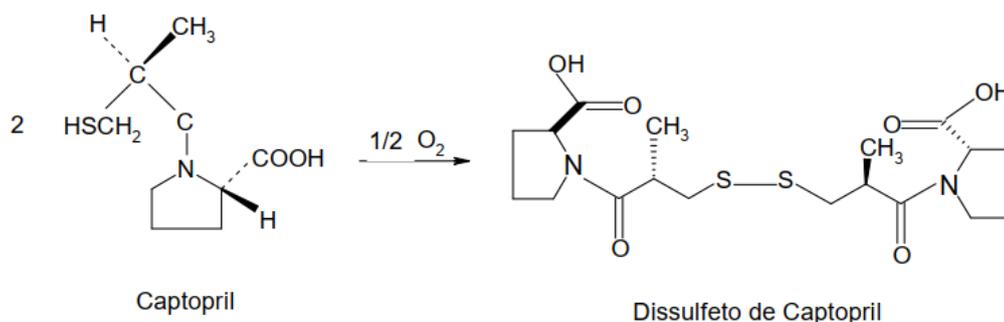
A renina-angiotensina faz parte da classe dos peptídeos e enzimas para o controle da pressão arterial no organismo humano. A renina é liberada na corrente sanguínea quando acontece uma diminuição na pressão arterial, sendo convertida em angiotensinogênio que se divide em angiotensina I (SOUZA, 2015). A angiotensina I é relativamente não reativa, sendo dividida pela enzima de conversão da angiotensina (ECA) em angiotensina II, fazendo com que as paredes musculares das pequenas artérias se contraíam, ocasionando um aumento na pressão arterial. O mecanismo dessas substâncias depende da Inibição da Enzima Conversora Angiotensina (IECA), pois são vasoconstritores potentes impedindo a conversão de angiotensina I em II, proporcionando a diminuição da pressão arterial (BRUNTON, KNOLL, 2018).

O fármaco captopril se comporta de modo a impossibilitar a enzima de conversão de angiotensina (ECA) interagir com a angiotensina I, impedindo a formação da angiotensina II, conseqüentemente, impedindo a elevação da pressão arterial. Este mecanismo está relacionado com as habilidades destas moléculas se ligarem ao sítio catalítico da enzima, no qual, sua atividade catalítica é dependente de zinco II (Zn^{2+}) e de uma seqüência tri peptídica específica que está relacionada ao reconhecimento molecular de seu substrato (TAVARES et. al., 2015). Este fármaco apresenta um diferencial por conseguir o efeito da inibição da ECA e pela presença no sítio ativo da enzima ligações de átomo de zinco (OIGMAN, FRITSH, 1998). Através do átomo de oxigênio o Zn^{2+} se coordena com o íon acetato, duas moléculas de histidina através do átomo de nitrogênio do anel imidazólico e pelo átomo e enxofre com o fármaco captopril. Além do que, o fármaco realiza interações com aminoácidos da ECA através de ligações de hidrogênio, sendo suficientes para a interação efetiva com a enzima (NATESH et al., 2004).

3.1.1 Dissulfeto de Captopril

O dissulfeto de captopril é a principal impureza presente no fármaco captopril, possuindo massa molecular de 432,58 g/mol, sendo considerado o seu principal produto de degradação (CONNORS; AMIDON; STELLA, 1986). O limite de dissulfeto de captopril em comprimidos de captopril é de no limite de 3% em relação a quantidade do ativo. A reação de degradação do captopril com a formação de seu dímero, sendo mais estável em soluções ácidas que em soluções básicas é demonstrada na Figura 2.

Figura 2. Reação de degradação do captopril com a formação de dissulfeto de captopril



Fonte: Lamim (2006).

Outra condição importante de ser mencionada é o fato desse fármaco ter um aumento significativo de tióis em $\text{pH} > 4$, estando sujeito à degradação oxidativa em seu grupo sulfidríla, causado por elevadas temperaturas e umidade, promovendo a formação do seu produto de degradação (LAMIM, 2006). Assim, o composto farmacêutico na forma sólida, necessita conter em sua embalagem primária barreiras efetivas contra a umidade e oxigênio, para permitir um prazo de validade adequado conforme as condições ambientais de elevada umidade e calor, tais condições presentes no Brasil (MARCATTO, 2005).

Uma vez que produtos farmacêuticos disponíveis no mercado estejam com problemas de não conformidade, acima do especificado pelas competências oficiais da legislação da vigilância sanitária, o produto é interrompido para comercialização, tendo em vista que podem acarretar danos a qualidade de vida das pessoas e acarretar danos ao tratamento terapêutico. (Brasil, Provene, 2008). Portanto, valores acima do permitido para o dissulfeto de captopril provocará efeitos indesejáveis no organismo, ocasionando diminuição do teor do ativo (MARCOLINO JÚNIOR, 2007).

3.2 Região espectral do infravermelho próximo

A técnica de espectroscopia no infravermelho próximo (NIRS) mede a energia absorvida pelas moléculas para que os grupos funcionais presentes nessas moléculas contenham informações analíticas quantitativas e qualitativas sobre a amostra. A técnica é baseada em espectroscopia vibratória no infravermelho próximo que resulta da interação de radiação eletromagnética com os constituintes da amostra na faixa de comprimento de onda entre 750 e 2500 nm (13300 a 4000 cm^{-1}) (ZAREEF et al., 2020). Esta técnica está sendo bastante utilizadas nos setores alimentícios, agrícolas e farmacêuticos, como também em diversas outras áreas, devido a sua versatilidade analítica, respectivamente, a sua potencialidade em medições quantitativas e qualitativas (ALVES; POPPI, 2013).

Os principais fatores que causam a formação das bandas de absorção no NIR são combinações e sobretons de modos vibratórios, que são acompanhados por uma variação no momento dipolo, assim, acontecendo a absorção da radiação NIR. O grau de anarmonicidade de grupos envolvendo o átomo de hidrogênio (mais leve) e o elemento mais pesado, como carbono, enxofre, nitrogênio e oxigênio, que correspondem às respectivas ligações CH, SH, NH e OH, determina a intensidade de uma dada absorção. A espectroscopia NIR é uma técnica não destrutiva, rápida e não invasivas, tendo um alto poder de penetração

na faixa de radiação de sondagem e sem a necessidade de uma preparação de amostra com exigências (PASQUINI, 2018).

As bandas de combinação se formam quando os modos vibracionais interagem com mudanças de energia simultâneas, cuja frequência é igual à soma das frequências capazes de excitar cada modo vibracional e localizadas na região do infravermelho próximo entre 1900 e 2500 nm. Dada a sua complexidade, bandas de absorção no NIR possuem características típicas como comprimento, sobreposição e intensidade fraca. Isso dificulta a avaliação dos dados espectrais. Assim, a molécula absorve a radiação infravermelha quando possui uma variação no momento dipolo durante os movimentos vibracionais e rotacionais, podendo acontecer de diferentes formas nas ligações das moléculas (SKOOG; HOLLER; NIEMAN, 2009).

A espectroscopia NIR fornece informações das propriedades químicas e físicas de uma amostra, conseguindo determinar vários parâmetros de qualidade em uma única medida (SIMÕES, 2008). Contudo, o uso da técnica necessita da combinação com ferramentas quimiométricas para tratar os dados espectrais, fornecendo métodos não destrutivos, diretos e de baixo custo, não havendo a necessidade do pré-tratamento da amostra e nem do uso de reagentes químicos (TIBOLA et al., 2018).

A análise dos dados espectrais NIR, juntamente com o uso de ferramentas quimiométricas são utilizadas com o objetivo de esclarecer as informações químicas presentes no espectro na forma de bandas ampliadas e sobrepostas. Essas técnicas auxiliam na interpretação das variáveis espectrais, bem como no desenvolvimento de modelos que possam compreender a distribuição de amostras enquanto procuram tendências de comportamento. Esses modelos foram desenvolvidos para quantificar e/ou classificar amostras usando as informações químicas extraídas dos espectros (JACKSON; MANTSCH, 2001).

A análise espectral NIR, juntamente com técnicas quimiométricas, vem se mostrando promissora em estudos de quantificações em formulações farmacêuticas por meio da construção de modelos de classificação, contribuindo com a retirada e tratamento de dados úteis para análise, como podemos ver em alguns estudos apresentados a seguir.

Maluf et al. (2010) em estudo com 22 misturas contendo sibutramina e usando a espectroscopia infravermelha de refletância difusa (DRIFTS, do inglês: *Diffuse Reflectance Infrared Fourier Transform Spectroscopy*) juntamente com regressão de mínimos quadrados parciais, obteve compatibilidade com as normas da Agência Nacional de Vigilância Sanitária (ANVISA-Brasil) e os processos de calibração multivariada utilizados no estudo viabilizam a determinação de sibutramina em formas farmacêuticas.

Souza et al. (2012) propuseram o uso da espectroscopia NIR para quantificar as impurezas do comprimido de captopril. Os autores empregaram modelos de mínimos quadrados parciais para determinar o produto de degradação que obtiveram um erro quadrático médio de predição (RMSEP) de 0,074 mg por comprimido. Concluindo que a metodologia foi validada e pode ser facilmente utilizada para análises de rotina.

