

PROGRAMA DE PÓS-GRADUAÇÃO EM VIGILÂNCIA SANITÁRIA
INSTITUTO NACIONAL DE CONTROLE DE QUALIDADE EM SAÚDE
FUNDAÇÃO OSWALDO CRUZ

Vanessa Emídio Dabkiewicz

**DESENVOLVIMENTO E VALIDAÇÃO DE UM MÉTODO NÃO DESTRUTIVO DE
DETERMINAÇÃO DE NITROGÊNIO PROTEICO EM VACINA DE FEBRE
AMARELA POR ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO
COM CALIBRAÇÃO MULTIVARIADA**

Rio de Janeiro

2016

Vanessa Emídio Dabkiewicz

DESENVOLVIMENTO E VALIDAÇÃO DE UM MÉTODO NÃO DESTRUTIVO DE
DETERMINAÇÃO DE NITROGÊNIO PROTEICO EM VACINA DE FEBRE
AMARELA POR ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO
COM CALIBRAÇÃO MULTIVARIADA

Dissertação apresentada ao Programa de Pós-Graduação em Vigilância Sanitária do Instituto Nacional de Controle de Qualidade em Saúde da Fundação Oswaldo Cruz como requisito parcial para obtenção do título de Mestre em Vigilância Sanitária.

Orientadores: Shirley de Mello Pereira Abrantes
Ricardo Jorgensen Cassella

Rio de Janeiro

2016

Catálogo na fonte

Instituto Nacional de Controle de Qualidade em Saúde

Biblioteca

Emídio Dabkiewicz, Vanessa

Desenvolvimento e Validação de método não destrutivo de determinação de nitrogênio proteico em vacina de febre amarela por espectroscopia de infravermelho próximo com calibração multivariada / Vanessa Emídio Dabkiewicz.- Rio de Janeiro: INCQS/FIOCRUZ, 2016.

169f.:il

Dissertação (Mestrado em Vigilância Sanitária) - Programa de Pós-Graduação em Vigilância Sanitária, Instituto Nacional de Controle de Qualidade em Saúde, Fundação Oswaldo Cruz. Rio de Janeiro, 2016.

Orientadores: Shirley de Mello Pereira Abrantes
Ricardo Jorgensen Cassella

1. Determinação de Nitrogênio. 2. Infravermelho Próximo. 3. Quimiometria. 4. Validação do Método. 5. Controle de Qualidade. 6. Calibração Multivariada. I Título

Development and Validation of the non-destructive method of determining protein nitrogen in the yellow fever vaccine by near infrared spectroscopy with multivariate calibration

Vanessa Emídio Dabkiewicz

DESENVOLVIMENTO E VALIDAÇÃO DE UM MÉTODO NÃO DESTRUTIVO DE
DETERMINAÇÃO DE NITROGÊNIO PROTEICO EM VACINA DE FEBRE
AMARELA POR ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO
COM CALIBRAÇÃO MULTIVARIADA

Dissertação apresentada ao Programa de Pós-Graduação em Vigilância Sanitária do Instituto Nacional de Controle de Qualidade em Saúde da Fundação Oswaldo Cruz como requisito parcial para obtenção do título de Mestre em Vigilância Sanitária

Aprovado em 20/10/2016

BANCA EXAMINADORA

Maria Helena Wohlers Morelli Cardoso (Doutora)
Instituto Nacional de Controle de Qualidade em Saúde-INCQS

Hiram da Costa Araújo Filho (Doutor)
Instituto Federal do Rio de Janeiro-IFRJ

Daniel da Silva Guedes Junior (Doutor)
Instituto de Tecnologia em Imunobiológicos-Biomanguinhos

Shirley de Mello Pereira Abrantes (Doutora) - Orientadora
Instituto Nacional de Controle de Qualidade em Saúde-INCQS

Ricardo Jorgensen Cassella (Doutor) - Orientador
Universidade Federal Fluminense-UFF

Dedico o presente trabalho
ao meu esposo e aos meus filhos
por serem minha motivação,
agora e sempre.

AGRADECIMENTOS

A Deus, meu principal caminho.

Aos meus orientadores, professora Dra. Shirley de Mello Pereira Abrantes, pelo grande carinho, motivação, compreensão, confiança e conhecimentos transmitidos sobre validação, de grande importância para o desenvolvimento deste conteúdo e ao professor Dr. Ricardo Jorgensen Cassella, pelo grande incentivo, confiança, paciência, dedicação, atenção e conhecimentos transmitidos sobre a técnica de infravermelho e quimiometria.

