



**UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
INSTITUTO DE QUÍMICA**

TRABALHO DE CONCLUSÃO DE CURSO

**SELEÇÃO DE VARIÁVEIS, DETECÇÃO DE OUTLIER E FIGURAS DE
MÉRITO PARA A DETERMINAÇÃO SIMULTÂNEA DE ENSAIOS DE
DISSOLUÇÃO DE COMPRIMIDOS CONTENDO ISONIAZIDA, RIFAMPICINA,
PIRAZINAMIDA E ETAMBUTOL USANDO ESPECTROSCOPIA NO
INFRAVERMELHO PRÓXIMO (NIRS) E CALIBRAÇÃO MULTIVARIADA.**

Fernanda Saadna Lopes da Costa

Orientador: Prof. Dr. Kássio Michell Gomes de Lima

Natal – RN
2012

Fernanda Saadna Lopes da Costa

SELEÇÃO DE VARIÁVEIS, DETECÇÃO DE OUTLIER E FIGURAS DE MÉRITO PARA A DETERMINAÇÃO SIMULTÂNEA DE ENSAIOS DE DISSOLUÇÃO DE COMPRIMIDOS CONTENDO ISONIAZIDA, RIFAMPICINA, PIRAZINAMIDA E ETAMBUTOL USANDO ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO (NIRS) E CALIBRAÇÃO MULTIVARIADA.

Monografia apresentada junto ao curso de Química Bacharelado da Universidade Federal do Rio Grande do Norte, como requisito obrigatório à obtenção do título de bacharela.

Orientador: Prof. Dr. Kássio Michell Gomes de Lima

Natal – RN
2012

Aos maiores responsáveis por esta conquista. Especialmente minha mãe, Ana, e meu pai Francisco, por toda dedicação, confiança e paciência.

Com amor, dedico.

AGRADECIMENTOS

- Primeiramente a Deus por me conceder determinação, saúde e sabedoria para concluir com sucesso esta jornada.
- Aos meus pais Ana Maria e Francisco Benjamim, por todo amor, companheirismo, apoio e dedicação. Essa graduação eu devo e ofereço a vocês. Essa não é a minha, mas sim a nossa vitória. Eu os amo incondicionalmente.
- Às minhas avós Maria Ester e Maria Costa, ao meu avô Antônio de Assis, às minhas tias, tios, primos e primas. À todos os outros familiares que me encorajaram e incentivaram a seguir em frente.
- Ao professor e orientador Prof. Dr. Kássio Michell Gomes de Lima, por toda a sua constante atenção, apoio, cobrança, conhecimento e, antes de tudo, pela confiança que me foi depositada.
- À minha amiga inseparável Wendy Marina, que no ensino médio acreditou e fez do que era para todos os outros uma “loucura”, um sonho e, deste sonho, realidade. Você foi, por muitas vezes, a minha coragem para continuar. Muito obrigada por ter estado sempre ao meu lado, amiga.
- À minha família GPQA: Carolina Neves, Raquel Cavalcanti, Karine Fonseca, Stéphanie Cavalcante, Gustavo Mesquita, Rosangela Costa, Rodrigo Rocha, Ricardo Pedroza, Ana Luiza, Jábine Talitta, Rafael Fernandes e Manuel Carneiro. Além de companheiros de trabalho, vocês tornaram-se muito bons amigos.
- Aos meus amigos que, direta ou indiretamente, contribuíram com a concretização deste sonho, e em especial: Thaíse, Priscilla, Jorivan, Everton, Tâmyna, Sheila e Marcos Antônio. Por toda compreensão, incentivo e carinho.
- À minha amiga Aline Marques, por ter dividido comigo nos últimos anos, os melhores e piores momentos da minha vida. Por sempre estar ao meu lado, me apoiar e acreditar na minha capacidade. Você é a melhor amiga que eu poderia ter. Eu te amo, amiga!
- Ao meu amigo Álvaro Gustavo, não tenho palavras para descrever a importância que você tem em minha vida, tudo que aprendi e muito menos para agradecer tudo que você tem feito por mim. Amo você!
- À Universidade Federal do Rio Grande do Norte, ao Instituto de Química, ao NEPPGN e ao NUPLAM.

CURRICULUM VITAE

Fernanda Saadna Lopes da Costa

E-mail: nanda_saadna@yahoo.com.br

Curriculo Lattes: <http://lattes.cnpq.br/5652600800390930>

FORMAÇÃO ACADÊMICA

2006 – 2010 Técnico em Geologia e Mineração

Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte

Orientador: Prof. MSc. Marcondes Mendes de Souza

Título: TÉCNICO

2009 – 2012 Graduação em QUÍMICA

Universidade Federal do Rio Grande do Norte

Orientador: Prof. Dr. Kássio Michell Gomes de Lima

Título: BACHARELADO

Bolsista: Iniciação Tecnológica (CNPq)

FORMAÇÃO COMPLEMENTAR

2008 - 2008 Curso de curta duração de Geoprocessamento. Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte, IFRN, Brasil.

2010 - 2010 Curso de curta duração em Inovações Tecnológicas Aplicadas às Indústrias Farmacêuticas. Universidade Federal do Rio Grande do Norte, UFRN, Brasil.

2010 - 2010 Curso de curta duração em Tratamento Sistemático de Equilíbrios. Universidade Federal do Rio Grande do Norte, UFRN, Brasil.

- 2010 - 2010** Extensão universitária em OLIMPÍADAS DE QUÍMICA. Universidade Federal do Rio Grande do Norte.
- 2011 - 2011** Curso de curta duração em Compostos Lamelares e Suas Aplicações. Universidade Federal do Rio Grande do Norte, UFRN, Brasil.
- 2011 - 2011** Espectroscopia Vibracional Aplicada a Compostos Inorgânicos. Universidade Federal do Rio Grande do Norte, UFRN, Brasil.

BOLSA OBTIDA

01/2011 – 06/2012 Iniciação Tecnológica (CNPq)

TEXTOS EM JORNAIS DE NOTÍCIAS/REVISTA

LIMA, K. M. G. ; MORAIS, S. C. ; NEVES, A. C. O. ; PORTO, D. L. ; SOARES, G. M. ; COSTA, F. S. L. Metodologia Espectroscópica Inovadora aplicada no NUPLAM-UFRN. Ciência Sempre, RN, p. 21 - 25, 01 jul. 2010.

APRESENTAÇÃO DE TRABALHOS EM EVENTOS CIENTÍFICOS

Neves, A.C.O.; Soares, G.M.; Morais, S.C.; Costa, F.S.L.; Porto, D.L.; Lima, K.M.G.; Prediction of dissolution profile of tablets for tuberculosis using NIR spectroscopy and multivariate calibration; 15th Internacional Conference on Near Infrared Spectroscopy, 2011, Cidade do Cabo, África do Sul.

ARTIGOS EM REVISTA CIENTÍFICA INDEXADA

Neves, A.C.O.; Soares, G.M.; Morais, S.C.; Costa, F.S.L.; Porto, D.L.; Lima, K.M.G.; Dissolution testing of isoniazid, rifampicin, pyrazinamide and ethambutol tablets using near-infrared spectroscopy (NIRS) and multivariate

calibration; Journal of Pharmaceutical and Biomedical Analysis, Volume 57, 5 January 2012, Pages 115-119.

Fernandes, R.S.; Costa, F.S.L.; Valderrama, P.; Março, P.H.; Lima, K.M.G.; Non-destructive detection of adulterated tablets of glibenclamide using NIR and solid-phase fluorescence spectroscopy and chemometric methods; Journal of Pharmaceutical and Biomedical Analysis, Volume 66, July 2012, Pages 85-90.

*Os conformistas transformam
fracassos em medo; os determinados
transformam derrotas em garra.*

(Augusto Cury)

RESUMO

O objetivo deste trabalho foi demonstrar como figuras de mérito multivariadas, em conjunto com detecção de outlier, métodos de regressão por mínimos quadrados parciais (PLS) e algoritmos de seleção de variáveis, tais como intervalo de mínimos quadrados parciais (iPLS), algoritmo genético (GA) e algoritmo das projeções sucessivas (SPA), podem ser utilizados para avaliar e melhorar o desempenho de espectroscopia no infravermelho próximo na determinação simultânea de isoniazida, rifampicina, pirazinamida e etambutol em produtos farmacêuticos acabados produzidos pelo NUPLAM na Universidade Federal do Rio Grande do Norte (Brasil). O coeficiente de correlação (R^2) para a determinação por cromatografia líquida de alta eficiência (CLAE) versus valores previstos (NIRS) variou de 0,94 - 0,98 para os 4 fármacos investigados. A raiz quadrada do erro médio quadrático da previsão (RMSEP) para cada parâmetro, obtido a partir de modelos PLS, PLS-iPLS e PLS-AG foram melhorados 37,18%, 8,95%, 69,15% e 6,03% para isoniazida, rifampicina, etambutol e pirazinamida, respectivamente, quando comparados aos iniciais, indicando que o método NIR é uma ferramenta eficaz e não destrutiva para a medição de dissolução dos fármacos a partir dos comprimidos. Foi feito um teste t-pareado e a um nível de 95% de confiança comprovou-se que não existe diferença estatística entre o método de referência e a técnica NIR.

Palavras-chave: NIRS; os ensaios de dissolução; seleção de variáveis; figuras de mérito.

ABSTRACT

The aim of this study was to demonstrate how multivariate figures of merit, together with outlier detection methods, partial least squares regression (PLS) and variable selection algorithms, such as interval partial least squares (iPLS), genetic algorithm (GA) and the successive projections algorithm (SPA), can be used to evaluate and improve the performance of near infrared spectroscopy for simultaneous determination of isoniazid, rifampicin, pyrazinamide and ethambutol in tablets produced by NUPLAM at the Federal University of Rio Grande do Norte (Brazil). The correlation coefficient (R^2) for determination by high performance liquid chromatography (HPLC) versus predicted (NIRS) ranged from 0.94 to 0.99 for the four drugs investigated. The root mean squared error of prediction (RMSEP) for each parameter, obtained from PLS, PLS-and PLS-GA iPLS were improved 37.18%, 8.95%, 69.15% and 6.03% for isoniazid, rifampicin, ethambutol and pyrazinamide, respectively, when compared to the initial, indicating that the method is a NIR efficient non-destructive for the measurement of drug dissolution from tablets. We conducted a paired t-test and a 95% level of confidence proved that there is no statistical difference between the reference method and NIR technique.

Keywords: NIRS, dissolution testing, variable selection; figures of merit.