Grangeiro Junior et al. (2014) investigaram o potencial da espectroscopia NIR associadas a duas técnicas de calibração multivariada (PLS e MLR) para monitorar a produção de comprimidos combinados de dose fixa de zidovudina (AZT) e lamivudina (3TC) em três etapas diferentes de produção: no material granulado, no núcleo e no comprimido fabricado. Imagens químicas NIR foram usadas para avaliar a distribuição química nas superfícies dos comprimidos não revestidos e suas estruturas de correlação interna. Como também, foram monitoradas o teor de umidade no material granulado e a dureza dos comprimidos antes da etapa de revestimento. A metodologia proposta mostrou que os modelos apresentaram resultados precisos e erros de previsão abaixo dos permitidos pelos órgãos reguladores.

Mesmo com todas suas vantagens e aplicabilidades, a interpretação do NIR, para muitos casos, ainda é complexa, e o desenvolvimento de métodos quimiométricos facilitou a compreensão e trouxe um tratamento estatístico e químico para os espectros, contribuindo com a retirada e tratamento de dados úteis para análise.

O desenvolvimento de métodos quimiométricos ajudou a simplificar essa complexidade e trouxe um tratamento estatístico e químico para os espectros, na qual, ajudou a processar os dados para sua análise. No entanto, algumas variáveis espectrais não são esclarecedoras na construção do modelo, sendo necessário excluí-las. Nesse sentido, os métodos de seleção de variáveis incentivam a extração de informações químicas importantes e redução de ruídos para a construção de modelos mais precisos do que os modelos globais, que utilizam todas as variáveis (CHEN et al. 2017).

3.3 Quimiometria

Com o avanço tecnológico e a necessidade de ampliar métodos para adquirir informação digital adquirida, surge a quimiometria, como ferramenta para esclarecer problemas de classificação, se tornando decisiva para a química analítica (FERREIRA, 2022).

Essa área surgiu com a evolução de equipamentos de análise de dados, sendo voltada aos estudos de dados químicos, ou seja, com o progresso do avanço científico possibilita a análise de uma grande quantidade de dados e em diferentes níveis de complexidade. Assim, a

quimiometria propicia um estudo de forma multivariada, retirando informações de dados químicos através de métodos matemáticos e estatísticos com a finalidade de otimizar ou planejar procedimentos experimentais, viabilizando a análise de dados de maior complexidade, assim obtendo melhores resultados analíticos (OLIVEIRA, 2019). Dentro dessa perspectiva, a quimiometria vem crescendo no Brasil e sendo inserida em diversos campos de pesquisa, sendo mais empregada na química analítica, além do mais beneficiando setores industriais e acadêmicos (BEEBE; PELL; SEASHOLTZ, 1998).

De acordo, com o interesse de estudo, ferramentas quimiométricas são utilizadas para análise qualitativas ou quantitativas. Dentre as áreas principais temos: planejamento e otimização de experimentos, reconhecimento de padrões e calibração multivariada. No qual, as duas últimas citadas conquistaram visibilidade nas aplicações analíticas (STEKI; NOUROOZI; SHAHSAVARI, 2015).

3.3.1 Calibração Multivariada

A calibração multivariada é dos métodos mais utilizados na quimiometria. Na química analítica a calibração multivariada vem sendo muito utilizada, sobretudo quando os componentes de uma mistura precisam ser determinados (BRO, 2003). A calibração multivariada tem como objetivo investigar toda informação fornecida pelas técnicas instrumentais, construindo modelos matemáticos através dos dados multivariados, favorecendo a previsão de uma resposta de interesse, sejam elas de caráter química ou física (NAES et al., 2002). Os dados do modelo são coletados através da espectroscopia do ultravioleta e do visível.

A aplicabilidade deste método é direcionada para encontrar um algoritmo matemático que produza informações quantitativas. A partir disso, prever a concentração do analito de interesse, através dos resultados instrumentais obtidos por um determinado equipamento (WANG et al., 2022). Posteriormente, é possível fazer uma relação com a propriedade de interesse através da obtenção dos dados instrumentais, onde as respostas são as intensidades medidas (FERREIRA, 2015).

O método de calibração multivariada demonstra uma relação entre duas matrizes de dados, isto é, quando suas propriedades demonstram dependência entre elas. A partir da matriz é possível estabelecer todas as variações existentes, procurando as relações entre as amostras e as variáveis do conjunto de dados e convertê-las em novas variáveis latentes. (KUMAR et al., 2014).

As técnicas mais conhecidas na análise multivariada são: a regressão pelo método dos mínimos quadrados parciais – PLS, do inglês, *Partial Least Squares*; a regressão linear múltipla – MLR, do inglês, *Multiple Linear Regression*; e a análise de componentes principais – PCA, do inglês *Principal Component Analysis*).

O método PLS, técnica de regressão supervisionada, relaciona variáveis independentes (matriz X) com as propriedades de interesse (vetor Y), utilizando a análise de componentes principais como objetivo de reduzir a dimensionalidade da matriz (FERREIRA, 2015). Já, o MLR é caracterizado por estabelecer uma ligação entre um número reduzido de variáveis e uma propriedade das amostras, resolvendo uma série de equações simultâneas. Por outro lado, a PCA, análise exploratória, consiste em transformar um conjunto de variáveis originais em outro conjunto com as mesmas dimensões. Assim, sendo possível explicar, visualizar e resolver as diferenças entre as variáveis e as ligações que podem existir entre as amostras, bem como, identificar o aparecimento de amostras anômalas (VINZI et al., 2010).

Assim, o processo da calibração multivariada consiste basicamente em três fases: calibração, validação e predição (SILVA et al., 2013). Neste trabalho serão detalhados os métodos PLS e MLR, que foram aplicados aos dados deste trabalho.

3.3.1.1 Regressão Parcial de mínimos quadrados (PLS)

A regressão pelo método dos mínimos quadrados parciais tem como base o PCA, em que a construção dos modelos é feita a partir de dados de primeira ordem, sendo capaz de prever e modelar amostras mesmo na presença de interferentes (FERREIRA, 2015). Este método considera a substituição das variáveis originais por variáveis latentes em combinações lineares (BEEBE; PELL; SEASHOLTZ, 1998). A construção de modelos é feita com base na matriz contendo as informações espectroscópicas e as informações de concentração ou propriedade de interesse, ou seja, consiste na regressão entre os escores das matrizes X e Y. (RODGERS et al., 2013). De forma geral, este método modela a máxima correlação entre os conjuntos de variáveis independentes (espectros ou outros sinais analíticos) e variáveis dependentes (concentrações ou propriedades de analitos), isto é, as informações das variáveis independentes (X) e as concentrações (Y) são aplicadas simultaneamente na fase de calibração para encontrar as variáveis latentes em X (CIROVIC, 1996).

Deste modo, no PLS as variáveis são compactadas tornando a técnica robusta, isto significa que os parâmetros são menos dispendiosos de interpretar, sendo capaz de analisar um número grande de variáveis, dados com alto ruído e problemas com colinearidade (GELADI; KOWALSKI, 1986). Em termos práticos, este é um método de calibração inversa,

na qual as variáveis a serem previstas são dependentes dos sinais analíticos. Assim, as informações das matrizes X e Y são decompostas em suas matrizes de pesos e escores, conforme as equações 1 e 2 (BRERETON, 2003).

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

Em que T é a matriz de score, P é a matriz de pesos e E é a matriz de resíduos espectrais.

$$Y = U.C^T + F \quad \text{Equação (2)}$$

Em que U é a matriz de score, C é a matriz de pesos e F é a matriz dos resíduos de concentração.

A correlação das matrizes X e Y acontece de forma linear através do vetor de coeficiente de regressão linear e pode ser observada nas equações 3 e 4.

$$U = BT + G \quad \text{Equação (3)}$$

$$Y = BTQ^T + H \quad \text{Equação (4)}$$

Em que B é a matriz dos coeficientes de regressão; G é a matriz de resíduos dos escores e H a matriz de resíduos de concentração.

Assim, a etapa de definição do número de variáveis latentes é de extrema importância, pois o número adequado de variáveis latentes a serem incluídos no modelo PLS é baseado na validação cruzada, evitando o sobreajuste do modelo (WOLD, 1978).

3.3.1.2 Regressão Múltipla Linear (MLR)

A Regressão Linear Múltipla é uma técnica estatística capaz de resolver uma série de equações simultâneas, isto é, analisar a relação linear de uma variável dependente e diversas variáveis independentes. Isso pode ser representado matematicamente como (HAIR et al., 2009):

$$Y = Xb + e \quad \text{Equação (5)}$$

Em que Y é a propriedade de interesse, X é a matriz com os dados instrumentais e e é os resíduos.