Ao meu esposo Sérgio, pela total compreensão e apoio para a realização deste trabalho.

Aos meus familiares e amigos que compreenderam as ausências em virtude da dedicação a este estudo.

As pessoas especiais como, o professor Sérgio pelos conhecimentos de estatística, a professora Maria Denise pelos conhecimentos de incerteza e as amigas Camila e Melissa, pelo total incentivo durante todo o trabalho.

Aos meus amigos de trabalho do laboratório físico químico (LAFIQ), que contribuíram para o sucesso e desenvolvimento do projeto.

Aos professores, coordenadores, aos profissionais da secretaria acadêmica e aos colegas de turma do curso de Mestrado Profissional do Programa de Pós-Graduação em Vigilância Sanitária do Instituto Nacional de Controle de Qualidade em saúde.

Ao Instituto Nacional de Controle de Qualidade em Saúde e ao Instituto de Tecnologia de Imunobiológicos por proporcionarem a realização deste trabalho.

Obrigada a todos.

*“O conhecimento pertence à humanidade
e é a tocha que ilumina o mundo”*

Louis Pasteur

*“Nós somos o que repetidas vezes fazemos
Excelência então, não é um ato, mas um hábito”*

Aristóteles

RESUMO

A vacina febre amarela (atenuada) é o meio mais eficaz para prevenir e controlar essa doença. E segundo registros recentes, preocupa sua possível reemergência, refletindo no aumento da solicitação da vacina, tanto pelo Ministério da Saúde para áreas nacionais, quanto para a exportação. O conteúdo de proteína total é um fator de influência na tolerância a este imunobiológico, devendo ser monitorado o limite máximo exigido pela legislação, por meio do teste de nitrogênio proteico; realizado através do método de Kjeldahl, cujas desvantagens incluem o uso de reagentes químicos, a geração de resíduos e um longo tempo de análise. Diante da importância atribuída ao teste de nitrogênio proteico, no controle do produto final da vacina, é clara a motivação para o desenvolvimento e validação deste trabalho. Dentro das diretrizes da boa prática de fabricação (BPF) e fundamentos de vigilância sanitária, foi contemplada a espectroscopia do infravermelho próximo (NIR) com reflectância difusa, associada à calibração multivariada, como uma alternativa para o método clássico. A calibração multivariada mais apropriada foi alcançada para o método mínimos quadrados parciais (PLS) com tratamento de correção multiplicativa de sinal (MSC) e dados centrados na média (CM), utilizando um número mínimo de variáveis latentes igual a 1, com menor valor da raiz quadrada do erro médio quadrático de previsão (0,00334) associado ao maior valor de percentual (90%) de amostras com exatidão na faixa de recuperação de 95 a 105%, na região de 4000-5184 cm^{-1} . O modelo teve comportamento nas suas previsões, considerado repetitivo com um valor de *HorRat* menor que 2, e com desvio padrão relativo menor que 5%. Os critérios de aceitação foram satisfatórios, indicando precisão intermediária adequada. O método não foi considerado robusto para a temperatura de 27°C, mas robusto para as temperaturas avaliadas (22°, 24° e 25°C), para todos os tempos de degelo avaliados (10 minutos a 6 horas), para a troca da lâmpada do instrumento e para a variação da posição de leitura do frasco da amostra. A recuperação para valores de incremento de 0,02 mg/dose em nitrogênio proteico, com adição do padrão de ovoalbumina às amostras, apresentou percentuais entre 95 a 105%. A incerteza dos resultados da metodologia analítica do Kjeldahl (0,002345 mg/dose), representou uma contribuição significativa de 54% no resultado da incerteza dos resultados da calibração multivariada (0,002867mg/dose). Essa avaliação reflete a forte influência da confiabilidade dos valores de referência. A

espectroscopia NIR apresentou vantagens sobre o método Kjeldahl, pois permitiu a eliminação do uso de reagentes perigosos, a redução do tempo total de análise e eliminou a geração de resíduos. A análise pela técnica espectroscópica, por ser não destrutiva, possibilita o reuso das amostras para estudo de estabilidade da vacina. Finalmente, este trabalho apresentou a implementação e validação de modelos multivariados, através de uma robusta técnica de espectroscopia de infravermelho próximo, para determinação de nitrogênio proteico. Mostrando-se uma alternativa viável ao método tradicional e uma possível ferramenta orientativa, para estudos posteriores na área de controle de qualidade, voltadas às exigências da vigilância sanitária.