SUMÁRIO

	PÁGINA
LISTAS DE ABREVIATURAS	xiii
LISTAS DE FIGURAS	xv
LISTA DE TABELAS	xvii
PREFÁCIO	xviii
1. INTRODUÇÃO	1
1.1 NIR E FÁRMACOS	1
1.1.1 Estudos de dissolução	3
1.2 MÉTODO DE REFERÊNCIA	4
1.3 MÉTODOS QUIMIOMÉTRICOS	6
1.3.1 Pré-tratamentos espectrais	7
1.3.2 Seleção de variáveis	7
1.3.3 Detecção de outliers	9
1.4 FIGURAS DE MÉRITO	10
1.4.1 Linearidade	11
1.4.2 Precisão	12
1.4.3 Exatidão	13
1.4.4 Sensibilidade	14
1.4.5 Seletividade	14

1.4.6 Sensibilidade do método: Limite de detecção (LD) e Limite de quantificação (LQ)	15
2. OBJETIVOS	16
3. EXPERIMENTAL	16
3.1 MATERIAIS	16
3.2 NIR	17
3.3 CLAE E APARELHO DE DISSOLUÇÃO	17
3.4 SOFTWARE E ANÁLISE DE DADOS	18
4. RESULTADOS E DISCUSSÃO	19
4.1 MODELO PLS	20
4.2 SELEÇÃO DE VARIÁVEIS	23
4.2.1 Isoniazida	23
4.2.2 Rifampicina	25
4.2.3 Pirazinamida	27
4.2.4 Etambutol	28
4.4 FIGURAS DE MÉRITO (FOM)	30
5. CONCLUSÃO	33
6. PERSPECTIVAS FUTURAS	34
7. REFERÊNCIAS	35

LISTA DE ABREVIATURAS

AG	Algoritmo genético
API	Ingredientes farmacêuticos ativos (<i>active pharmaceutical ingredient</i>)
APS	Algoritmo das projeções sucessivas
CLAE	Cromatografia líquida de alta eficiência
DPR	Desvio padrão relativo
EJCR	Região elíptica de confiança conjunta (<i>elliptic joint confidence region</i>)
FDC	Combinações de dose fixa (<i>fixed dose combinations</i>)
iPLS	Intervalo de mínimos quadrados parciais (<i>interval partial least squares</i>)
IV	Infravermelho
LD	Limite de detecção
LQ	Limite de quantificação
MLR	Regressão linear múltipla (<i>multiple linear regression</i>)
MSC	Correção de espalhamento multiplicativo de luz (<i>multiplicative scatter correction</i>)
NAS	Sinal líquido do analito (<i>net analyte signal</i>)
NEPPGN	Núcleo de Estudos e Pesquisas em Petróleo e Gás Natural
NIR	Infravermelho próximo (<i>near Infrared</i>)
NUPLAM	Núcleo de Pesquisa em Alimentos e Medicamentos
OSC	Correção ortogonal de sinal (<i>orthogonal signal correction</i>)
PLS	Mínimos quadrados parciais (<i>partial least squares</i>)
R ²	Coeficiente de correlação

RMSEC	Raiz quadrada da soma do quadrado dos erros da calibração <i>(root mean squared error of calibration)</i>
RMSEP	Raiz quadrada da soma do quadrado dos erros de previsão <i>(root mean squared error of prediction)</i>
S/N	Sinal–ruído
SEL	Seletividade
SEM	Sensibilidade
USP	<i>United States Pharmacopeia</i>
UV	Ultravioleta
VL	Variáveis latentes

LISTA DE FIGURAS

	PÁGINA
Figura 1 Visão esquemática do Cromatógrafo Líquido de Alta Eficiência: A) soluções, B) coluna cromatográfica, C) injetor, D) Computador para aquisição dos dados.	5
Figura 2 Espectros originais das 38 amostras.	19
Figura 3 Atribuição de bandas no espectro original de uma amostra.	20
Figura 4 Concentração medida x concentração prevista pela técnica de referência das amostras de calibração e validação para isoniazida utilizando o modelo iPLS. (○) conjunto de calibração; (*) conjunto de validação.	24
Figura 5 Concentração medida x concentração prevista pela técnica de referência das amostras calibração e validação de rifampicina utilizando o modelo AG-PLS. (○) conjunto de calibração; (*) conjunto de validação.	26
Figura 6 Concentração medida x concentração prevista pela técnica de referência das amostras de calibração e validação para pirazinamida utilizando o modelo AG-PLS. (○) conjunto de calibração; (*) conjunto de validação.	28
Figura 7 Concentração medida x concentração prevista pela técnica de referência das amostras de calibração e validação para etambutol modelo PLS utilizando. (○)	30

conjunto de calibração; (*) conjunto de validação.

Figura 8 Região elíptica de confiança conjunta para a inclinação e intercepto da regressão da densidade prevista versus medida de conjunto de validação externa utilizando os modelos iPLS, GA-PLS e PLS para: (a) isoniazida (linha tracejada verde); (b) rifampicina (linha contínua violeta); (c) pirazinamida (linha contínua azul), (d) etambutol (linha tracejada preta). A estrela vermelha marca o ponto ($a = 0$, $b = 1$) teórico.

32

LISTA DE TABELAS

	PÁGINA
Tabela 1 Resultados dos melhores modelos PLS utilizando toda a faixa espectral.	22
Tabela 2 Resultados para os modelos com seleção de variáveis na calibração e validação externa definida para isoniazida.	23
Tabela 3 Resultados para os modelos com seleção de variáveis na calibração e validação externa definida para rifampicina.	25
Tabela 4 Resultados para os modelos com seleção de variáveis na calibração e validação externa definida para pirazinamida.	27
Tabela 5 Resultados para os modelos com seleção de variáveis na calibração e validação externa definida para etambutol.	29
Tabela 6 Figuras de Mérito dos melhores modelos para: isoniazida, rifampicina, pirazinamida e etambutol.	31

PREFÁCIO

O processo de controle de qualidade na indústria farmacêutica é uma etapa fundamental para que se certifique a confiabilidade dos medicamentos. Por isso, diversos são os testes realizados, dentre eles um de importância significativa é o teste de dissolução, no qual se avalia a liberação do fármaco no organismo, o que é determinante na formulação de dosagem da droga.

O método padrão para determinação de perfis de dissolução é descrito por diversas farmacopeias. Consiste no teste de dissolução e quantificação por cromatografia líquida de alta eficiência (CLAE), entretanto exige muitas etapas, o que pode acarretar erros, gasto excessivo de reagentes, tempo de análise longo e destruição das amostras.

O presente trabalho tem como objetivo validar uma metodologia anteriormente descrita por nosso grupo de pesquisa, para realizar a determinação simultânea da dissolução de isoniazida, rifampicina, pirazinamida e etambutol em comprimidos anti-tuberculostáticos produzidos na UFRN, através da Espectroscopia no Infravermelho Próximo e calibração multivariada, que ao contrário da técnica padrão, minimiza os passos experimentais, tempo de análises, o gasto com reagentes e não destrói as amostras.

Os modelos multivariados foram feitos utilizando novas estratégias, tais como seleção de variáveis, detecção de outliers e por fim, tendo em mãos os melhores modelos para cada fármaco, foram calculadas as Figuras de Mérito multivariadas: seletividade (SEL), sensibilidade (SEN), precisão, exatidão (RMSE), linearidade (R^2), limite de quantificação, limite de detecção e relação sinal/ruído.

O cálculo destas figuras de mérito é imprescindível para a implantação desta nova metodologia na indústria farmacêutica, pois é o cálculo destes parâmetros que garante o bom desempenho e a aplicabilidade da nova técnica analítica que se deseja utilizar.

1. INTRODUÇÃO

1.1 NIR E FÁRMACOS

Formas farmacêuticas sólidas administradas por via oral são amplamente prescritas nas práticas médicas. O crescente interesse pelas formas farmacêuticas em cápsulas se deve as vantagens dos comprimidos durante a terapia medicamentosa [1]. Entretanto para que os medicamentos sejam aprovados e repassados para a população existem uma série de análises de controle de qualidade que devem ser feitas, tais como friabilidade, dureza, teste de desagregação, e entre eles um de fundamental importância é o teste de dissolução, que é considerado uma variável fundamental para descrever o processo de absorção do fármaco no trato gastrointestinal. As indústrias farmacêuticas seguem o método oficial descrito pela USP (do inglês, *United States Pharmacopeia*) na qual este parâmetro é analisado pelo teste de dissolução e quantificado pela técnica de Cromatografia Líquida de Alta Eficiência (CLAE).

Embora a aplicabilidade desta técnica seja incontestável devido a sua alta precisão, exatidão e reprodutibilidade, ainda existem algumas desvantagens, tais como uso de muitos reagentes, preparação das amostras, o fato de ser uma técnica destrutiva, o que torna necessário a busca por novas alternativas de análise.

A espectroscopia no infravermelho próximo (NIR, do inglês, *Near Infrared*) utiliza a região do espectro eletromagnético, que se situa entre as bandas visíveis e infravermelho médio. O espectro NIR é formado a partir dos sinais devido às vibrações moleculares de ligações CH, NH e OH [2,3]. Atualmente, a espectroscopia NIR é amplamente aceita pela indústria farmacêutica como um método de análise qualitativa e quantitativa para o controle dos parâmetros físico-químicos das amostras, proporcionando medição direta de amostras sólidas com uma análise rápida e precisa, necessitando de pouca ou nenhuma manipulação da amostra [4]. A utilização do novo método espectroscópico foi benéfica para as indústrias farmacêuticas, porque eliminou a etapa de diluição do analito, economizando tempo e

aumentando a produtividade. Além disso, a espectroscopia NIR pode ser usada para determinar várias propriedades físicas e químicas de medicamentos, tais como: o conteúdo, densidade, os ensaios de dissolução, dureza e de dispersão de moléculas, entre outros. Uma das vantagens mais importantes do método é a sua natureza não destrutiva, uma vez que evita passos importantes que são responsáveis para a introdução de erros e permite que o comprimido possa ser reutilizado após a medição, quando a amostra é pequena [5].

Bodson et.al (2007), utilizaram a espectroscopia NIR para fazer a validação do processo de fabricação de comprimidos de hidrocloreto de Diltiazem, medicamento pertencente ao grupo das benzotiazepinas, um tipo de bloqueadores dos canais de cálcio. A técnica foi empregada com a finalidade de diminuir o tempo consumido das operações envolvidas na produção do medicamento. O método de referência utilizado foi a espectroscopia de Ultravioleta (UV). O modelo foi validado através do cálculo de repetibilidade, precisão, linearidade e limite de quantificação. Os resultados mostraram que o NIR foi equivalente ao do método de referência, portanto sendo considerada uma ferramenta interessante no estudo do processo de produção em tempo real [6].

A espectroscopia NIR foi utilizada também por Trafford et. al. (1998) para determinação do conteúdo de comprimidos de paracetamol. A escolha da técnica levou em consideração as vantagens de ser não destrutiva e não precisar de preparação da amostra. Utilizaram-se 45 comprimidos de 500 mg, a técnica de referência foi o UV, conforme a recomendação da British Pharmacopoeia (1993). Com base em modelos MLR (do inglês, *Multiple Linear Regression*), foi obtido um modelo de calibração para apenas dois comprimentos de onda, para prever a quantidade de paracetamol no comprimido. Os resultados mais uma vez comprovaram que as duas técnicas eram equivalentes, obtendo valores de precisão e repetibilidade muito próximos entre si. [7]

1.1.1 Estudos de dissolução

Um dos parâmetros muito importantes para a indústria farmacêutica é o teste de dissolução. Dissolução é um processo dinâmico, fortemente dependente tanto da composição do meio quanto da hidrodinâmica. Como o ambiente do trato gastrointestinal luminal varia consideravelmente, é necessário medir diferentes variáveis para se chegar a uma imagem completa do comportamento do fármaco a partir das suas versões ativas. Os procedimentos para o teste de dissolução variam de acordo com o número de ingredientes ativos nas formulações. O método oficial é descrito pela United States Pharmacopeia (USP) para rifampicina, isoniazida e rifampicina cápsula, onde o meio de dissolução é de 0,1 mol L⁻¹ de HCl. O procedimento para três drogas com combinações de dose fixa (FDC, do inglês *fixed dose combinations*) tornou-se oficial na USP Suplemento 24 2, enquanto que para as quatro drogas (FDC) tornou-se oficial na USP 26 (2003) [8].

Existem alguns trabalhos [9-11] na literatura que descrevem a aplicação da técnica NIR em ensaios de dissolução em comprimidos. Freitas et al [9] realizou um estudo sobre o teste de dissolução dos comprimidos de clonazepam usando espectroscopia de reflectância difusa. O método quimiométrico dos mínimos quadrados parciais (PLS, do inglês *partial least squares*) foi usado, e comprovou ser um método de regressão adequado na construção dos modelos de calibração, sem recorrer a um pré-tratamento de dados. Todas as 21 curvas de taxas de dissolução previstas contra os valores medidos foram consistentes e mostraram correlação satisfatória, demonstrando o potencial deste método para auxiliar no desenvolvimento de formulações de ensaios adequados de dissolução. Tabasi et al [10] utilizou espectroscopia NIR e modelos de calibração multivariada para prever o comportamento de dissolução de teofilina. Eles analisaram 117 comprimidos separados em cinco lotes de diferentes proporções de Eudragit NE 30D. Para todas as leituras, o erro padrão de calibração foi menor do que o erro padrão da previsão. Portanto, os resultados foram considerados bem sucedidos e mostraram que a espectroscopia NIR, combinado com a modelagem multivariada é uma excelente alternativa.