A análise de regressão tem como objetivo constituir uma ligação entre a propriedade das amostras e o número reduzido de variáveis, podendo ser descrita pela equação dos mínimos quadrados:

$$b = (X^T X)^{-1} X^T y \quad \text{Equação (6)}$$

Em que os índices subscritos representam a transposição e inversão de um vetor ou matriz (FERREIRA, 2015). Vale ressaltar, que a análise de regressão múltipla postula que o número de variáveis seja menor que o número de amostras e que tais não sejam altamente correlacionáveis. Assim, a equação necessita a inversa da matriz $X^T X$ para evitar um sistema indeterminado (LIRA, 2010). Entretanto, em dados analíticos multivariados é frequente a permanência de um número elevado de variáveis altamente correlacionadas e de o número de variáveis ser maior que o número de amostras. Para isso, utiliza-se algoritmos de seleção de variáveis (BEEBE; PELL; SEASHOLTZ, 1998).

3.3.2 Seleção de Variáveis

O modelo de calibração multivariada demonstra mais eficiência com a escolha de regiões espectrais específicas. Assim, o método de seleção de variáveis permite a escolha de regiões específicas, selecionando as variáveis relevantes para o modelo. Assim, diminuindo o erro, permitindo modelos mais robustos e de fácil interpretação, bem como melhor precisão nas previsões, aumentando o desempenho de alguns métodos de reconhecimento de padrões (MORO, 2021). Além disso, a seleção de variáveis viabiliza a utilização de instrumentos espectroscópicos portáteis (usados em campo), favorecendo a interpretação química dos espectros (SCHENONE; GALERA; GOICOECHEA, 2013).

A seleção de variáveis é bastante expandida em diversos campos da quimiometria. A técnica é classificada conforme o subconjunto de variáveis selecionado (seleção de variáveis individuais discretas e intervalos) e a presença de variáveis aleatória (randômicas ou determinísticas). A seleção de variáveis individuais discretas escolhe o conjunto de variáveis que contém as informações adequadas dentro de toda faixa espectral completa. Já por intervalos, procura uma região espectral que contenha melhores resultados que a faixa espectral completa, dividindo o espectro em iguais intervalos definidos pelo usuário (GOMES et al., 2013).

Em contrapartida, os métodos randômicos estão associados a certo grau de probabilidade, ou seja, o subconjunto de variáveis aleatórias selecionada é estocástico (LARDI; GONZALEZ, 1998). Neste método, várias escolhas são feitas com base em números aleatórios, isto é, o método não executará a mesma sequência de operações, assim seguindo o seu próprio caminho, levando a resultados finais que geralmente convergem para as mesmas soluções ótimas com dada precisão. Já os métodos determinísticos, retrata como solução um único subconjunto de variáveis, dessa forma, não estando relacionados a nenhuma probabilidade a priori. Neste método, executando diferentes operações com as mesmas

condições iniciais chegará à mesma resposta, havendo a reprodutibilidade do subconjunto de variáveis selecionadas (YANG, 2010).

Diversos métodos têm sido propostos para a finalidade de seleção de variáveis. Para este trabalho, foram empregados apenas Algoritmo Genético (GA, “Genetic Algorithm”), Algoritmo das Projeções Sucessivas (SPA) e Stepwise (SW) e os quais são descritos a seguir.

3.3.2.1 Algoritmo Genético (GA)

Em 1975, o algoritmo genético foi proposto por John H. Holland. É um algoritmo matemático e tem como base a teoria da biologia evolutiva, sendo uma técnica de otimização. O método projetava matematicamente o processo da evolução biológica com todas as características do processo. No decorrer dos anos, o algoritmo passou a ser aplicado na química para resolver diversos tipos de problemas (WILLIS et al., 1997).

O algoritmo genético permite que o subconjunto de variáveis selecionadas estudadas seja representado por um vetor de componentes randômicos binários (0 e 1), que se refere a variáveis não selecionadas e selecionadas. Seu mecanismo funciona a partir da seleção de variáveis com maior capacidade e alto desempenho de sobreviver quando submetidos a operações de combinação e mutação, as quais fornecem aleatoriedade ao algoritmo (BROWN; TAULES; WALKZAK, 2009).

O funcionamento do algoritmo genético coincide com o processo evolutivo dos seres vivos, seguindo o princípio que as informações são codificadas de maneira análoga aos cromossomos biológicos (composto por genes responsáveis pelas características dos indivíduos) (NUNES, 2008).

Logo, o GA interpreta as informações em cinco passos. O primeira é a codificação das variáveis que são semelhantes aos cromossomos e que são representadas através da codificação binária (GOMES et al., 2013). Assim, trazendo uma abordagem para a química analítica, para solucionar problemas de seleção de variáveis, cada comprimento de onda (variável) de um espectro é caracterizado por um gene que pode ser selecionado ou não selecionado. Assim, o tamanho do cromossomo equivale a todos os comprimentos de onda do espectro selecionado (KONZEN, 2003).

A segunda etapa é a geração de população inicial, que executa o algoritmo de forma randômica para o desenvolvimento de uma população de cromossomos (indivíduos) que constitui possíveis soluções do problema, no qual, a geração de indivíduos é completamente aleatória, sendo definida na próxima etapa de execução do algoritmo (GOMES et al., 2013).

Avaliar a aptidão é a terceira etapa e está relacionada com a capacidade de sobrevivência na teoria da evolução. Essa etapa representa a capacidade de gerar melhores respostas, deste modo, para a seleção de variáveis, quanto maior a aptidão menor o erro de predição, sendo o motivador pela ausência ou presença de uma determinada característica do indivíduo (GOMES et al., 2013). Portanto, sendo a etapa mais importante do GA, pois os indivíduos são selecionados como aptos na população para gerar novos cromossomos (indivíduos) aptos para a próxima geração (NUNES, 2008).

Os próximos passos estão relacionados ao cruzamento e mutação. No algoritmo o cruzamento é realizado entre material genético de diferentes indivíduos, a partir do cruzamento aleatório novos indivíduos são formados. Já, a mutação cria novos indivíduos através da troca aleatória de algumas características genéticas do indivíduo (GOMES et al., 2013).

Portanto, para que o GA seja executado de forma satisfatória deve-se fazer uma escolha adequada da codificação dos cromossomos em conjunto com a medida de aptidão. Diante disto, a escolha correta da codificação dos indivíduos resultará em maiores número de soluções para o problema, caracterizando quais cromossomos oferecem as soluções mais satisfatórias (PONTES, 2009). Para o funcionamento do algoritmo é muito importante que os parâmetros sejam bem definidos para não ocasionar um efeito indesejável no processo e refletir nos possíveis resultados (NUNES, 2008).

O GA é muito utilizado na área de química. Todavia, a aplicabilidade é mais difundida em técnicas analíticas, como na espectroscopia, pelo fato de aplicações de calibração multivariada envolver um grande número de variáveis (SANTOS, et al., 2022). Dessa forma, os algoritmos genéticos são indispensáveis para solucionar problemas de otimização, escalonamento e alocação de tarefas e seleção de rotas.

3.3.2.2 Algoritmo das Projeções Sucessivas (SPA)

O algoritmo das projeções sucessivas (SPA, do inglês *successive projections algorithm*) resolve problemas de colinearidade em regressão linear múltipla, construindo subconjunto de variáveis que sejam minimamente correlacionadas entre si (SOARES et al., 2013). Para atingir um número N (forward selection) o algoritmo vai inserindo variáveis com maior projeção ortogonal uma a uma. Além disso, o SPA tem sido utilizado para solução de problemas associados ao LDA, MLR e PLS, uma vez que os dois primeiros também se propõem a resolver problemas de multicolinearidade (GALVÃO et al., 2008).

Contudo, o algoritmo SPA realiza operações para obter subconjuntos de variáveis que apresentem menor colinearidade (XIAOBO et al., 2010). A seleção inicia com apenas um comprimento de onda e vai inserindo novos comprimentos, de modo a escolher os comprimentos de onda com menor repetição (ARAÚJO et al., 2001).

Basicamente, o algoritmo é processado em três etapas. A primeira é a geração das cadeias de variáveis que verifica os subconjuntos de variáveis com menor colinearidade. No segundo processo é a verificação das cadeias geradas que analisa a etapa anterior, a partir da capacidade preditiva do modelo, através dos valores de RMSEv ou RMSEcv, a depender do tipo de validação utilizada. Finalmente, na etapa de eliminação, acontece a exclusão de variáveis que não contribuí significativamente para o modelo. Assim, comparando os valores de RMSE (Raiz Quadrada do Erro Médio) com o RMSE mínimo alcançado por meio de um teste, no qual, são selecionadas variáveis que possuem valores de RMSE maiores que o RMSE mínimo (GALVÃO et al., 2008).

Neste contexto, para a fase de avaliação na classificação a função de custo (equação 7) é utilizada como critério para orientar a escolha das variáveis (SOARES et al., 2013):

$$J_{cost} = \frac{1}{N_{val}} \sum_{n=1}^{N_{val}} g_n \quad \text{Equação 7}$$

Onde g_n é o risco médio de classificação incorreta para o conjunto de validação, sendo calculado pela razão do quadrado da distância de Mahalanobis expressa na equação 8 (SOARES et al., 2013):

$$g_n = \frac{MD^2 [x_{val,n}, \bar{x}(I_n)]}{\min_{I_j \neq I_n} MD^2 [x_{val,n}, \bar{x}(I_j)]} \quad \text{Equação 8}$$

Em que MD^2 corresponde ao quadrado da distância de Mahalanobis, considerando a n ésima amostra de validação da classe I e o valor médio da sua classe verdadeira $\bar{x}(I_n)$; $\min_{I_j \neq I_n} MD^2 [x_{val,n}, \bar{x}(I_j)]$ corresponde ao quadrado da distância de

Mahalanobis entre a i -ésima amostra de validação da classe I e o centro da classe incorreta mais próxima, I_j .