Palavras-chaves: Determinação de nitrogênio. Infravermelho próximo. Quimiometria. Validação do método. Controle de qualidade. Calibração multivariada.

ABSTRACT

The yellow fever vaccine (attenuated) is the most effective way to prevent and control the disease. And according to recent records, concerns its possible reemergence, reflecting the increased request by both the Ministry of Health to national areas, and for export. The total protein content is a factor influencing the tolerance to this immunobiological and should be monitored the limit required by law, through the protein nitrogen test; performed by the Kjeldahl method, which disadvantages include the use of chemical reagents, waste generation and a long analysis time. Given the importance attached to the protein nitrogen test in control of the final product of the vaccine is a clear motivation for the development and validation of this work. Within the guidelines of good manufacturing practice (GMP) and health surveillance foundations, was awarded the near infrared spectroscopy with diffuse reflectance, connected with multivariate calibration as an alternative to the traditional method. The most appropriate multivariate calibration has been achieved for the PLS method and MSC, CM treatment, using a minimum number of latent variables equal to 1, less RMSEP value (0.00330) associated with the highest percentage value (91%) samples exactly a 95 to 105%, recovery band in the region of 4000-5184 cm^{-1} . The model had behavior in their forecasts, considered repetitive with less Horrat value than 2, and relative standard deviation less than 5%. The acceptance criteria were satisfactory, indicating adequate intermediate precision. The method is not considered to be robust to 27° C, but robust to measured temperatures (22°, 24° and 25° C) for all time evaluated thaw (10 minutes to 6 hours) for the exchange instrument lamp and the variation of the sample vial reading position. The recovery increased to values of 0.02 mg/dose in protein nitrogen, with standard addition of the ovalbumin samples showed percentages between 95 and 105%. The uncertainty of the result of the Kjeldahl analytic methodology (0,002345 mg/dose), was a significant contribution of 54% in the uncertainty of the result of multivariate calibration (0,002867 mg/dose). This assessment reflects the strong influence of the reliability of the reference values. NIR spectroscopy showed advantages over the Kjeldahl method, because it allowed the elimination of the use of hazardous reagents, the reduction in the total analysis time and eliminated the generation of residues (leavings). The analysis by spectroscopic technique by been non-destructive, enables the reuse of samples for study of vaccine stability. Finally, this work

presented the implementation and validation of multivariate models, through a robust technique of near infrared spectroscopy for the determination of protein nitrogen, proving a viable alternative to the traditional method and a possible tool for guidance for further studies in the quality control area, geared to the demands of health surveillance.

Keywords: Nitrogen determination. Near infrared. Chemometrics. Validation method. Quality control. Multivariate calibration.

LISTA DE ILUSTRAÇÕES

Figura 1. Fluxograma de produção simplificada da vacina de febre amarela.....	27
Figura 2. Principais bandas de absorção em NIR e sua localização	33
Figura 3. Modos vibracionais das moléculas pertencentes ao grupo AX2	34
Figura 4. Representação esquemática de reflectância difusa num sólido.....	36
Figura 5. Ilustração gráfica dos procedimentos da análise univariada versus a análise multivariada.....	38
Figura 6. Fluxograma do método de calibração.	39
Figura 7. Organização de dados para calibração multivariada em espectrometria NIR.....	39
Figura 8. A) Espectrômetro Antaris II FT-NIR Analyser Thermo Scientific, associado ao sistema de microcomputador. B) Compartimento de leitura, com o frasco de VFC.	52
Figura 9. A) Digestor Buchi Speed digester K-436. B) Purificador de gases Buchi Scrubber B-414. C) Destilador Marconi.....	53
Figura 10. Matriz dos dados espectrais executada em Excel.....	74
Figura 11. Histograma resultante da análise cluster (HCA).....	74
Figura 12. Espectros médios das amostras sem tratamento.....	77
Figura 13. Espectros médios das amostras com tratamento.....	77
Figura 14. Espectros médios das amostras com tratamento SNV.	78
Figura 15. Espectros médios das amostras aplicada a primeira derivada com alisamento Savitsky-Golay(5p e 2ª Ord.).....	78
Figura 16. Espectros médios das amostras aplicada a segunda derivada com alisamento Savitsky-Golay(5p e 2ª Ord.).....	79
Figura 17. Espectros médios tratados com MSC e CM na região proposta pela literatura	80
Figura 18. Análise discriminante na região 4545-5000cm ⁻¹ , com tratamento MSC e CM. A) Distâncias entre os níveis 2 e 1. B) Distâncias entre os níveis N1 e N3.	81
Figura 19. Análise discriminante na região 4545-5000cm ⁻¹ , com tratamento MSC e CM. C) Distâncias entre os níveis N2 e N3	82
Figura 20. Espectros médios tratados com MSC e CM	82
Figura 21. Espectros médios tratados com MSC e CM na região 9091-9999 cm ⁻¹ propostas pela literatura.....	83