Abe e Otsuka [11] investigaram a relação entre o lubrificante de mistura de tempo e tempo de dissolução por espectroscopia NIR e análise quimiométrica. O pó a granel que consistia de teofilina, lactose e amido foi pré-misturado. O material particulado foi compactado para formar comprimidos e testes de dissolução foram realizados. A partir de cada curva de dissolução, o tempo de dissolução de 50% (T50) foi calculado. Estes estudos mostram a utilidade de medições NIR para prever o parâmetro de dissolução.

Nosso grupo de pesquisa recentemente empregou a técnica de espectroscopia NIR e modelos PLS [12], para medir o percentual de dissolução dos fármacos: isoniazida, rifampicina, pirazinamida e etambutol em preparações farmacêuticas produzidas na UFRN para o tratamento da tuberculose pulmonar. A raiz quadrada da soma do quadrado dos erros da previsão (RMSEP), obtido a partir de modelos PLS foram: 9,99%, 8,63%, 8,57% e 9,97% para isoniazida, rifampicina, etambutol e pirazinamida, respectivamente.

1.2 MÉTODO DE REFERÊNCIA

O controle de qualidade dos medicamentos são análises de rotina que necessitam de métodos analíticos rápidos e que forneçam parâmetros apropriados para análises quantitativas de fármacos. Por este motivo faz-se necessário o uso de técnicas que forneçam resultados precisos, exatos, rápidos e de baixo custo.

Diversos estudos, tais como testes de dissolução, fazem uso da técnica de cromatografia líquida de alta eficiência (CLAE) para a quantificação de fármacos, tal como regulado pela USP. Atualmente, a utilização de métodos de CLAE em estudos de dissolução tem vindo a aumentar devido à dissolução *in vitro* do princípio ativo ser um parâmetro importante para as indústrias farmacêuticas em diferentes fases do desenvolvimento e produção de um medicamento.

A técnica CLAE submete as amostras juntamente com a fase móvel, que é eluída a grandes pressões, através de uma fase estacionária que são colunas preenchidas com materiais especialmente preparados para cada tipo de amostra. Na qual as forças físicas e químicas que atuam entre os solutos e as duas fases são responsáveis pela retenção dos solutos sobre a coluna cromatográfica. A diferença na magnitude dessas forças que determina a resolução e, portanto a separação dos solutos individuais.

É uma técnica que tem a capacidade de realizar separações e análises quantitativas de uma grande quantidade de compostos presentes em vários tipos de amostras, em escala de tempo de poucos minutos, com alta resolução, eficiência e sensibilidade.

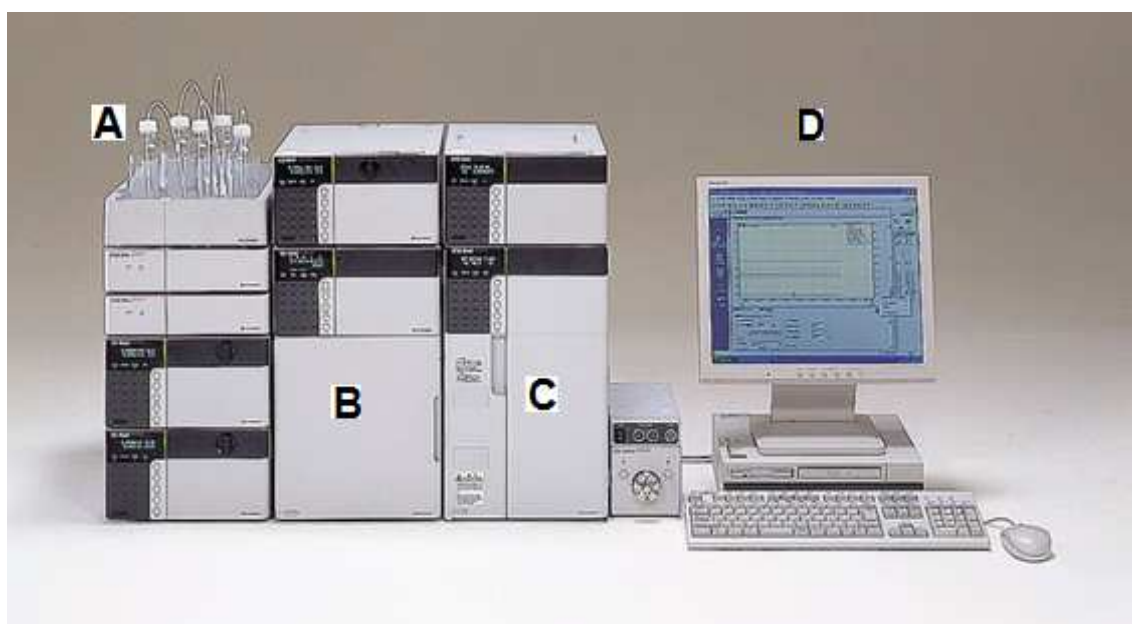


Figura 1: Visão esquemática do Cromatógrafo Líquido de Alta Eficiência: A) soluções, B) coluna cromatográfica, C) injetor, D) Computador para aquisição dos dados.

Panchagnula e colaboradores (1999) utilizaram a técnica CLAE de fase reversa para determinar rifampicina e seu principal metabolismo no plasma e na urina, em presença de pirazinamida e isoniazida. O trabalho visou desenvolver um método sensível e simples para determinar simultaneamente

rifampicina e desacetil rifampicina em fluídos biológicos, em presença de isoniazida e pirazinamida. Como resultado, observou-se que o método de fase reversa discutido é simples, sensível e com grande repetibilidade para análises de rifampicina e desacetil rifampicina em amostras de plasma e urina. O método é capaz de fazer estimações seletivas, além de necessitar de um volume bem menor de amostras e simplicidade nos procedimentos de extração, bem como uma fase móvel de simples preparo [13].

1.3 MÉTODOS QUIMIOMÉTRICOS

A quimiometria é a área da ciência que se refere à aplicação de métodos estatísticos e matemáticos, assim como aqueles baseados em lógica, à problemas de origem química. [2] De uma maneira geral, a quimiometria fornece meios de transformar dados, tais como espectros, pH, pressão, temperatura, em previsões, que podem ser concentração, classificação, entre outros.

A área da quimiometria que tem como objetivo encontrar um algoritmo capaz de produzir uma propriedade de interesse, com base em resultados fornecidos por um equipamento é a calibração multivariada. Uma vez determinado este algoritmo, chamado de etapa de calibração, é possível prever a propriedade de interesse em amostras não utilizadas na construção do modelo, desde que se forneçam os dados com os quais o algoritmo foi calculado.

A espectroscopia NIR é um método promissor no processo de controle de qualidade de formulações farmacêuticas, porém necessita de ferramentas de calibração multivariada para fazer a interpretação dos dados obtidos. O uso de calibração multivariada possibilita a redução de erros e obtenção de resultados satisfatórios.

1.3.1 Pré-tratamentos espectrais

Os espectros NIR apresentam um grande número de sobreposição das bandas de absorção dos diferentes constituintes das amostras. A interpretação desses espectros torna-se difícil graças ao espalhamento da luz, tamanho de partículas e multicolineridade dos dados, o que gera diferentes fontes de erro [14]. Os métodos de pré-tratamentos espectrais são desenvolvidos para redução dos efeitos das variações espectrais que não são relatadas nas variações químicas das amostras.

Em geral, os dados obtidos por reflectância difusa apresentam ruídos instrumentais e erros no espalhamento de luz, o que gera a necessidade de se utilizar no espectro alguns pré-tratamentos. Entre eles destacam-se a suavização Savitzky Golay, que diminui os ruídos instrumentais aleatórios e aumenta a razão sinal/ruído; a correção de espalhamento multiplicativo de luz (MSC), que corrige o espalhamento de luz, devido a não uniformidade na granulação e geometria das partículas dos comprimidos e por fim a derivada Savitzky Golay que faz a correção da linha base.

Azzouz et. al. (2003) compararam diferentes pré-tratamentos visando aperfeiçoar os modelos para obtenção dos valores de umidade e conteúdo de proteínas em amostras de alfafa. [14] Foram avaliados entre outros pré-tratamentos: primeira e segunda derivadas, MSC, padrão normal de variáveis e correção ortogonal de sinal (OSC, do inglês *orthogonal signal correction*). Para avaliar a eficiência de cada pré-tratamento foram calculadas as figuras de mérito para cada método de remoção de ruído utilizado. O melhor pré-tratamento observado foi o OSC, por remover da matriz dos espectros apenas um vetor que é unicamente correlacionado com a concentração do analito.

1.3.2 Seleção de variáveis

Os modelos de calibração multivariada são comumente calculados utilizando toda a faixa espectral fornecida ao modelo, entretanto, ao construir o modelo com todas as variáveis espectrais, tem-se o risco de modelar

informações que não são úteis ou que não correspondem ao analito de interesse. Os métodos de seleção de variáveis são úteis na tentativa de obter um modelo com a mesma eficiência do que utilizaria todo o espectro, porém utilizando menos variáveis, de forma a remover do espectro as variáveis que não contém informações significativas. Classicamente a escolha dessas variáveis é feita com base nas propriedades espectroscópicas das amostras. Vem sendo comprovado que a utilização de estratégias matemáticas, possibilita a seleção de variáveis de forma mais eficiente.

Existem vários algoritmos de seleção de variáveis, os mais utilizados são o algoritmo das sucessões progressivas (APS) que retira a colinearidade das variáveis. O princípio da seleção de variáveis por APS é que a nova variável selecionada é aquela que possui o valor de projeção máxima dentre todas as variáveis restantes no espaço sub-ortogonal da variável previamente selecionada [15]. Entretanto, apresenta algumas desvantagens como baixa razão sinal/ruído e ser em alguns casos, insuficiente para a calibração multivariada, o que resulta em perda no modelo de predição das amostras.

O iPLS (do inglês, *interval partial least squares regression*) é outro algoritmo de seleção de variáveis [15]. Neste método os dados são subdivididos em seções não sobrepostas, em cada seção separada são feitos modelos PLS para determinar a gama de variáveis mais úteis.

O Algoritmo Genético (AG, do inglês *genetic algorithm*) é uma técnica heurística popular de otimização que emprega uma probabilística e o processo de pesquisa não-local é inspirado na teoria de Darwin da seleção natural [15]. AG é bastante utilizado atualmente em muitos campos e foi aplicado com sucesso na solução de problemas de seleção de frequência, em que o AG manipula cadeias binárias chamadas cromossomos que contêm genes que codificam fatores experimentais ou variáveis. A otimização utilizando AG aliada à regressão PLS combina as vantagens dos dois algoritmos, por isso é o método mais comumente utilizado para analisar conjuntos de dados na espectroscopia NIR. AG aplicado a PLS têm se mostrado muito eficiente para procedimentos de otimização. Eles foram aplicados em muitos conjuntos de

dados espectrais e mostrou ser capaz de fornecer melhores resultados do que os que utilizaram o espectro total.

Xu et. al. (2012) utilizaram seleção de variáveis em espectros de IV e visível na determinação de teor de açúcar em peras. Foram testados os algoritmos APS, iPLS e AG. Os resultados mostram que a precisão da análise quantitativa realizada pela espectroscopia Vis-NIR pode ser melhorada através da adequada seleção de variáveis. O algoritmo AG, por exemplo, foi capaz de atingir resultados tão satisfatórios quanto os feitos com todo o espectro, utilizando apenas 2% do espectro [16].