Entretanto, a função faz a relação entre a distância da amostra n da classe I com relação ao valor médio das classes correta e incorreta. Isto é, avaliando a classificação verdadeira da classe certa em comparação com a classe errada mais próxima. Assim, quando o valor de g_n é pequeno representa que a amostra (X_{val},n) está próxima do centro de sua classe verdadeira e distante das demais (SOARES et al., 2013).

3.3.2.3 Stepwise (SW)

O Stepwise é um método que seleciona subconjuntos de variáveis visando o melhor desempenho para o modelo, seja de classificação ou regressão. Isto é, o mesmo pode adicionar (*forward selection*) ou remover (*backward elimination*) uma variável por vez, como também fazer combinações com os dois processos para a construção do melhor modelo (BROWN, TAULER, R & WALCZAK, 2009).

Este algoritmo considera o valor individual de cada variável espectral conforme o fator de discriminação (D_i) entre as classes em consideração. A discriminabilidade (D_i) pode ser quantificada conforme a equação 9 (BROWN, TAULER; WALCZAK, 2009).

$$D_i = \frac{S_{Bi}}{S_{wi}} \quad \text{Equação 9}$$

Em que S_{Bi} e S_{wi} são as medidas de dispersão entre e dentro das classes. Dessa forma, dispersão entre as classes é definida pela seguinte equação:

$$S_{Bi} = \sum_{j=1}^c n_j [n_j [m_{ij} - m_i]^2] \quad \text{Equação 10}$$

Onde, onde m_i é a média de x_i sobre todas as amostras de treinamento. E m_{ij} é o valor médio de x_i na classe j , calculado conforme a equação 11:

$$m_{ij} = \frac{1}{n_j} \sum_{K \in I_j} x_i^k \quad \text{Equação 11}$$

Já a dispersão entre classes S_{wi} é definida por:

$$S_{wi} = \sum_{j=1}^c S_{ij} \quad \text{Equação 12}$$

onde s_{ij} é a dispersão de x_i na classe j , sendo calculada, conforme:

$$S_{ij} = \sum_{K \in I_j} [x_i^k - m_{ij}]^2 \quad \text{Equação 13}$$

Em que x_{ik} denota o valor de x_i no k -ésimo objeto e m_{ij} é o valor médio de x_i na classe j .

A variável x_i é selecionada a partir da maior discriminabilidade D_i para a realização da validação cruzada (*leave – one- out*), para obter o número de erros. Deste modo, as variáveis que possuem maior correlação com as variáveis selecionadas são retiradas para evitar problemas com a colinearidade. O conjunto de variáveis que apresentar um resultado com menor número de erros na validação cruzada é usado para construir o modelo de classificação, ou seja, o algoritmo repete o processo com todas as variáveis até que o menor erro seja escolhido (PONTES, 2009).

4 MATERIAL E MÉTODOS

A parte experimental descrita é de autoria de Souza (2011). O foque da metodologia deste trabalho é a realização do tratamento de dados.

4.1 Amostras

As amostras de comprimidos de Captopril (25 mg) foram produzidas pelo Laboratório Farmacêutico de Pernambuco (Lafepe). Neste trabalho foi utilizado 23 lotes, sendo que 400 comprimidos pertencentes ao mesmo lote e 220 comprimidos pertencentes aos 22 lotes distintos (10 comprimidos de cada lote). Inicialmente, usou uma amostra de 120 comprimidos (lote 10041492) para realização de ensaios físico-químicos de qualidade, conforme a monografia da Farmacopeia Brasileira 6ª edição. Após esta etapa inicial os comprimidos foram analisados antes do processo de degradação (tempo zero) e no último dia do processo de degradação (184 dias). Portanto, para a realização deste estudo os seguintes ensaios físico-químicos de qualidade foram realizados: determinação do peso, dureza, friabilidade, teste de desintegração, uniformidade de doses unitárias e teste de dissolução.

Segundo a Farmacopeia Brasileira (2019), para determinar o peso de comprimidos deve-se pesar 20 comprimidos individualmente e determinar o peso médio. Assim, devendo considerar $\pm 7,5\%$ do limite de variação para comprimidos com peso médio de 80 a 250 mg. Para o parâmetro da dureza o comprimido é submetido à ruptura ou esmagamento sob pressão radial, sendo a dureza inversamente proporcional a porosidade e proporcional à força de compressão. A dureza deste comprimido em estudo foi determinada pelo durômetro Vankel VK-200 da Varian, no qual 10 comprimidos foram submetidos pelo equipamento.

4.2 Degradação dos comprimidos de Captopril

O estudo de degradação teve a duração de seis meses. Do quantitativo de comprimidos produzidos pela LAFEPE (620), do lote 10041492, 200 comprimidos foram preservados dentro de sua embalagem primária e 200 comprimidos foram distribuídos em placa de petri (sem tampa) para ficarem expostos às condições de degradação. Todos os comprimidos foram submetidos à degradação forçada utilizando uma câmara de estabilidade acelerada (Nova Ética Modelo 420/CLD-300, à $40^{\circ}\text{C} \pm 2^{\circ}\text{C}$ e $75\% \text{UR} \pm 5\% \text{UR}$).

4.3 Instrumentação analítica

As medidas espectrais de reflectância difusa foram adquiridas por meio o espectrofotômetro FTIR, Bomem Modelo FTLA 2000-160 (**figura 1**), com faixa espectral de 14000 a 3800 cm^{-1} , resolução de 8 cm^{-1} e 32 varreduras e com acessório para reflectância difusa Powder Sampl/R. Antes do processo de degradação dos comprimidos, espectros Nir de 10 unidades de comprimidos foram registrados, sendo o registro feito a cada 7 dias no primeiro mês e a cada 15 dias nos meses restantes.

Figura 3: Bomem Modelo FTLA 2000-160 com acessório para reflectância difusa Powder Sampl/R.



Fonte: Adatado de Souza(2011).

Conforme descrito no manual da Farmacopeia Brasileira os comprimidos devem ser submetidos ao ensaio de doseamento pelo método de cromatografia líquida de alta eficiência. Assim, de forma simultânea aos registros dos espectros NIR utilizou-se um cromatógrafo Shimadzu Class-VP V6.14 SP2 com detector ultravioleta de 220 nm com coluna de 250 nm e 4,6 mm de diâmetro interno, empacotada com sílica quimicamente liada a grupo ostadecilsilano (5 μ m) da Shimadzu. Como fase móvel a mistura de ácido fosfórico 0,11% (v/v) e metanol (45:55).

4.4 Solução Padrão de dissulfeto de captopril

Utilizando uma balança analítica Mettler AE 260 foi preparada uma solução padrão de dissulfeto de captopril na concentração de 1mg/mL em fase móvel. O dissulfeto de captopril utilizado foi produzido pela FAFAR-UFMG (Faculdade de Farmácia da Universidade Federal de Minas Gerais), lote 120108. A partir dessa solução foi preparada a solução teste do padrão de dissulfeto de captopril, retirando uma alíquota de 3 mL e transferindo para um

balão volumétrico de 100 mL e o volume completado com fase móvel, chegando a concentração final de 30 µg/mL.

4.5 Solução Padrão de captopril

A solução padrão de captopril foi preparada pesando-se 100 mg da substância química de referência (SQR) de captopril da Farmacopeia Brasileira – FB, lote 1001 (lote corrente – 2011) e transferida para um balão volumétrica de 100 mL. Foi adicionada uma alíquota de 3 mL da solução padrão de dissulfeto de captopril (1mg/mL) e o volume completo com fase móvel. As amostras foram preparadas em triplicatas. Os comprimidos foram pesados, pulverizados e foi transferido o equivalente a um peso médio do comprimido para balão volumétrico (25 mL) e adicionado 15 mL de fase móvel e posteriormente levado à agitação em banho de ultra-som por 15 minutos e agitação mecânica por mais 15 minutos. Após esta etapa o volume foi completado, homogeneizado e filtrado.

As amostras e as soluções padrões foram colocadas em vials e levadas ao cromatógrafo onde foram injetados separadamente 20 µL de cada. Os cromatogramas foram registrados suas áreas medidas pelos picos obtidos. O teor de captopril e dissulfeto de captopril presentes nas amostras foram calculados em relação aos valores obtidos das soluções padrões de captopril e dissulfeto de captopril, os quais apresentam valores de especificação de no mínimo 90% e no máximo 110% para o captopril e para o dissulfeto de captopril o limite máximo é de 3% segundo a Farmacopeia Brasileira.

4.6 Tratamento dos dados

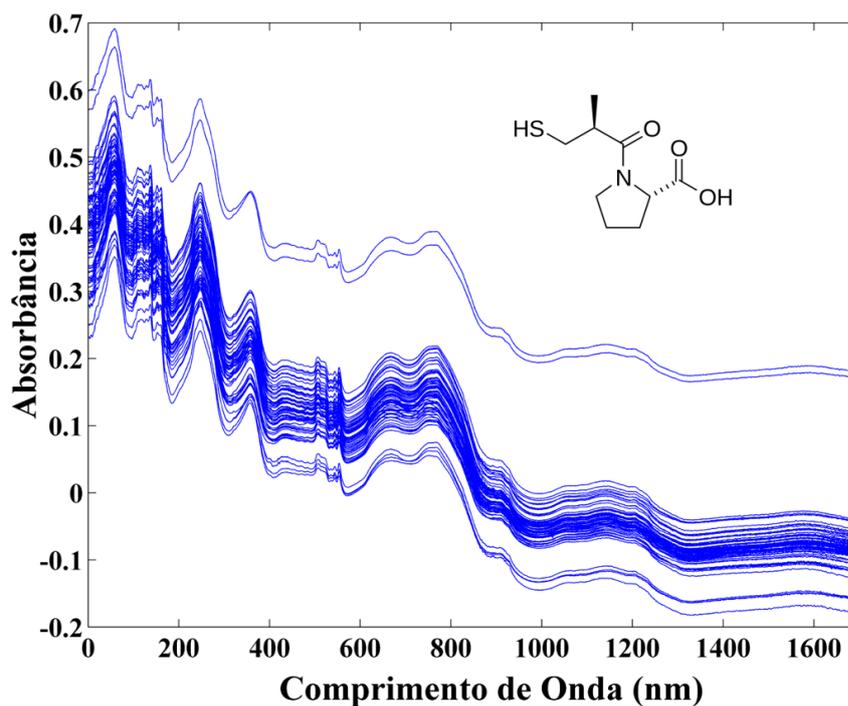
Os dados espectrais foram tratados utilizando o software MATLAB® da marca Eigenvector.

5 RESULTADOS E DISCUSSÃO

5.1 Espectros NIR

A **Figura 4** apresenta os espectros NIR das amostras de comprimidos de captopril na faixa de 1000 – 2.500 nm.

Figura 4: Espectros originais



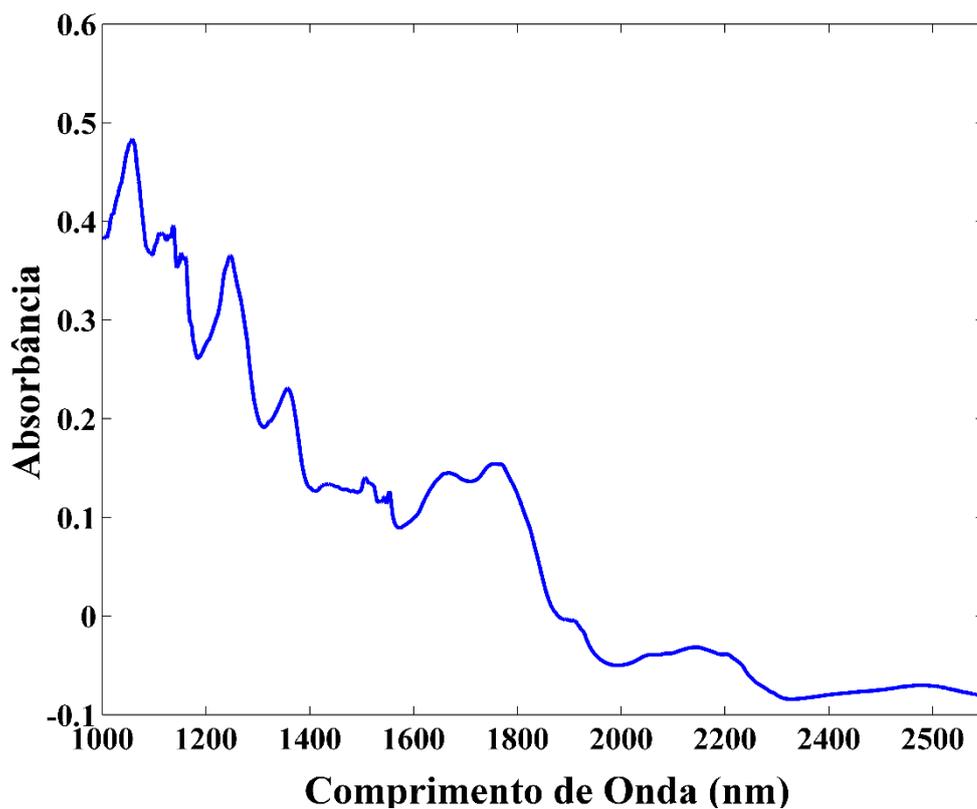
Fonte: Elaborada pela autora, 2022.

Conforme a figura 4 é possível observar, os espectros originais apresentam duas amostras com uma absorbância e linha de base maior que as outras.

Os sobretons e sobretons de combinação observados nos espectros NIR são devidas as ligações C-H, N-H e O-H (PAIVA; LAMPMAN; VYVYAN, 2013). O intervalo de 1000-1300 nm são comprimentos de onda informativos para captopril que corresponde as ligações de C-H. Nas regiões entre 1300- 1600 nm é mais informativo para o primeiro sobreton de N-H e O-H, como pode ser observado na molécula de captopril.

Para melhor visualização do espectro a figura 5 apresenta o espectro médio na região do infravermelho próximo.

Figura 5. Espectros médio NIR das amostras de captopril.



Fonte: Elaborada pela autora, 2022.

5.2 Modelos de Calibração

Os modelos de calibração foram construídos relacionando os espectros originais e pré-processados com as concentrações de captopril proposto (capítulo 3).

A Tabela 1 apresenta os resultados obtidos pelo PLS, Ispa/PLS, MLR/SPA, MLR/SW e MLR/GA para os conjuntos de calibração na determinação de captopril para os diferentes pré-processamentos.

Tabela 1: Resultados obtidos para os conjuntos de calibração na determinação de captopril para os diferentes pré-processamentos.

CAPTOPRIL -						
Pré-Processamento	Parâmetros	PLS	iSPA/PLS	MLR/SPA	MLR/SW	MLR/GA
Originais	r_{cv}	0,9722	0,9065	0,9417	0,9586	0,9828
	RMSECV	0,0525	0,1340	0,1068	0,0904	0,0587
Correção de linha de base <i>offset</i>	r_{cv}	0,8840	0,9935	0,9589	0,9904	0,9820
	RMSECV	0,0527	0,0376	0,0940	0,0459	0,0626
2º Derivada S.G [11*]	r_{cv}	0,9820	0,9961	0,9550	0,9986	0,9875
	RMSECV	0,0087	0,0294	0,0992	0,0179	0,0526
2º Derivada S.G [41*]	r_{cv}	0,9832	0,9961	0,9550	0,9986	0,9768
	RMSECV	0,0400	0,0345	0,0992	0,0179	0,0717
SNV	r_{cv}	0,8708	0,8661	0,9438	0,9865	0,9813
	RMSECV	0,1123	0,1683	0,1041	0,0516	0,0606

*Número de pontos empregados na derivação Savitzky-Golay

Fonte: Elaborada pela autora, 2022.

Com base nos resultados apresentados na Tabela 1 para determinação de captopril, é possível perceber que diferentes pré-processamentos apresentaram resultados satisfatórios para os modelos empregados, com baixos valores de RMSECV. Sendo assim, o pré-processamento utilizando a 2ª derivada de Savitzky-Golay com janela de 11 pontos foi escolhido por apresentar menores valores de RMSECV e maiores valores de r_{cv} . Os resultados mais detalhados para determinação de captopril são apresentados nas próximas seções.

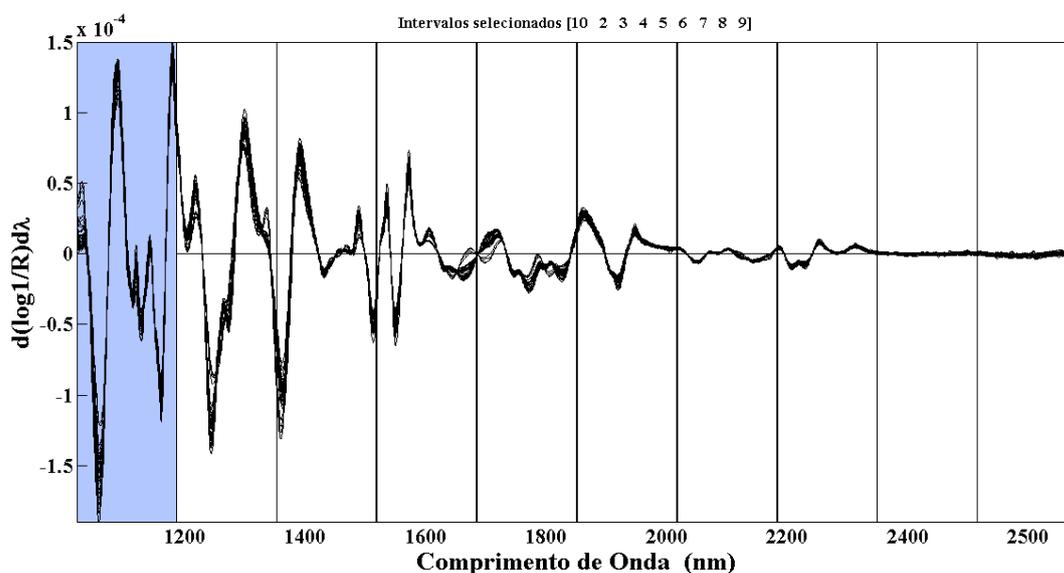
5.2.1. Avaliação dos modelos de Calibração desenvolvido

Nesta seção só serão apresentados os modelos utilizando seleções de variáveis e/ou intervalos que estão concatenados com o objetivo deste trabalho.