1.3.3 Detecção de outliers

Durante o processo de calibração é possível observar a presença de amostras, chamadas de outliers, que são aquelas que apresentam valores atípicos, e em geral são prejudiciais ao modelo. Podemos reconhecer dois tipos de valores amostrais extremos. O primeiro tipo de outliers é uma amostra que representa uma composição extrema em relação ao restante do conjunto de calibração. Estas amostras têm um elevado efeito de influência no modelo de calibração. Podem ser determinadas usando a equação abaixo:

$$h_i = \frac{1}{n} + \sum_{k=1}^a \frac{t_{ik}^2}{\sum_{i=1}^n t_{ik}^2} \quad \text{Equação 01}$$

Onde:

h_i = influência da amostra i

n = número de amostras

a = número de componentes principais consideradas

t_{ik} = o escore de cada amostra i e componente k

Um valor baixo do efeito de influência h_i , indica que as amostras consideradas são semelhantes à média das amostras de calibração no espaço de poucas dimensões. As amostras que possuem um efeito de influência muito alto sobre os resultados da regressão têm grandes valores de h_i . A inclusão dessas amostras na calibração poderia dar estimativas errôneas de amostras similares, se o valor de referência para as amostras de alta influência estiver errado. Esta estatística é usada para a detecção de outliers, durante calibrações, para a detecção de extrapolações do modelo durante a análise de predição, e para estimar a incerteza sobre um valor de estimativa [14].

O segundo tipo de outliers é aquele em que o valor do parâmetro estimado difere estatisticamente de forma significativa do valor de referência. Isto pode ser um indicativo de que pode ter havido um erro na obtenção do valor de referência, o que implica em erros no modelo. Este outro tipo de outlier pode ser detectado através da análise dos resíduos, onde os valores previstos pelo modelo são confrontados com os valores de referência.

1.4 FIGURAS DE MÉRITO MULTIVARIADAS

A qualidade nas medidas experimentais e confiabilidade estatística dos cálculos envolvidos nos processos são responsáveis pelo bom desempenho de qualquer técnica analítica. Para certificar a boa qualidade e aplicabilidade dos métodos é necessária que a nova técnica que se deseja implantar seja validada. A validação é uma averiguação do desempenho de um método, com a finalidade de avaliar se este apresenta um desempenho adequado para as condições nas quais será aplicado. [17] Essa validação pode ser feita através do cálculo de alguns parâmetros conhecidos como figuras de mérito, dentre

elas: seletividade, sensibilidade, linearidade, limite de detecção (LD), limite de quantificação (LQ), precisão, exatidão e relação sinal-ruído.

A maneira como essas figuras de mérito devem ser determinadas é estabelecida pelos órgãos de fiscalização e encontra-se descrita em normas específicas, guias de validação e trabalhos científicos. Entretanto, a maioria dos guias, normas e trabalhos científicos, referem-se apenas à calibração univariada e são poucos os trabalhos científicos que realizam a determinação de figuras de mérito para validação de modelos de calibração multivariada.

As análises quantitativas utilizando espectroscopia NIR têm crescido a cada ano, e estas determinações são feitas com base em modelos de calibração multivariada. A ausência de validação desses modelos, devido à carência de normas oficiais que descrevem como essa validação deva ser realizada tem restringido sua implementação. [17]

A etapa de validação é demorada, necessita de um número elevado de experimentos e cálculos estatísticos, o que leva a um alto custo nas análises. Tendo em vista esse impasse, tem-se intensificado a busca por procedimentos de validação com menor custo, em tempo mais curto e com a mesma confiabilidade.

Sarraguça e Lopes (2009) utilizaram o sinal líquido do analito (NAS, do inglês *net analyte signal*) na espectroscopia NIR em aplicações farmacêuticas, como uma estratégia de pré-processamento para retirar informações dos analitos a serem calibrados, devido a existência de muita sobreposição de dados nos espectros. Foram calculadas as figuras de mérito: sensibilidade e seletividade com base no vetor NAS. [18]

1.4.1 Linearidade

Expressa a faixa na qual o sinal analítico (variável dependente X) é linearmente proporcional à sua concentração (variável independente Y), bem como a determinação da equação matemática que rege a dependência, denominada de curva analítica.

O ajuste é feito pelo método dos mínimos quadrados, a melhor curva é aquela que apresentar menor resíduo e maior coeficiente angular, entre os valores de Y medidos e os valores de Y preditos. Uma vez conhecidos os coeficientes de regressão, é possível prever a concentração de novas amostras, com base apenas em seu sinal analítico. O valor R^2 pode ser calculado através da seguinte equação:

$$R^2 = \frac{\sum (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{(\sum (x_i - \bar{X})^2) \sum (y_i - \bar{Y})^2}} \quad \text{Equação 02}$$

Onde:

R^2 = coeficiente de correlação

X_i = variável independente

\bar{X} = valor médio de X

Y_i = parâmetro medido

\bar{Y} = valor médio do parâmetro medido

Entretanto apenas a análise do coeficiente de correlação não é suficiente para garantir a linearidade do modelo, que pode possuir um resíduo alto, má distribuição dos pontos e mesmo assim possuir uma correlação bem próxima a 1. Uma estratégia para minimizar os efeitos de falta de ajuste do modelo é a análise do gráfico dos resíduos do sinal analítico. É importante que se investigue neste gráfico possíveis correlações entre as amostras.

1.4.2 Precisão

Refere-se à dispersão dos valores medidos em torno de um valor médio, calculado através do desvio padrão relativo (DPR). Pode ser estimada em três

níveis: repetibilidade, precisão intermediária e reprodutibilidade [17]. Em que a repetibilidade, diz respeito à precisão nas medidas feitas em um mesmo laboratório, em um único equipamento e pelo mesmo analista; a precisão intermediária é a precisão dentro de um mesmo laboratório para medidas feitas por diferentes analistas ou equipamentos e a reprodutibilidade é a precisão dos resultados obtidos por diferentes laboratórios para uma mesma análise, que tenham utilizado a mesma metodologia. Poderá ser calculada através da seguinte equação:

$$Precisão = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (\hat{y}_{ij} - \bar{\hat{y}}_i)^2}{n(m-1)}} \quad \text{Equação 03}$$

Onde:

n = número de amostras

m = número de replicatas

\hat{y} = valor do parâmetro previsto

$\bar{\hat{y}}$ = média do valor do parâmetro previsto

1.4.3 Exatidão

Mede o quão próximo se encontra o valor medido do valor tido como referência, tem relação com o erro absoluto. Pode ser determinada através deste estudo frente a uma substância conhecida, padrões certificados, por exemplo. Ou então, pela diferença entre o valor previsto e o valor verdadeiro.

Em calibração multivariada é expressa através da raiz quadrada do erro médio quadrático de previsão (RMSEP, do inglês, *Root Mean Squares Error of Prediction*), que é calculado segundo a Equação 04.

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

Equação 04

Onde:

\hat{y} = valor predito

y_i = valor de referência

l_v = número de amostras do conjunto de validação

1.4.4 Sensibilidade

É a razão entre a inclinação da curva analítica e os desvio padrão do sinal analítico. Pode ser calculado através da equação abaixo:

$$SEN = \frac{1}{\|b_k\|}$$

Equação 05

Onde:

b_k = coeficientes de regressão do modelo PLS

1.4.5 Seletividade

Mede o grau de sobreposição entre o sinal da espécie de interesse e os interferentes presentes na amostra, indicando também, a parte do sinal que é perdida por essa sobreposição. Para modelos de calibração multivariada, a seletividade pode ser calculada através da seguinte equação:

$$SEL_{ki} = \frac{n\hat{\alpha}s_{k,i}}{\|x_{k,i}\|}$$

Equação 06

Onde:

$n\hat{\alpha}s_{k,i}$ = valor escalar do sinal analítico líquido para a amostra i

$x_{k,i}$ = vetor de respostas instrumental para a amostra i.

1.4.6 Sensibilidade do método: Limite de detecção (LD) e Limite de quantificação (LQ)

Define-se como sensibilidade de um método os limites de detecção (LD) e quantificação (LQ), em que o LD é a menor concentração do analito que a técnica permite detectar e que pode ser distinguível do sinal analítico e retira o ruído, enquanto que o LQ é a mais baixa concentração que pode ser quantificada dentro dos limites de precisão e exatidão. Podem ser estimados a partir da comparação do sinal analítico obtido para uma amostra que possui baixas concentrações do analito com o sinal de uma amostra do branco, estimando assim, a concentração mínima com a qual a espécie de interesse pode ser determinada de maneira confiável. Os valores de LD e LQ podem ser calculados através das equações:

$$LD = 3\sigma\delta x \|b_k\|$$

Equação 07

$$LQ = 10\sigma\delta x \|b_k\|$$

Equação 08

Onde:

LD = limite de detecção

LQ= limite de quantificação

δ = desvio padrão

b_k = coeficientes de regressão do modelo PLS

Os valores de LD e LQ podem ser estimados através do método visual, relação sinal/ruído e baseado em parâmetros da curva analítica. Os dois primeiros métodos apesar da rapidez com que geram os resultados apresentam a desvantagem de serem baseados em parâmetros qualitativos.

2. OBJETIVOS

O objetivo deste trabalho foi demonstrar como figuras de mérito multivariadas, juntamente com detecção de outlier, métodos de regressão por mínimos quadrados parciais (PLS) e algoritmos de seleção de variáveis podem ser utilizados para avaliar e melhorar o desempenho de espectroscopia no infravermelho próximo na determinação simultânea da dissolução de comprimidos contendo isoniazida, rifampicina, pirazinamida e etambutol. Comparar através de um teste t-pareado os resultados obtidos pelo método proposto com os da técnica de referência.

3. EXPERIMENTAL

3.1 MATERIAIS

Para a determinação dos quatro princípios ativos farmacêuticos através do método de dissolução, os seguintes reagentes e soluções foram utilizados: fosfato de sódio dibásico anidro (99,0%, Labsynth), ácido fosfórico (85,0%, Reagente Analítico), acetonitrila (0,7791 gm L⁻¹, JT Baker Solusorb),

trietilamina (Vetec), metanol (JT Baker Solusorb), isoniazida (98,21%), rifampicina (93,09%), pirazinamida (99,10%), e etambutol (99,87%). A formulação farmacêutica utilizada para estudar a dissolução consiste de isoniazida (7,92%), rifampicina (15,81%), pirazinamida (42,20%) e etambutol (29,03%) como ingredientes ativos e Aerosil (0,14%), explocel (2,80%), talco (1,4%) e estearato de magnésio (0,7%) como excipientes em forma de comprimido.

Foram preparadas três fases móveis (A, B e C) para a análise dos produtos farmacêuticos, utilizando CLAE. As fases móveis A e B consistiam em tampão de fosfato e acetonitrila em um 96:4 e uma razão de 55:45, respectivamente. Para fase móvel C, tampão de trietilamina foi primeiramente misturado com água a uma razão de 1:1000. O pH foi ajustado para $7,0 \pm 0,1$ usando ácido fosfórico. A solução foi filtrada a vácuo e desgaseificada durante 10 minutos. A solução tampão foi então misturada com acetonitrila numa razão de 50:50. Todas as fases móveis foram filtradas à vácuo e desgaseificada durante 10 minutos antes da utilização.

3.2 NIR

Os espectros NIR de reflectância foram medidos utilizando um modelo de espectrofotômetro Bomem MB 160 acoplado a um acessório de reflectância difusa. Cada espectro medido (em triplicata) foi a média de 50 scans obtido com uma resolução de 8 cm^{-1} e ao longo do intervalo de 1100-2500 nm. O espectro de uma amostra de politetrafluoretileno (PTFE) foi usado como a referência.