5.2.1.1. iSPA/PLS

A figura 6 apresenta o procedimento de seleção de intervalos utilizando o iSPA/PLS para captopril. Os comprimentos de onda estão no eixo horizontal e correspondem aos limites do lado esquerdo dos 10 intervalos espectrais determinados. Cada barra do gráfico indica o intervalo obtido pelo modelo PLS local para o intervalo correspondente. O número de variáveis latentes para cada modelo foi de 10 variáveis. O intervalo que apresentou menor RMSECV obtido corresponde aos comprimentos de onda na faixa de 1000-1200 nm conforme está mostrando na cor azul da figura.

Figura 6. Intervalos de variáveis selecionadas pelo Ispa/PLS.



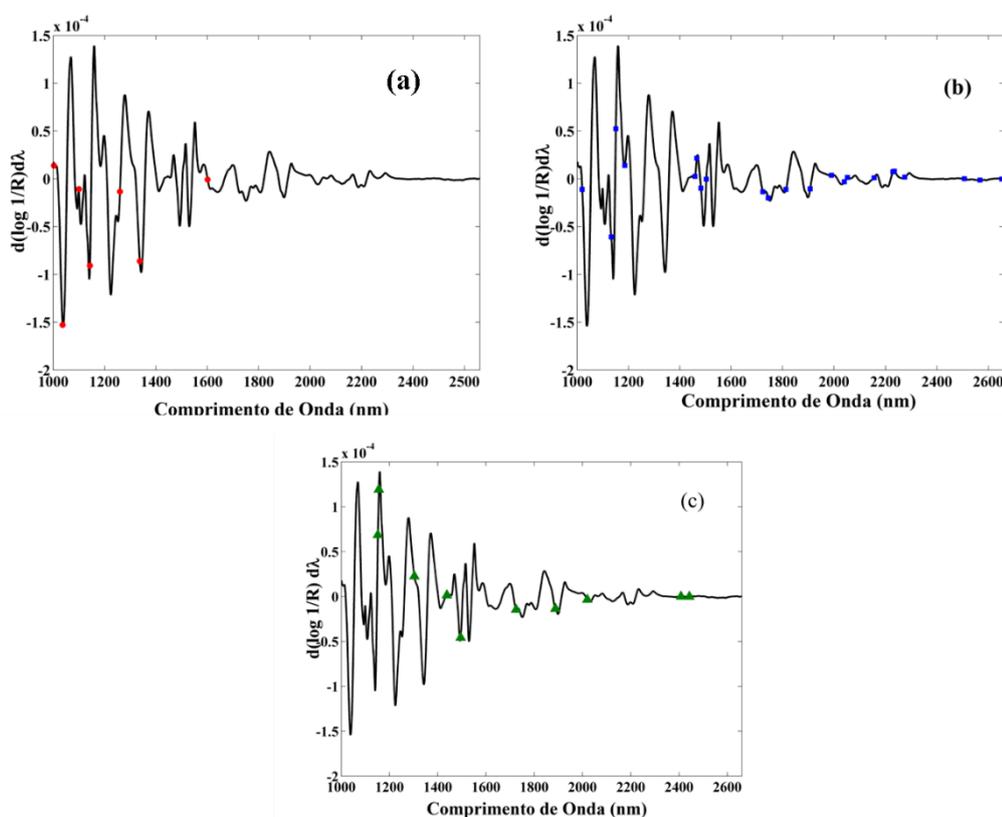
Fonte: Elaborada pela autora, 2022.

Como pode ser observado, a região selecionada pelo iSPA/PLS é bastante informativa conforme foi apresentada na seção 4.1.

5.2.1.2. Seleções de Variáveis

A figura 7 mostra as variáveis selecionadas pelo SPA, SW e GA para o modelo MLR para o captopril.

Figura 7. Espectros NIR médios do conjunto de dados com indicação das variáveis selecionadas pelos algoritmos SPA (●), SW (■) e GA (▲), para captopril, a,b e c respectivamente.



Fonte: Elaborada pela autora, 2022.

Como observado na Figura 7 para o captopril, todos os modelos selecionaram variáveis na região de 1000-1600 nm, porém o MLR/SPA e o MLR/GA selecionaram menos variáveis e em locais mais informativos, que corroboram com a região selecionada pelo iSPA/PLS.

5.2.2. Análise dos modelos de predição

Os resultados de predição dos modelos de regressão iSPA/PLS e MLR com seleções de variáveis desenvolvidos para determinar o teor de captopril usando a espectroscopia NIR estão resumidos na Tabela 2.

Tabela 2. Resultados obtidos pelos modelos iSPA/PLS, MLR/SPA, MLR/SW e MLR/GA para o conjunto de predição em termos de RMSEP, valores de correlação e de determinação. N é o número de amostras de predição.

Tabela 2. Resultados obtidos pelos modelos iSPA/PLS, MLR/SPA, MLR/SW e MLR/GA para o conjunto de predição em termos de RMSEP, valores de correlação e de determinação. N é o número de amostras de predição.

CAPTOPRIL						
Modelos	N	RMSEP (mg)	r_{pred}	R²	Rep (%)	VS ou VL*
iPLS	28	0,08	0,96	0,92	23,08	10
MLR-SPA	28	0,015	0,85	0,65	40,59	7
MLR-SW	28	0,03	0,99	0,98	8,19	22
MLR-GA	28	0,11	0,93	0,80	30,81	10

Fonte: Elaborada pela autora, 2022.

*VS: Variáveis

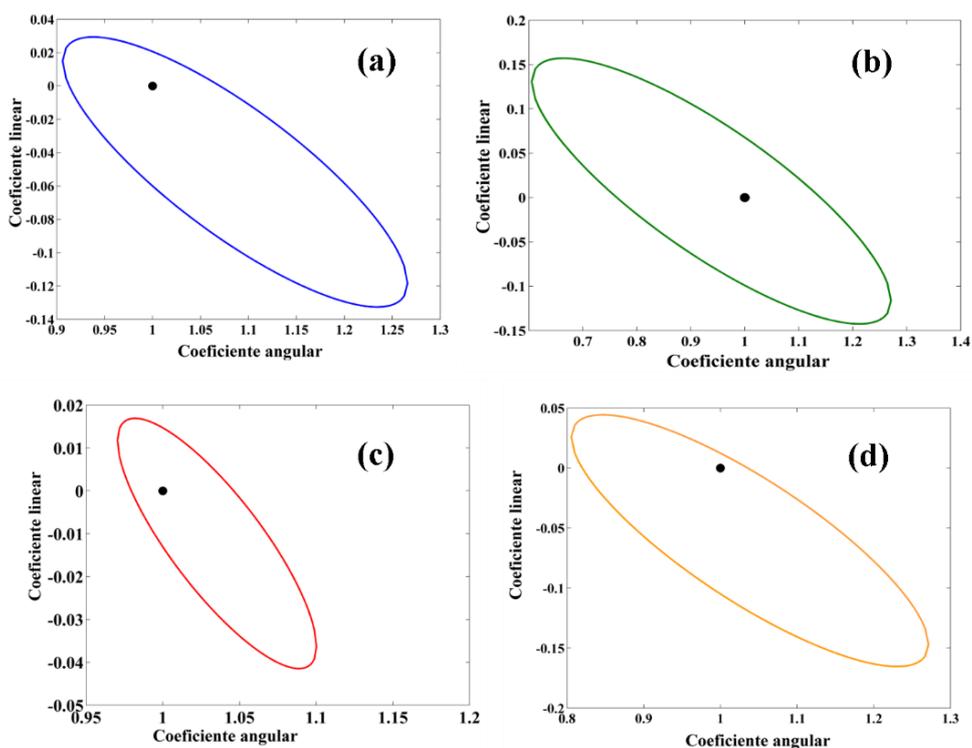
Selecionadas/ VL: Variáveis Latentes

Na comparação entre os modelos PLS (SOUZA, 2011) e iSPA/PLS para os dados estudados, os valores de RMSEP e REP, apresentaram melhores resultados semelhantes para os modelos com RMSEP 0,08.

Para os modelos MLR utilizando seleções de variáveis aplicados aos medicamentos, resultados satisfatórios foram encontrados para o SW e GA que apresentaram valores aproximados ao iPLS. Comparando o SPA com o iPLS apenas duas variáveis foram selecionadas, uma sendo indicada na região selecionada pelo iPLS, o que torna o modelo mais simples. O modelo MLR/SW apresentou um menor valor de RMSEP e maiores valores para r de correlação e R^2 quando comparado aos outros modelos desenvolvidos.

A acurácia do modelo foi avaliada por meio do teste EJCR (Figura 8), que permite a inclusão dos pontos 0 e 1 na região de confiança da articulação elíptica da inclinação e no intercepto a ser investigado.

Figura 8. Elipse de confiança para o captopril obtidas para os modelos iSPA/PLS (linha azul), MLR/SPA (linha verde), MLR/SW (linha vermelho) e MLR/GA (linha laranja).

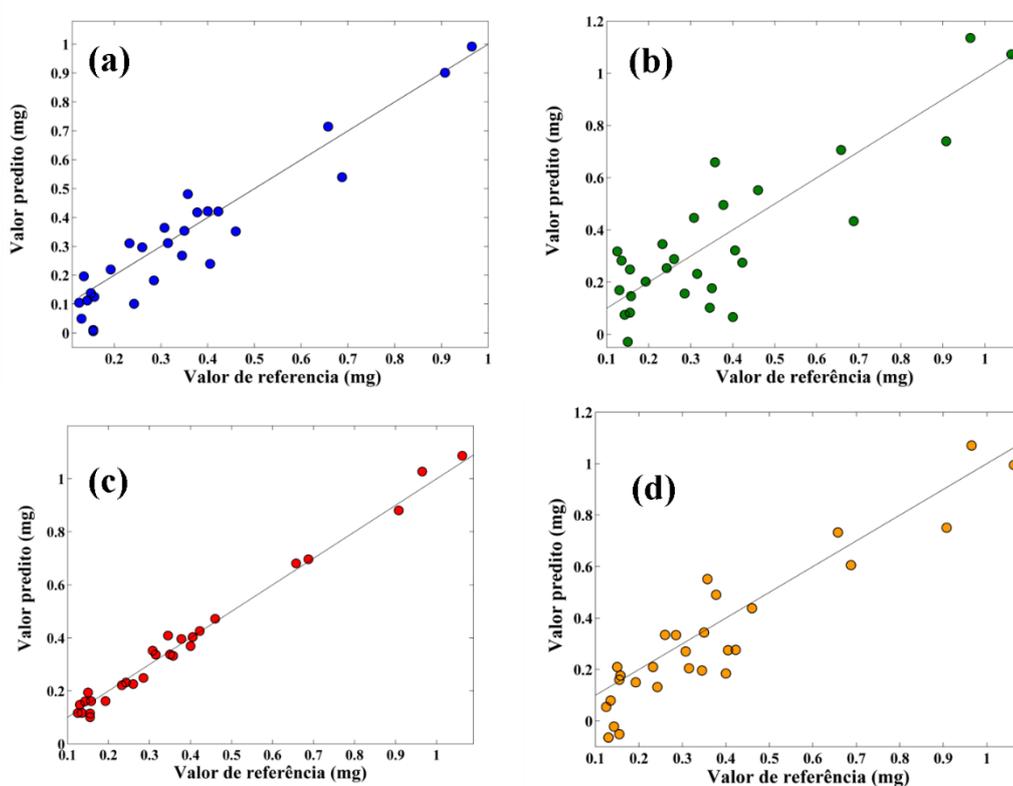


Fonte: Elaborada pela autora, 2022.

A elipse aplicada para todos os modelos propostos mostrou que os modelos estudados não apresentam diferenças estatisticamente significativas entre os valores de referência e os valores preditos pelo modelo. Este fato é indicativo de que erros proporcionais e constantes não estejam presentes e, assim, a capacidade preditiva dos modelos seja satisfatória.

A figura 9 apresenta os valores preditos *versus* os valores de referência para os modelos estudados.

Figura 9. Gráfico dos valores preditos versus os valores de referências obtidos pelos modelos iSPA/PLS (a), MLR/SPA (b), MLR/SW (c) e MLR/GA (d).



Fonte: Elaborada pela autora, 2022.

6 CONCLUSÕES

Neste trabalho foram desenvolvidas metodologia baseada na espectroscopia NIR para estudo de calibração multivariada de produto de degradação em captopril.

Os modelos de calibração multivariada baseados em iSPA/PLS, MLR/SPA, MLR/SW e MLR/GA foram desenvolvidos, validados e comparados entre si. Os modelos MLR/SW apresentou um melhor desempenho com um RMSEP de 0,03, r de correlação 0,99 e um R^2 de 0,98, essa técnica emprega seleção de variáveis, onde tem-se uma reafirmação que o uso de seleções de variáveis frente ao uso do espectro completo na aplicação de modelos de calibração se torna mais viável, uma vez que se trabalha com um número restrito de variáveis, porém mais informativa.

Como vantagens, a metodologia proposta e desenvolvida neste trabalho apresentaram bons desempenhos para quantificar produto de degradação em captopril. Tais característica é relevante para que se realize esta metodologia, por órgãos de vigilância, o controle da qualidade de captopril no mercado.

REFERÊNCIAS

Alves, Júlio César L.; Poppi, Ronei J. **Determinação do teor de biodiesel em misturas de óleo diesel usando espectroscopia de infravermelho próximo (NIR) e máquinas de vetores de suporte (SVM)**. *Talanta*, v. 104, pág. 155-161, 2013.

ANVISA – Agência Nacional de Vigilância Sanitária - **Guia para validação de métodos analíticos e bioanalíticos**. Resolução - RE nº 899, de 29 de maio, 2003. Disponível pelo site: http://anvisa.gov.br/legis/resol/2003/re/899_03re.htm. Acesso em: 02 de novembro de 2022.

Araújo, C.G.S. **Fisiologia do exercício físico e hipertensão arterial: uma breve introdução**. *Rev. Soc. Bras. Hipertensão* 4, 2001. Acesso pelo site: http://www.sbh.org.br/revista/2001_N3_V4/index.asp, em 3 de maio de 2004.

Araújo, M.C.U.; Saldanha, T.C.B. ; Galvão, R.K.H.; Yoneyama, T.; Chame, H.C.; Visani, V. **The successive projections algorithm for variable selection in spectroscopic multicomponent analysis**. *Chemometrics and Intelligent Laboratory Systems*, 57, 2001, 65–73.

Barroso, Weimar Kunz Sebba et al. **Diretrizes Brasileiras de Hipertensão Arterial**. 2021. Arquivos Brasileiros de Cardiologia, v. 116, p. 516-658,2021.

Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. **Chemometrics: a practical guide**. New York: John Wiley and Sons, Inc., 1998.

Brasil. **Agência Nacional de Vigilância Sanitária**. Provedor. Brasília. [Acesso em: 19 nov. 2022]. Disponível em: <<http://www.anvisa.gov.br/Inspeção/Fiscalização>>

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

Bro, R. **Multivariate calibration. What is in chemometrics for the analytical chemistry?** *Anal. Chim. Acta* 500, 2003,185.

Brown, S.D.; Tauler, R.; Walkzar, B. **Comprehensive Chemometrics-Chemical and biochemical data analysis**, Elsevier: Broomfield, 2009.

Brunton, Laurence L.; KNOLLMANN, Björn C. **As Bases Farmacológicas da Terapêutica de Goodman e Gilman-13**. Artmed Editora, 2018.

Cirovic, D. A.; Brereton, R. G.; Walsh, P. T.; Ellwood, J. A.; Scobbie, E. **Application of Partial Least Squares Calibration to Measurements of Polycyclic Aromatic Hydrocarbons in Coal Tar Pitch Volatiles**. *Data Process.*, vol. 121, p. 575–580, 1996.

Conselho Nacional de Secretários de Saúde – CONASS. **Assistência farmacêutica no SUS**. Brasília: Conselho Nacional de Secretários de Saúde; 2007.

Connors, K. A.; Amidon, G. L.; Stella, V. J.; **Chemical stability of pharmaceuticals: a handbook for pharmacists**. 2nd ed., WileyInterscience: New York, 1986.

Ferreira, M. M. C.; **Quimiometria: Conceitos, Métodos e Aplicações**. Ed. UNICAMP: Campinas, 2015.

Galvão, R. K. H.; Araújo, M. C. U.; Fragoso, W. D.; Silva, E. C.; José, G. E, Soares, S. F. C.; Paiva, H. M. **Variable elimination method to improve the Parsimony of MLR Models Using the Successive Projections Algorithm**. *Che Intell Lab Sys.* 92 83-91, 2008.

Geladi, P.; Kowalski, B. R., “**Partial least-squares regression: a tutorial**”, *Anal. Chim. Acta*, vol. 185, p. 1–17, 1986.

Gomes, A. A; Galvão, R. k. H.; Araújo, M. C. U.; Vêras, G.; Silva, E. C. **The sucessive projections algorithm for interval selections in PLS**. *Microchemical Journal* 100: (2013) 202-208.

Hair, J. F., Anderson, R. e Tatham, R.: BLACK WC. **Análise Multivariada de Dados**. Tradução Adonai Schlup SantAnna e Anselmo Chaves Neto, 6ª ed., 2009.

JACKSON, M.; MANTSCH, H.H.; **Pathology by infrared and Raman spectroscopy. Handbook of Vibrational Spectroscopy**, vol. 4, pp. 5-6, 2001.

Konzen. *et al.* **Otimização de métodos de controle de qualidade de fármacos usando algoritmo genético e busca de tabu**. *Pesqui. Operacional* 23, (2003) 189–207.