3.3 CLAE E APARELHO DE DISSOLUÇÃO

O dissolutor automático Erweka Modelo BT-80 foi utilizado para realizar a dissolução dos comprimidos da amostra contendo os princípios ativos. O equipamento de CLAE (Shimadzu), modo isocrático, utilizado para obter os cromatogramas consiste nos seguintes módulos: desgaseificador (DGU-20A5),

da bomba (LC-20AT), auto-injetor (SIL-20A), forno de coluna (CTO-20A), matriz de iodo detector (SPD-M20A) e comunicação (CBM-20A). Para pesagem das amostras e padrões, uma balança de precisão analítica alta (Aventura Ohaus) foi usada. Um modelo FANEM 257 de agitador foi utilizado para a homogeneização das soluções preparadas. Um medidor de pH Tecnal Tec-3MP foi utilizado para controlar o pH das soluções tampão. Um modelo de bomba à vácuo 504 (Eletrolab) foi utilizado para auxiliar a filtração das soluções preparadas. O dissolutor automático foi definido com as seguintes especificações: tempo de 45 minutos, temperatura de 37 °C, o aparato utilizado foram as pás, 900 mL de tampão fosfato com pH 6,8 como meio de dissolução e uma taxa de rotação de 100 rpm. Os comprimidos foram dissolvidos automaticamente sob as condições relatadas. Alíquotas de cada amostra dissolvida foram recolhidas e filtradas utilizando papel de filtro e transferidas para um frasco de 1,5 mL para cromatografia. Os valores calculados para a dissolução dos comprimidos analisados variaram de 5% a 90%, através dos dados de CLAE.

3.4 SOFTWARE E ANÁLISE DE DADOS

O carregamento de dados, pré-processamento (correção de linha de base, centragem da média, alisamento e derivação Savitzky-Golay), método de seleção de variáveis (iPLS, AG e APS), construção do modelo de regressão quimiométrica (full-PLS, PLS-iPLS, PLS-AG , PLS-APS, MLR-APS), e de validação foram realizados em MATLAB ® versão 7.10 (Math-Works, Natick, EUA) com o PLS-toolbox (Investigação Eigenvector, Inc., Wenatchee, WA, EUA, versão 4.0). As amostras foram divididas em conjuntos de calibração e validação através da aplicação do clássico Kennard-Stone, algoritmo de seleção (KS) para os espectros NIR [20]. Foram selecionadas 28 amostras para o conjunto de calibração e 10 para o conjunto de previsão. A detecção de outlier foi realizada para melhorar a precisão do modelo e para remover as amostras com valores extremos para a sua influência e de resíduos não-modelados no Y e X nas respostas de dados. Uma região elíptica de confiança conjunta (EJCR) foi calculada para avaliar a inclinação e interceptar

a regressão da referência e prever os valores a um intervalo de confiança de 95%. Aspectos da matemática e aplicações da presente EJCRC podem ser encontrados nas seguintes referências [22-23]. Todos estes cálculos foram realizados utilizando rotinas adequadas MATLAB 7.10.

4. RESULTADOS E DISCUSSÃO

Inicialmente foram obtidos os espectros em triplicata das 38 amostras de comprimidos da produção, registrados na região de 1100 a 2500 nm. Na figura abaixo são mostrados as médias das medições em triplicata para cada comprimido:

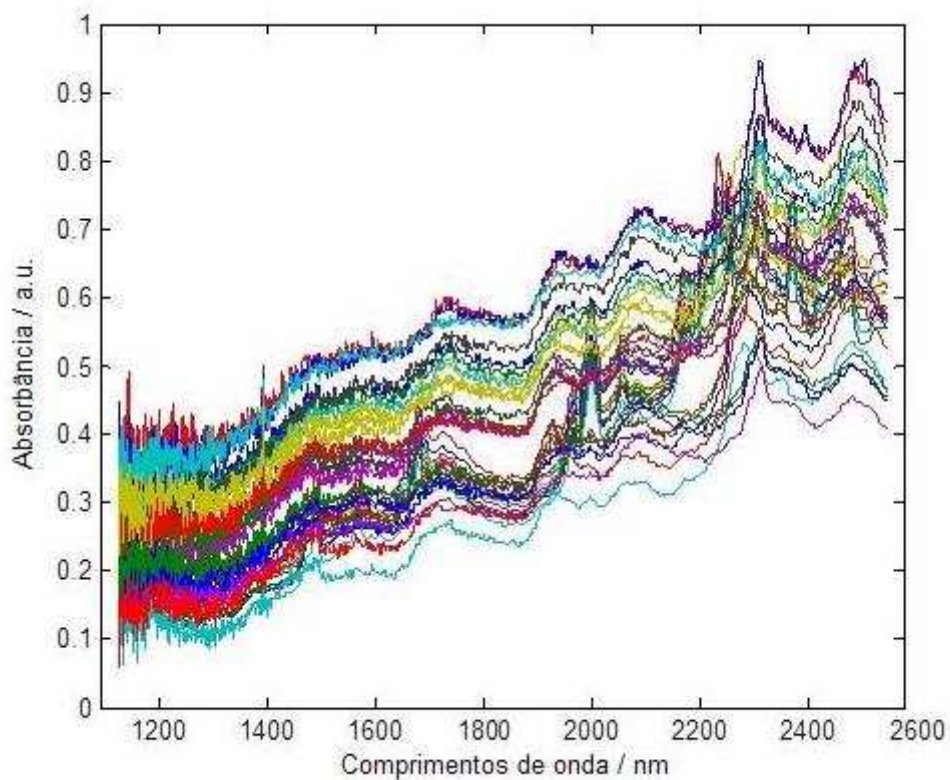


Figura 2: Espectros NIR das 38 amostras originais.

Conforme é possível identificar na Figura 2 os espectros possuem muita sobreposição, de maneira que se torna impossível identificar os constituintes

das amostras, os 4 princípios ativos e excipientes, distinguir características semelhantes entre eles, muito menos atribuir alguma propriedade química.

No máximo pode-se atribuir alguns sobretoms e bandas de combinação evidenciadas no espectro, destacando-se: região A 1450 - 1460 nm é atribuída a NH (amidas) para o primeiro sobretom; região B: 1600-1800 nm é atribuída a CH para o primeiro sobretom; a região C: 1910 - 1980 nm é atribuída a C=O (amida) para o segundo sobretom, região D: 2110 a 2205 nm e 2250 nm a 2260 regiões são atribuídos a NH (aminas) e OH (água), e é relacionada com uma combinação de vibrações. Como mostrado na Figura 3:

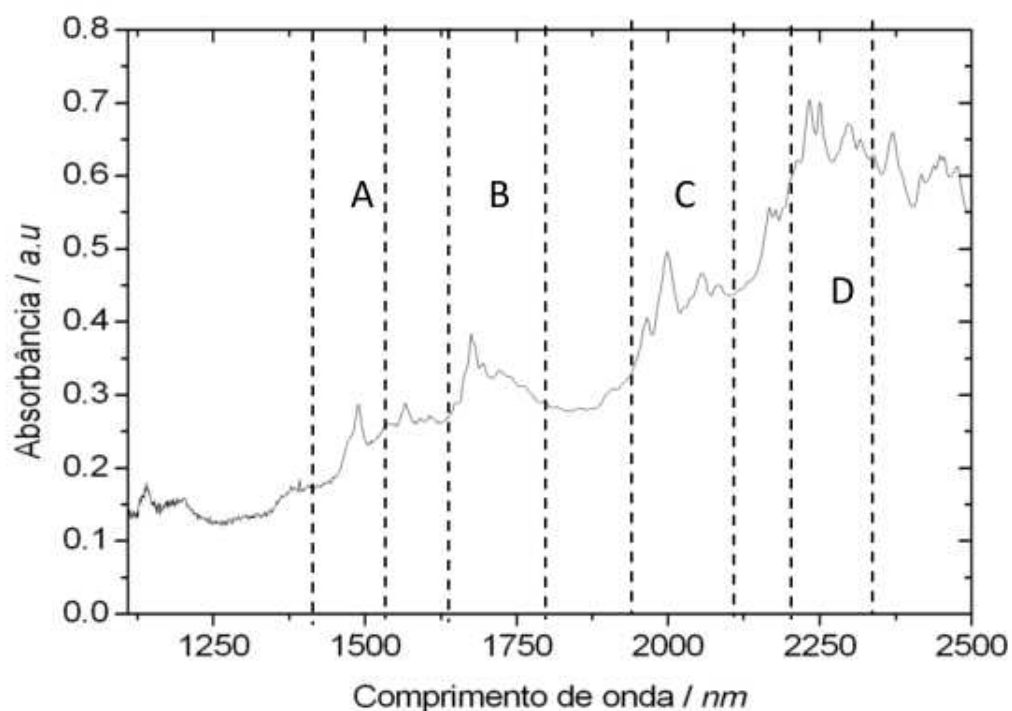


Figura 3: Atribuição de bandas no espectro original de uma amostra.

4.1. MODELO PLS

Após a aquisição dos dados espectrais e já de posse dos valores de dissolução para cada um dos 4 princípios ativos, medidos através do método de referência, o próximo passo foi a construção dos modelos multivariados.

Para tanto, fez-se necessários alguns pré-tratamentos espectrais afim de facilitar a interpretação das propriedades relacionadas aos sinais analíticos.

O primeiro pré-processamento realizado foi o alisamento Savitzky-Golay, para remover dos espectros a presença de ruídos instrumentais aleatórios, aumentando, assim, a razão sinal/ruído. As janelas variam de 5 a 11 pontos com polinômio de primeiro grau. Através da aplicação do MSC foi realizada a correção do espalhamento de luz, provocada pela diferença na geometria e tamanho das partículas na amostra, que causam a dispersão da luz em todas as direções, o que gera alterações na relação entre a intensidade das medições de reflectância e a concentração das espécies analisadas. Por fim, foi feito a derivada Savitzky-Golay, para a correção da linha base, na qual a janela variou 7 a 11 pontos com polinômio de primeira ordem.

As amostras foram divididas em dois conjuntos: 28 na calibração e 10 na previsão. A primeira categoria é utilizada na construção dos modelos multivariados e a segunda na estimação do desempenho do modelo, para prever a dissolução de amostras futuras. Esta seleção foi feita com o algoritmo Kennard-Stone, que faz o cálculo da distância euclidiana de todas as amostras, com o centro da distribuição, e com base nisso sugere para calibração as amostras mais distantes, aquelas cujos valores são os extremos, máximos e mínimos, e as de previsão os valores intermediários. Não foram usadas amostras de validação, por que o modelo faz uma validação cruzada, na qual ele retira uma amostra da calibração faz o modelo com as restantes e prevê a que não foi incluída, de forma a fazer já a validação interna.

Os melhores modelos de PLS obtidos durante a fase de pré-tratamento utilizando alisamento Savitzky-Golay (com uma janela de 7 e 9 pontos), MSC e a derivada Savitzky-Golay de primeira ordem (com uma janela de 7 e 9 pontos). Com janelas superiores a este valor, a correlação caía bastante e o erro aumentava consideravelmente. Foram testados modelos com segunda derivada, mas os resultados não foram satisfatórios, por isso não foram incluídos neste trabalho, conforme verificamos na Tabela 1.

O melhor modelo foi determinado utilizando a raiz quadrada da soma do quadrado dos erros da calibração (RMSEC) e raiz quadrada da soma do

quadrado dos erros da previsão (RMSEP) versus o número de variáveis latentes (VL) do PLS. O número ideal de VL foi selecionado com base no ponto de inflexão de incrementos RMSE (ou variância). Bem como levando em consideração o acordo entre RMSEC e valores RMSEP.

Tabela 1 – Resultado dos modelos PLS. O número dentro do parêntese corresponde ao número de variáveis latentes.

Modelos	Calibração		Previsão		
	R^2	RMSEC (%)	R^2	RMSEP (%)	Tamanho
ISONIAZIDA PLS1 (7) 9 pts.	0,99	3,92	0,91	12,24	1284
RIFAMPICINA PLS1 (7) 9 pts.	0,99	5,34	0,96	11,40	1284
PIRAZINAMIDA PLS1 (2) 9 pts.	0,81	14,22	0,64	14,90	1284
ETAMBUTOL PLS2 (4) 7 pts.	0,83	10,25	0,86	18,28	1284

Para os modelos PLS, foram utilizados de 2 a 7 VL para os 4 API. Os coeficientes de correlação obtidos foram: 0,91, 0,96, 0,64 e 0,86, e os valores de RMSEP: 12,19%, 11,40%, 14,90% e 18,28% para isoniazida, rifampicina, pirazinamida e etambutol, respectivamente. A faixa espectral utilizada contemplava 1284 comprimentos de onda.

Usando o conjunto de previsão (10 amostras), uma boa correlação e baixos RMSEP foram encontrados entre o percentual de dissolução previsto por NIR e o percentual de dissolução determinada pelo equipamento de laboratório (aparelhos de dissolução e determinação por CLAE). Estes resultados demonstram uma boa qualidade de previsão para o teste de dissolução de quatro comprimidos API para modelos de PLS. Entretanto, os modelos atuais, utilizando apenas os pré-tratamentos não conseguiram superar os modelos do artigo de Neves et al [12], por isso, algumas novas estratégias como a seleção de variáveis e detecção de outliers foram utilizados.

4.2 SELEÇÃO DE VARIÁVEIS

Nesta seção serão discutidos os três algoritmos de seleção de variáveis, aplicados aos modelos PLS, na tentativa de otimizar os resultados alcançados, por cada um deles.

4.2.1 Isoniazida

Os resultados obtidos para os modelos utilizando PLS-APS, PLS-AG e iPLS de calibração na região do NIR para isoniazida em comprimidos são mostrados na Tabela 2. Apenas os melhores resultados para os pré-processamentos testados são apresentados. Observamos que a seleção de variáveis foi importante neste trabalho porque permitiu o uso de modelos que exigem um pequeno número de variáveis espectrais. Os coeficientes de correlação para o conjunto de previsão variaram de 0,86 a 0,94 para todos os modelos, apenas com exceção do modelo com iPLS, para apenas 64 variáveis.

Tabela 2: Resultados para calibração e validação externa definida para isoniazida.

Modelos ^a	Calibração		Previsão		
	R^2	RMSEC (%)	R^2	RMSEP (%)	Tamanho
iPLS (2)	0,85	14,68	0,31	28,36	64
iPLS (3)	0,93	9,78	0,86	15,61	128
iPLS (2)	0,74	18,60	0,94	10,24	256
PLS-APS (4)	0,95	8,48	0,93	14,04	27
PLS-AG (6)	0,99	4,49	0,91	13,49	352
iPLS (5) ¹	0,99	2,94	0,94	6,28	256

^a Pts, pontos; deriv., derivada; ¹ Aplicação de detecção de outliers.

O número de VL utilizados para os modelos de iPLS, APS e AG utilizando espectros NIR foi de 2 a 6. Para isoniazida, pode-se ver (na Tabela 2) que a maioria dos modelos após a seleção de variável por AG obtiveram

valores baixos de RMSEP comparado com o modelo original, o que implica e, bom poder preditivo. As melhorias no RMSEP ficaram entre 6,28% e 12,24%. Os modelos com AG e APS tiveram valores baixos de RMSEP. Os que utilizaram iPLS exibiram melhorias nos valores RMSEP, em comparação com o modelo original. Todas as técnicas de seleção de variáveis contribuíram para a capacidade preditiva do modelo, mas duas das técnicas mostraram melhoras maiores do que as outras, o AG e o iPLS.

Algumas amostras de valores extremos foram excluídas do conjunto de calibração no melhor modelo para isoniazida utilizando iPLS, cujo pré-tratamento foi alisamento e derivada, com janela de 9 pontos e MSC. Para este modelo, o menor RMSEC e RMSEP foram 2,94 e 6,28, respectivamente. O coeficiente de correlação para o conjunto de validação foi de 0,94 e foi obtido utilizando-se cinco VL.

A Figura 4 mostra a regressão entre os valores medidos e os previstos para o teste de dissolução usando o iPLS para isoniazida.

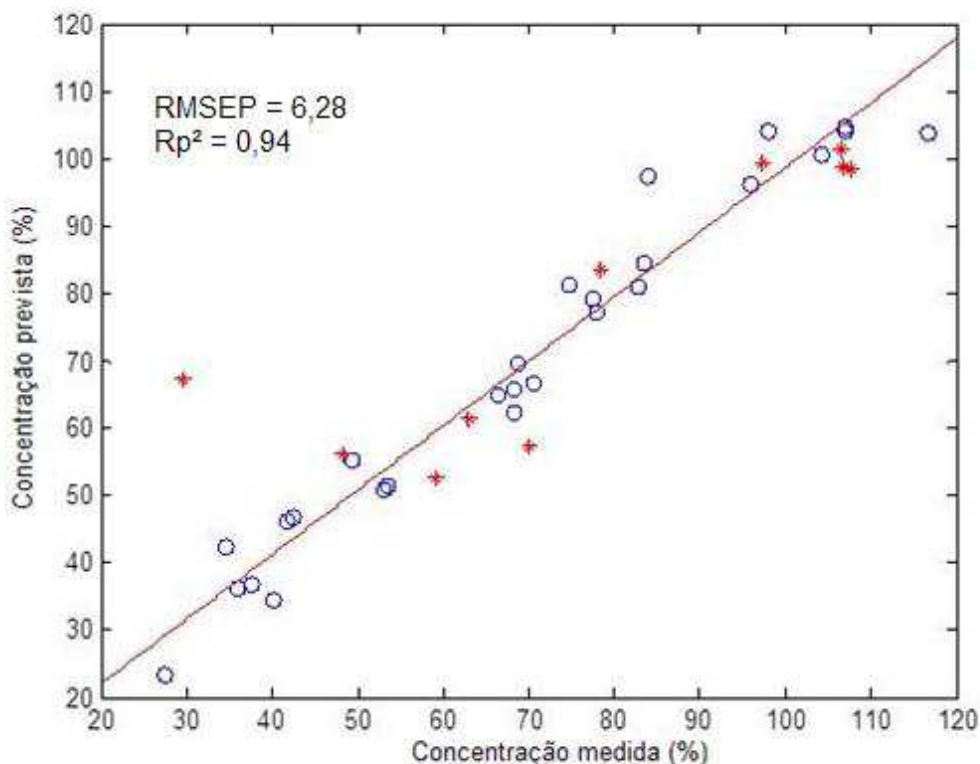


Figura 4: Concentração medida x concentração prevista pela técnica de referência das amostras de calibração e validação para isoniazida utilizando o modelo iPLS. (○) conjunto de calibração; (*) conjunto de validação.

Foi feito um teste t-pareado e comprovou-se que a um nível de 95% de confiança não há diferença estatística entre os resultados previstos pelo NIR e os valores medidos pela técnica de referência.

4.2.2 Rifampicina

A Tabela 3 mostra os resultados para a análise de rifampicina em comprimidos. Para este parâmetro foi feita a seleção de variáveis usando os iPLS, APS e algoritmo AG o que gerou bons resultados. Por exemplo, de acordo com a Tabela 3, para o modelo com iPLS (2), um coeficiente de correlação de 0,96 foi obtido para a predição definida usando apenas uma faixa com 64 variáveis espectrais, e um RMSEP foi 11,33%. Podemos observar que com o acréscimo de mais variáveis, a capacidade preditiva desse algoritmo aumenta.

Tabela 3: Resultados para calibração e validação externa definida para rifampicina.

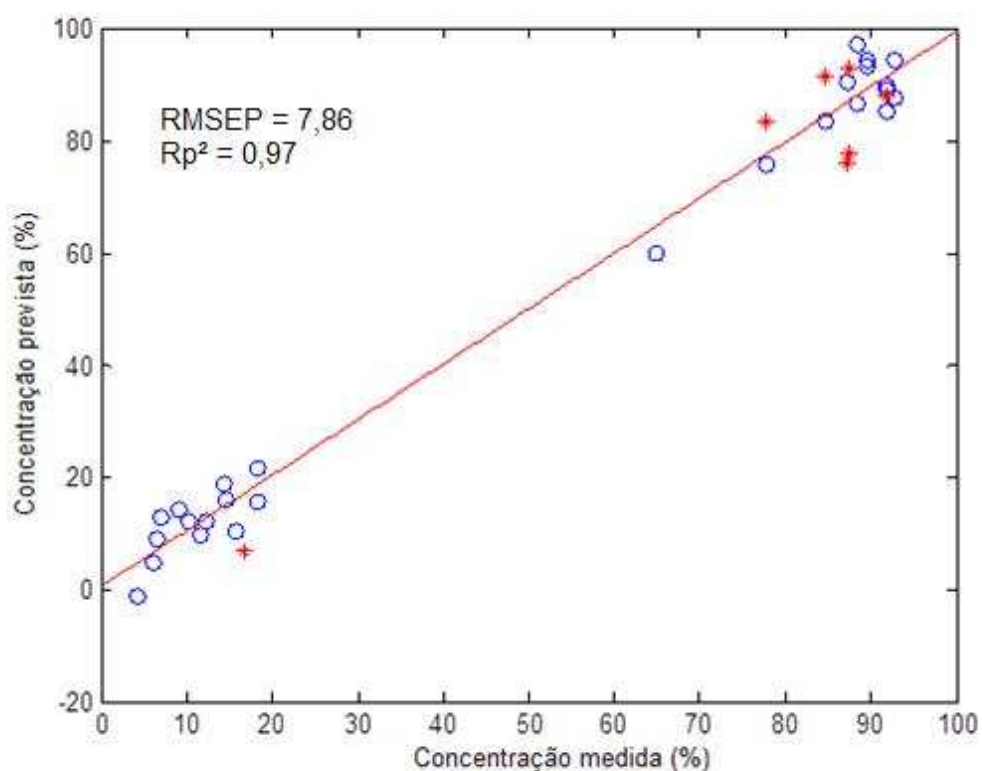
Modelos ^a	Calibração		Previsão		Tamanho
	R^2	RMSEC (%)	R^2	RMSEP (%)	
iPLS (2)	0,94	10,53	0,96	11,33	64
iPLS (3)	0,98	6,45	0,97	10,12	128
iPLS (4)	0,99	5,95	0,98	9,79	256
PLS-APS (2)	0,95	10,19	0,97	10,34	27
PLS-AG (5)	0,99	2,82	0,93	13,81	341
PLS-AG (4) ¹	0,99	4,03	0,97	7,86	341

^a Pts, pontos; deriv., derivativa; ¹ Aplicação de detecção de outliers

Todas as técnicas de seleção de variáveis melhoram a capacidade preditiva do modelo, mas dentre elas destacaram-se o iPLS e AG.

O melhor algoritmo de seleção de variáveis para este parâmetro foi o algoritmo PLS-AG. Quando 341 variáveis espectrais foram usados para construir o modelo PLS-GA (5), um coeficiente de correlação para a previsão definida de 0,93 foi encontrado, este valor foi próximo ao obtido para o melhor modelo. Sobre este modelo foi feito a detecção de outliers

Duas amostras de outlier foram excluídas do conjunto de calibração do melhor modelo AG para rifampicina. Para este modelo, o menor RMSEC e RMSEP foram 4,03 e 7,86, respectivamente. O coeficiente de correlação para o conjunto de validação foi de 0,97 e foi obtido utilizando-se quatro variáveis latentes. A Figura 5 mostra a regressão entre os valores medidos e os previstos para os testes de dissolução utilizando o modelo AG para rifampicina.