Kumar, N. et al. **Chemometrics tools used in analytical chemistry: An overview**. *Talanta*, v. 123, p. 186-199, June 2014.

Lamim, R. **Quitosana e n-carboximetilquitosana: desenvolvimento e biofilmes para aplicações farmacêuticas**. 2006. Dissertação (Mestrado em Ciências Farmacêuticas) – Universidade do Vale do Itajaí, Itajaí.

Leardi, R.; Gonzalez, A. L. **Genectic algorithms applied to feature selection in PLS regression: how and when to use them**. *Chemom. Intell. Lab. Syst.*41: 195-207, 1998.

Lira, L. de F. B. **Desenvolvimento de Métodos Analíticos para Monitoramento da Qualidade do Biodiesel e suas Misturas**. Recife, Programa de Pós-Graduação em Química, UFPE. 2010. Tese de doutorado, 124p.

Maluf et al. **Determinação de sibutramina em formas farmacêuticas através de espectroscopia no infravermelho com refletância difusa e métodos de calibração multivariada**. *Quim. Nova*, Vol. 33, No. 3, 649-652, 2010.

Marcatto, A. C. **Análise de cápsulas de captopril manipuladas em farmácias**. *Rev. Ciênc. Farm. Básica Apl.*, v. 26, n. 3, p.221-225, 2005.

Marcolino Júnior, L. H. **Eletrodos voltamétricos e amperométricos para a determinação de espécies de interesse farmacêutico**. 2007. 138f. Tese (Doutorado e Ciências). Universidade Federal de São Carlos, São Carlos.

Mengue SS, Bertoldi AD, Boing AC, NUL Tavares, da Silva Dal Pizzol T, Oliveira MA, et al. **Pesquisa Nacional sobre Acesso, Utilização e Promoção do Uso Racional de Medicamentos (PNAUM): métodos do inquérito domiciliar.** *Rev Saude Publica.* 2016.

Moro, M.K., dos Santos, F.D., Folli, S.G., Romão, W., Filgueiras, P.R. **A review of chemometrics models to predict crude oil proprieties from nuclear magnetic resonance and infrared spectroscopy.** *Fuel*, v. 303, n. 121283. 2021. DOI: 10.1016/j.fuel.2021.121283.

Naes, T.; Isaksson, T.; Fearn, T.; Davies, T. **A User-Friendly Guide to Multivariate Calibration and Classification.** Chichester: NIR Publications, 2002.

Natesch, R.; Schwager, S. L. U.; Evans, H. R.; Sturrock, E. D.; Acharua. R. **Structural Details on the Binding of Antihypertensive Drugs Captopril and Enalaprilat to Human Testicular Angiotensin I-Converting Enzyme.** *Biochemistry*, v.43, p. 8718-8724, 2004.

Nóbrega, I. M. F. **Desenvolvimento farmacotécnico e estudo de estabilidade de comprimidos à base de captopril.** 2006. 58f. Dissertação (Pós-Graduação em Ciências Farmacêuticas)- Universidade Federal de Pernambuco, Recife.

Nunes, P. G. A. **Uma nova técnica para seleção de variáveis em calibração multivariada aplicada às espectrometrias uv-vis e nir.** 2008. Tese (Doutorado em Química)-Curso de Pós Graduação em Química, Universidade Federal da Paraíba.

Oigman, W.; Fritsch, M. T. **Drogas que intervêm no sistema renina-angiotensina.** *HiperAtivo* 5 (2): 84-90, 1998.

Oliveira, T. A. D. **Identificação de fraudes em documentos por adição de texto e obliteração utilizando imagens hiperspectrais e quimiometria.** Dissertação de Mestrado, Brasília – DF. 2019.

Pasquini, Célio. **"Espectroscopia de infravermelho próximo: uma técnica analítica madura com novas perspectivas – uma revisão."** *Analytica chimica acta* 1026 (2018): 8-36.

Pontes, M. J. C. **Algoritmo das Projeções Sucessivas para Seleção de Variáveis Espectrais em Problemas de Classificação.** 2009.144p. Programa de Pós-Graduação em Química. (Tese de Doutorado) - Universidade Federal da Paraíba, João Pessoa.

Rafael K, Faria MGI. **Controle de qualidade dos comprimidos de captopril: uma breve revisão literária.** *Rev Uningá Rev.* 2013;16(2):49-53

Rodgers, J.; Schild, G.; Cui, X.; Delhom, C. **Textile Technology: Portable Color Spectrophotometer Measurements of Cotton Color in Remote Locations.** *Journal of Cotton Science*, 17:202,2013.

Santos, S. R. et al (2017). **Qualidade de vida do idoso na comunidade: aplicação da Escala de Flanargan.** *Revista Latino-americana de Enfermagem.* 10(6), 757- 764.

Santos, Grazielle R. et al. **Avanços analíticos baseados em modelos de calibração de primeira ordem e espectroscopia uv-vis para avaliação da qualidade da água: uma revisão - parte 1.** Química Nova [online]. 2022, v. 45, n. 03 [Acessado 7 Dezembro 2022], pp. 314-323. Disponível em: <<https://doi.org/10.21577/0100-4042.20170830>>. Epub 20 Maio 2022. ISSN 1678-7064. <https://doi.org/10.21577/0100-4042.20170830>.

Shenone, A. V.; Culzoni, M. J.; Galera, M. M.; Goicoechea, H. C. **Secondorder Advantage achieved by modeling excitation-emission fluorescence matrices affected by inner filter effects using a strategy which combines standardization and calibration: Reducing experimental and increasing analytical sensitivity.** Talanta, 109:107, 2013.

Silva, C. S.; Borba, F de S. L.; Pimentel, M. F.; Pontes, M. J. C.; Honorato, R. S.; Pasquini, C. **Classifications of blue pen ink using infrared spectroscopy and linear discriminante analysis.** Microchemical Journal, 109:122, 2013.

Simões, S. S. **Desenvolvimento de métodos validados para a determinação de captopril usando espectroscopia NIRR e calibração multivariada,** 2008. 83p. Tese (Doutorado em Química) – universidade Federal da Paraíba, João Pessoa.

Skoog, Douglas A., F. James Holler, and Timothy A. Nieman. **"Principles of instrumental analysis 5th edition."** Saunders College Pub. Co.: Philadelphia (1998).

Soares, S. F. C.; Gomes, A. A.; Galvão Filho, A. R.; Araújo, M. C. U.; Galvão, R. K. H. **The successive projections algorithm.** Trends in Analytical Chemistry, v. 42, 2013.

SOCIEDADE BRASILEIRA DE CARDIOLOGIA; SOCIEDADE BRASILEIRA DE HIPERTENSÃO; SOCIEDADE BRASILEIRA DE NEFROLOGIA. V diretrizes brasileiras de hipertensão arterial. São Paulo: Ministério da Saúde. 2006. Disponível: http://publicacoes.cardiol.br/consenso/2010/Diretriz_hipertensao_associados.pdf. Acesso em: 20 nov. 2022.

Souza, M. C. **Análise Estrutural de complexos de Zn com Inibidores da Enzima Carboxipeptidase – ECA.** 2015. Tese (Doutorado) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas, Juiz de Fora, 2015.

Souza, J. A. L. **Desenvolvimento de uma metodologia baseada na espectroscopia no infravermelho próximo para determinação do produto de degradação do captopril em comprimidos.** 2012. Dissertação (Mestrado). Programa de Pós-Graduação em Ciências Farmacêuticas, Universidade Federal de Pernambuco, Recife, 2012.

Tavares, M.T.; Primi, M. C.; Polli, M.C.; Ferreira, E.I.; Parise-Filho, R. **Interações fármaco-receptor: aplicações de técnicas computacionais em aula prática sobre a evolução dos inibidores da enzima conversora de angiotensina.** Quím Nova [Internet]. 2015Sep;38(Quím. Nova, 2015 38(8)). Available from: <https://doi.org/10.5935/0100-4042.20150113>

Tibola, Casiane Salete, et al. **"Economically motivated food fraud and adulteration in Brazil: Incidents and alternatives to minimize occurrence."** Journal of Food Science 83.8 (2018): 2028-2038.

Vinz, V. Esposito et al. **Handbook of partial least squares**. Berlin: Springer, 2010.

Willis, M. J.; Hiden, H. G.; Harenbach, P.; McKay, B. & Montage, G. A. **Genetic programming: an introduction and survey of applications**. The Second International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications 1997 (GALESIA 97). pp 314-319. IEEE Press, 1997.

Yang, X.. **Firefly algorithm, stochastic test functions and design optimisation**. *Int. J. Bio-Inspired Comput.* 2 (2010) 78–84.

Xiaobo, Z.; Jiewen, Z.; Povey, M.J.W.; Holmes, M.; Hanpin, M. **Variables selection methods in near-infrared spectroscopy**. *Analytica Chimica Acta*, 667, 14–32, 2010.

Zareef, Muhammad, et al. "**Uma visão geral sobre as aplicações de algoritmos não lineares típicos acoplados à espectroscopia NIR na análise de alimentos.**" *Food Engineering Reviews* 12 (2020): 173-190.