Para analisar a eficiência dos modelos foi feito um teste t-pareado e a um nível de 95% de confiança não há diferença estatística entre os resultados previstos pelo NIR e os valores medidos pela técnica de referência.

4.2.3 Pirazinamida

A Tabela 4 mostra os resultados dos modelos para pirazinamida. Em geral, os melhores valores de RMSEP foram obtidos para os dados derivados com janelas de 7 pontos. Os coeficientes de correlação para a conjunto de predição variaram de 0,38 a 0,98, indicando que alguns modelos foram insatisfatórios.

O algoritmo APS não mostrou ser muito eficiente e dos três, foi o que teve o pior desempenho. Os modelos com iPLS, obtiveram correlações na previsão próximas ao modelos com AG, entretanto apresentavam RMSEC maiores e correlação na calibração bem menores. Utilizando o algoritmo AG para escolher as principais variáveis espectrais para o uso nos modelos, o coeficiente de correlação obtido para o conjunto de previsão foi de 0,70, e o RMSEP foi 14,64 (341 variáveis espectrais).

Tabela 4: Resultados para calibração e validação externa definida para pirazinamida.

Modelos ^a	Calibração		Previsão		
	R^2	RMSEC (%)	R^2	RMSEP (%)	Tamanho
iPLS (3)	0,78	15,28	0,72	13,81	64
iPLS (2)	0,85	12,75	0,28	18,96	128
iPLS (3)	0,75	16,21	0,78	12,45	256
PLS-APS (2)	0,78	15,26	0,38	17,87	27
PLS-AG (6)	0,96	5,72	0,70	14,64	341
PLS-AG (3) ¹	0,99	2,98	0,98	9,37	341

^a Pts, pontos; deriv., derivada; ¹ Aplicação de detecção de outliers

Apenas as seleções de variáveis feitas por iPLS e AG melhoram a capacidade preditiva do modelo. Para o melhor modelo utilizando o AG, foi feita a detecção de outliers e três amostras foram excluídas do conjunto de calibração. Para este modelo, o menor RMSEC e RMSEP foram 2,98 e 9,37, respectivamente. O coeficiente de correlação para o conjunto de validação foi de 0,98 e foi obtido utilizando-se três variáveis latentes. A Figura 6 mostra a regressão entre os valores medidos e os previstos para o teste de dissolução utilizando o modelo AG para pirazinamida.

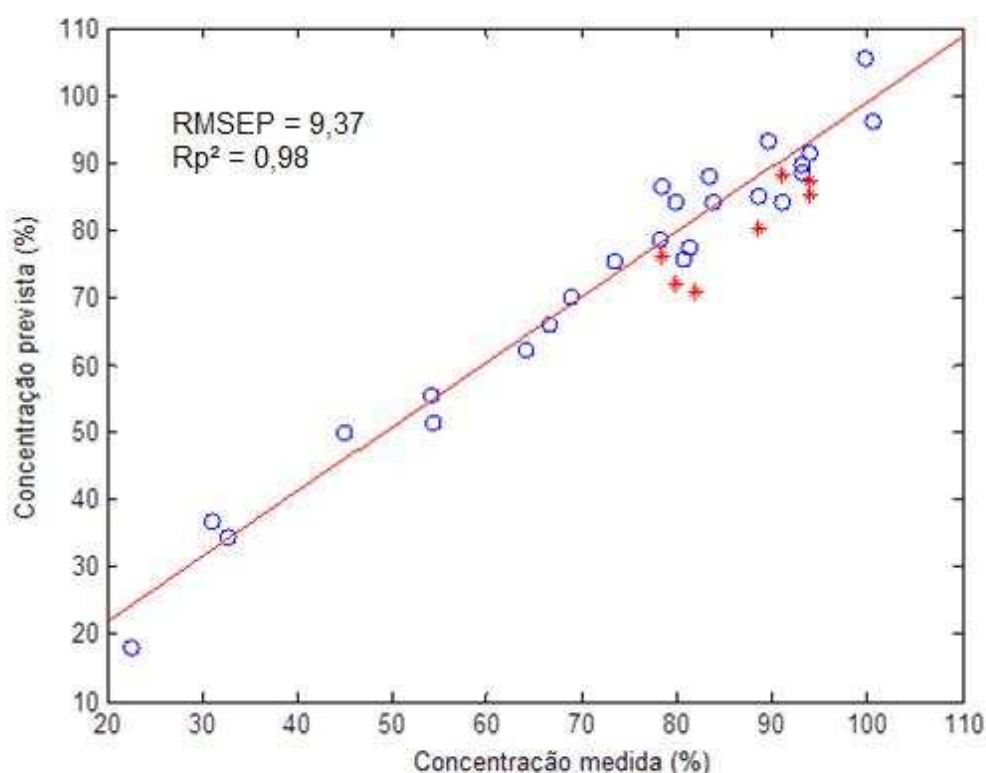


Figura 6: Concentração medida x concentração prevista pela técnica de referência das amostras de calibração e validação para pirazinamida utilizando o modelo GA-PLS. (○) conjunto de calibração; (*) conjunto de validação.

Semelhante aos modelos anteriores, para certificar a confiabilidade foi feito um teste t-pareado e a um nível de 95% de confiança, não houve diferença estatística entre as duas análises.

3.2.4. Etambutol

Os resultados para a análise de etambutol em comprimidos são mostrados na Tabela 5. Para este parâmetro foi realizada uma seleção de variáveis usando: iPLS, APS e algoritmos AG produzindo também bons resultados. Por exemplo, de acordo com a Tabela 5, para o modelo PLS-AG (4), um coeficiente de correlação de 0,82 foi obtido para a previsão definida usando apenas uma faixa com 307 variáveis espectrais, e o RMSEP obtido foi 16,47%.

Tabela 5: Resultados para calibração e validação externa definida para etambutol.

Modelos ^a	Calibração		Previsão		Tamanho
	R^2	RMSEC (%)	R^2	RMSEP (%)	
iPLS (2)	0,86	9,47	0,48	22,69	64
iPLS (2)	0,91	7,68	0,79	19,12	128
iPLS (6)	0,99	1,66	0,49	21,87	256
PLS-APS (2)	0,76	12,05	0,94	16,02	27
PLS-AG (4)	0,99	1,71	0,82	16,47	307
PLS (5) ¹	0,99	1,10	0,97	2,64	1284

^a Pts, pontos; deriv., derivada.; ¹ Aplicação de detecção de outliers

Como observado acima, o iPLS não forneceu bons resultados, as correlações na previsão foram baixas e RMSEP bastante elevado. O melhor algoritmo de seleção de variáveis para este parâmetro encontrado foi o APS. Quando 27 variáveis espectrais foram usadas para construir o modelo APS (2), um coeficiente de correlação para a previsão definida de 0,94 foi encontrado, este valor foi próximo ao obtido para o melhor modelo.

Apenas APS e AG foram capazes de melhorar a capacidade preditiva do modelo. Entretanto, ao fazer a detecção de outliers para os melhores modelos, o melhor resultado obtido, foi quando a seleção foi feita para o modelo com a

faixa espectral completa, deste modelo foram excluídas três amostras anômalas do conjunto de calibração. O RMSEC e RMSEP obtidos foram 1,10 e 2,64, respectivamente. O coeficiente de correlação para o conjunto de validação foi de 0,97 e foi obtido utilizando-se cinco VL. A Figura 7 mostra a regressão entre os valores medidos e os previstos para o teste de dissolução usando o PLS para etambutol.

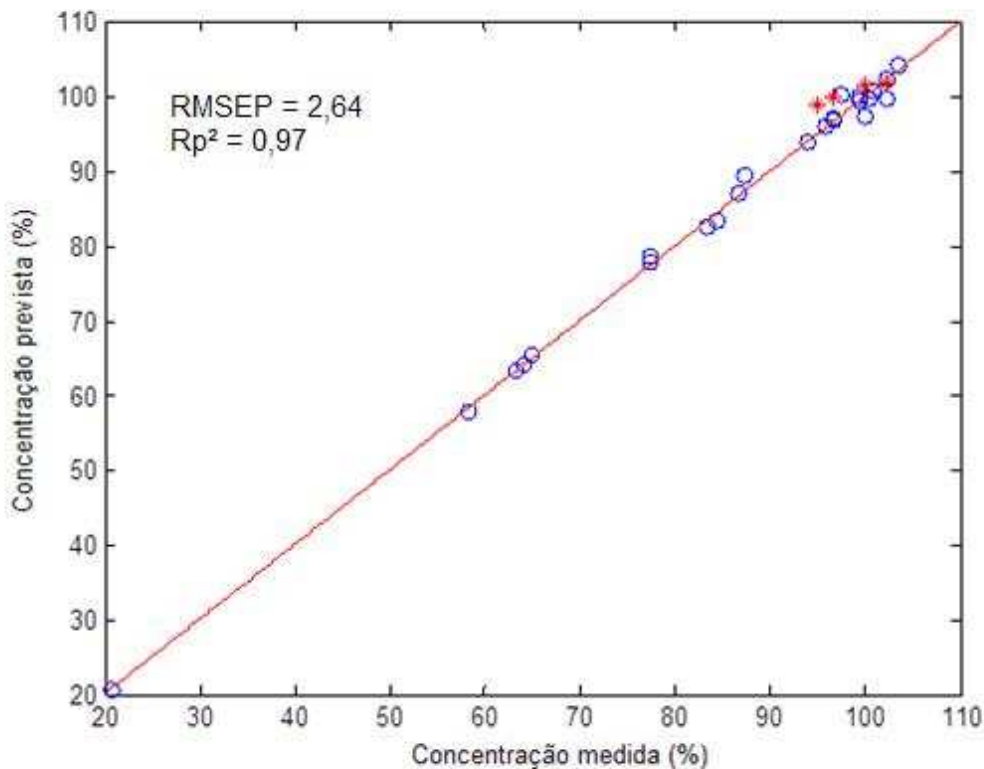


Figura 7: Concentração medida x concentração prevista pela técnica de referência das amostras de calibração e validação para etambutol modelo PLS utilizando. (○) conjunto de calibração; (*) conjunto de validação.

Foi feito um teste t-pareado e comprovou-se que a um nível de 95% de confiança não há diferença estatística entre os resultados previstos pelo NIR e os valores medidos pela técnica de referência.

4.3 FIGURAS DE MÉRITO (FOM)

Após a construção dos modelos multivariados, o passo seguinte foi o cálculo das figuras de mérito. Para tanto foi selecionado o melhor modelo para cada um dos parâmetros, com base no menor valor de RMSEP para cada princípio ativo. As figuras de mérito calculadas, como mostrado na Tabela 6, foram: precisão, linearidade, exatidão, sensibilidade (SEN), seletividade (SEL), LD, LQ e sinal-ruído (S/N). Estes cálculos basearam-se nos valores do vetor NAS para os modelos multivariados PLS. Os valores são mostrados na Tabela 6.

Tabela 6: Figuras de Mérito dos melhores modelos para: isoniazida, rifampicina, pirazinamida e etambutol.

FOM	Isoniazida	Rifampicina	Pirazinamida	Etambutol
RMSEC	2,94	2,98	4,03	1,10
RMSEP	6,28	9,37	7,86	2,64
Rc ²	0,99	0,99	0,99	0,99
Rp ²	0,94	0,98	0,97	0,97
Precisão	2,72	2,00	0,94	0,75
SEN	1,68 x 10 ⁻⁴	3,36 x 10 ⁻⁴	4,62 x 10 ⁻⁴	7,89 x 10 ⁻⁴
SEL	0,09	0,29	0,19	0,13
LD	9,63	14,30	12,33	7,23
LQ	32,13	47,66	41,09	24,09
S/N	5,39 x 10 ⁻⁴	1,6 x 10 ⁻³	1,9 x 10 ⁻³	1,9 x 10 ⁻³

As figuras de mérito multivariadas fornecem um grau adicional de caracterização ao método que não pode ser adquirido a partir de estatísticas de calibração de precisão e outros fatores devem ser levados em conta antes da implantação do método. Questões logísticas, tais como custo de implementação, a compatibilidade com o processo existente / equipamentos de produção e facilidade de transferência podem ditar a escolha.

Não há dados na literatura, com os quais possam ser comparadas as figuras de mérito para modelos PLS usados no estudo de dissolução de fármacos.

Uma região elíptica de confiança conjunta (EJCR, do inglês, *elliptic joint confidence region*) foi desenhada para inclinação e o intercepto ao plotar o previsto versus valores reais dos parâmetros (quatro API), como mostrado na Figura 8. Cálculos EJCR são uma forma conveniente para determinar se está presente viés na determinação de ambos os parâmetros quando se utiliza o modelos multivariados. A Figura 8 mostra o EJCR para a inclinação e intercepção da previsão de teste de dissolução de um conjunto de validação externa e o valor de referência a um intervalo de confiança de 95%. Como pode ser visto em ambas as células, a elipse contém o valor esperado teórico de (1,0), quando construído para a regressão de ensaios de dissolução.

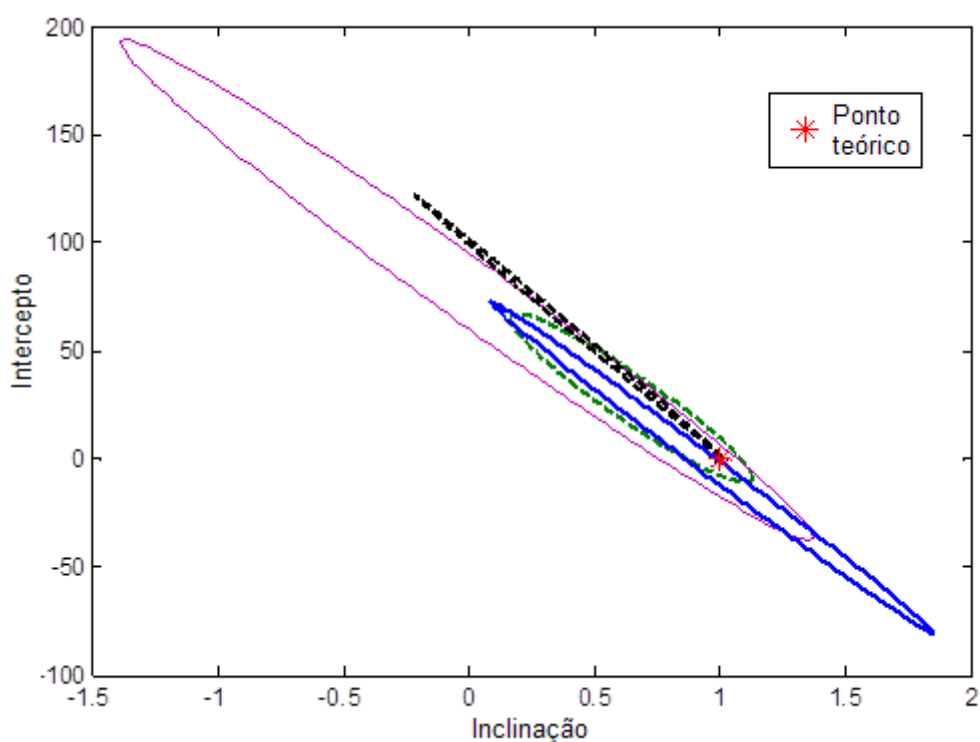


Figura 8: Região de confiança Elíptica conjunta para a inclinação e intercepto da regressão da densidade prevista versus medida de conjunto de validação externa utilizando os modelos iPLS, AG-PLS e PLS para: (a) isoniazida (linha tracejada verde); (b) rifampicina (linha contínua violeta); (c) pirazinamida (linha contínua azul), (d) etambutol (linha tracejada preta). A estrela vermelha marca o ponto (a = 0, b = 1) teórico.

Como é possível verificar na figura acima, dois dos resultados, pirazinamida e etambutol, na previsão não contemplaram o ponto teórico de referência. Isto se deve muito provavelmente a diferença no tamanho de partículas, ao não revestimento dos comprimidos e ao fato de na calibração, terem sido usados comprimidos da produção, ou seja, a variabilidade destas amostras era pequena.

5. CONCLUSÕES

Comprovou-se que a espectroscopia de infravermelho próximo por reflectância difusa, em conjunto com métodos quimiométricos pode ser utilizada como uma técnica analítica satisfatória e eficiente, especialmente no estudo dos perfis de dissolução de comprimidos, quando comparada à atual técnica de referência, CLAE, que é a utilizada pelas indústrias para este tipo de análise.

A porcentagem de dissolução de rifampicina, isoniazida, etambutol e pirazinamida, uma combinação atualmente utilizada em formulações de comprimidos para o tratamento de tuberculose, foram determinadas simultaneamente por meio de medidas espectrais, em conjunto com modelos de calibração multivariada e métodos de seleção de comprimento de onda (iPLS, APS e AG). Quando diferentes técnicas de seleção de variáveis foram conduzidas em dados NIR, foram obtidas melhoras na capacidade de previsão, em comparação com os resultados propostos por Neves et al [12].

A comparação mostra que as maiores reduções de valores de RMSEP foram encontradas quando se utiliza o algoritmo genético ou iPLS. O algoritmo genético melhorou o valor RMSEP por 6,03-8,95%. Enquanto que a seleção feita por iPLS levou a reduções do valor RMSEP de 37,18%. Apenas a detecção de outliers já foi suficiente para incrementar o modelo para etambutol.

As figuras de mérito multivariadas calculadas fornecem um meio de avaliar a capacidade do método e nos permite afirmar a aplicabilidade destes modelos, na indústria farmacêutica, para determinação simultânea dos 4 fármacos em estudo. Para que esta técnica alternativa seja implantada é

necessário ainda que fatores adicionais de logística sejam tomados em consideração.

6. PERSPECTIVAS FUTURAS

Este trabalho é fruto de um trabalho anterior e mostrou ser capaz de melhorar os modelos obtidos inicialmente. Entretanto, deixa algumas perspectivas para aperfeiçoar ainda mais os resultados, como investigação de modelos multivariados não lineares, visto que um algoritmo de seleção de variáveis (AG), cujo princípio é não linear, forneceu bons resultados; elaboração de um planejamento experimental para distribuir melhor o intervalo de concentração do conjunto de amostras de calibração; um número maior de amostras e comparação de outro tipo de pré-processamento para remoção de ruídos espectrais, a transformada Wavelet.

7. REFERÊNCIAS

- [1] R.C.P. Azevedo, G.P. Ribeiro, M.B. Araújo, Desenvolvimento e validação do ensaio de dissolução para captopril em cápsulas magistrais por CLAE, *Revista Brasileira de Ciências Farmacêuticas*, 44 (2008).
- [2] J. Luypaert, D.L. Massart, Y.V. Heyden, Near-infrared spectroscopy applications in pharmaceutical analysis, *Talanta* 72 (2007) 865-883.
- [3] M. Blanco, J. Coello, H. Iturriaga, S. MasPOCH, C. de La èzuela, Near-infrared spectroscopy in the pharmaceutical industry, *Analyst* (Cambridge U.K.) 123 (1998) 135-150.
- [4] M. Blanco, J. Coello, A. Eustaquio, H. Iturriaga, S. MasPOCH, Analytical control of pharmaceutical production steps by near infrared reflectance spectroscopy, *Anal. Chim. Acta* 392 (1999) 237-246.
- [5] M. Blanco, M.A. Romero, Near-infrared libraries in the pharmaceutical industry: a solution for identify confirmation, *Analyst* (Cambridge U.K.) 126 (2001) 2212-2217.
- [6] C. Bodson, E. Rozet, E. Ziemons, B. Evrard, Ph. Hubert, L. Delattre, Validation of manufacturing processo f Diltiazem HCl tablets by NIR spectrophotometry (NIR), *J. Pharm. Biomed. Anal.*, 45 (2007), 356-361.
- [7] A.D. Trafford, R.D. Jee, A.C. Moffat, P. Graham, A rapide quantitative assay of intact paracetamol tablets by reflectance near-infrared spectroscopy, *The Analyst*, 124 (1999), 163-167.
- [8] S. Agrawal, R. Panchagnula, Dissolution test as a surrogate for quality evaluation of rifampicin containing fixed dose combination formulations, *Int. J. Pharm.* 287 (2004) 97-112.

- [9] M.P. Freitas, A. Sabadin, L.M. Silva, F.M. gianotti, D.A. do Couto, E. Tonhi, R.S. Medeiros, G.L. Coco, V.F.T. Russo, J.A. Martins, Prediction of drug dissolution profiles from tablets using NIR diffuse reflectance spectroscopy: A rapid and nondestructive method, *J. Pharm. Biomed. Anal.* 39 (2005) 17-21.
- [10] S.H. Tabasi, V. Moolchandani, R. Fahmy, S.W. Hoag, Sustained release dosage forms dissolution behavior prediction: a study of matrix tablets using NIR spectroscopy, *Int. J. Pharm.* 382 (2009) 1-6.
- [11] H. Abe, M. Otsuka, Effects of lubricant-mixing time on prolongation of dissolution time and its prediction by measuring near infrared spectra from tablets, *Drug Dev. Ind. Pharm.* 38 (2012) 412-419.
- [12] A.C.O Neves, G.M. Soares, S.C. Morais, F.S.L. Costa, D.L. Porto, K.M.G. Lima, Dissolution testing of isoniazid, rifampicin, pyrazinamide and ethambutol tablets using near-infrared spectroscopy (NIRS) and multivariate calibration, *J. Pharm. Biomed. Anal.* 57 (2012) 115-119.
- [13] R. Panchagnula, A. Sood, N. Sharda, K. Kaur, C.L. Kaul, Determination of rifampicin and its main metabolite in plasma and urine in presence of pyrazinamide and isoniazid by HPLC method, 18 (1999), 1013-1020.
- [14] T. Azzouz, A. Puigdoménech, M. Aragay, R. Tauler, Comparison between different data pré-treatment in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method, *Analytica Chimica Acta*, 484 (2003), 121-134.
- [15] R.M. Balabin, S.V. Smirnov, Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data, *Analytica Chimica Acta*, 692 (2011), 63-72.
- [16] H. Xu, B. Qi, T. Sun, X. Fu, Y. Ying, Variable selection in visible and near-infrared spectra: Application to on-line determination of sugar content in pear, *Journal of Food Engineering*, 109 (2012) 142-147.

- [17] P. Valderrama, Avaliação de figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia no infravermelho próximo, Campinas, SP, 2005.
- [18] M.C. Sarraguça, J. A. Lopes, The use of net analyte signal (NAS) in near infrared spectroscopy pharmaceutical applications: Interpretability and figures of merit, *Analytica Chimica Acta*, 642 (2009), 179-185.
- [19] A.C. Olivieri, N.M. Faber, J. Ferre, R. Bosque, J.H. Kalivas, H. Mark, Uncertainty estimation and figures of merit for multivariate calibration, *Pure Appl. Chem.* 78 (2006) 653–661.
- [20] R.W. Kennard, L.A. Stone, Computer aided design of experiments, *Technometrics*. 111 (1969) 137-148.
- [21] H. Working, H. Hotelling, Application of the theory of error to the interpretation of trends, *J. Am. Stat. Assoc.* 24 (1929) 73-85.
- [22] A.G. González, M.A. Herrador, A.G. Asuero, Intra-laboratory testing of method accuracy from recovery assays, *Talanta* 48 (1999) 729-36.
- [23] H.C. Goicoechea, A.C. Olivieri, Sustained prediction ability of net analyte preprocessing methods using reduced calibration sets. Theoretical and experimental study involving the spectrophotometric analysis of multicomponent mixtures, *Analyst* 126 (2001) 1105-1112.