Figura 22. Histograma com as faixas de teores de nitrogênio proteico obtidas pelo método de referência kjeldahl	85
Figura 23. Histograma da distribuição das amostras do set de calibração.....	86
Figura 24. Representação gráfica (A) de correlação entre valores de referência e valores previstos pelo modelo PLS.MSC.CM (1 fator) e (B) resíduos dos valores previstos e valores de referência.....	89
Figura 25. Representação gráfica (A) de correlação entre valores de referência e valores previstos pelo modelo PLS.SNV.CM (1 fator) e (B) resíduos dos valores previstos e valores de referência.....	89
Figura 26. Representação gráfica para modelo PLS.MSC.CM	91
Figura 27. Representação gráfica para modelo PLS.SNV.CM.....	91
Figura 28. Representação da matriz para cálculo dos coeficientes de correlação ...	92
Figura 29. Correlação entre os coeficientes de correlação com números de onda. .	93
Figura 30. Regiões da literatura (STUART, 2004) que representam estiramentos de NH.....	93
Figura 31. Representação gráfica entre os coeficientes de regressão do modelo e número de onda(cm^{-1}), calculados através do Unscrambler.	96

LISTA DE TABELAS

Tabela 1. Composição da vacina para 5 e 10 doses.	27
Tabela 2. Especificações do produto final da vacina (atenuada).....	28
Tabela 3. Comprimentos de onda, para absorção de grupos funcionais na região do infravermelho próximo.....	33
Tabela 4. Classificação dos testes, segundo sua finalidade.....	44
Tabela 5. Ensaio necessários para a validação do método analítico, segundo sua finalidade.....	44
Tabela 6. Faixa de recuperação correspondente a concentração limite.....	63
Tabela 7. Variações avaliadas na verificação da robustez.....	64
Tabela 8. Procedimento para avaliação da recuperação de padrão de ovoalbumina nas vacinas liofilizadas.....	67
Tabela 9. Procedimento para avaliação da influência espectral dos componentes da matriz nas vacinas liofilizadas.....	68
Tabela 10. Classificação hierárquica (HCA) dos espectros obtidos no NIR, para as amostras de vacinas de febre amarela.....	75
Tabela 11. Classificação hierárquica dos espectros obtidos no NIR para as amostras de vacinas de febre amarela.....	76
Tabela 12. Tratamento de valores extremos.....	79
Tabela 13. Análise discriminante, utilizando o cálculo de Mahalanobis.....	83
Tabela 14. Faixa de resultados dos valores de Kjeldahl.....	84
Tabela 15. Faixas de concentrações (mg/dose) vs frequência.....	84
Tabela 16. Faixa de resultados dos valores de selecionados para cada conjunto. ..	85
Tabela 17. Faixas de concentrações do set de calibração.....	86
Tabela 18. Coeficientes calculados pelo teste de Ryan Joiner para normalidade dos dados da regressão.....	87
Tabela 19. Modelos avaliados com 99 valores para calibração e 99 para a validação, utilizando as 3 regiões sugeridas pela literatura.....	88
Tabela 20. Avaliação do número ideal de variáveis latentes (VL) para o modelo PLS.....	90
Tabela 21. Parâmetros estatísticos para PLS.MSC.CM.....	91
Tabela 22. Parâmetros estatísticos para PLS.SNV.CM.....	91

Tabela 23. Parâmetros estatísticos para as regiões limitadas pelos coeficientes de correlação.....	94
Tabela 24. Parâmetros estatísticos para regiões sugeridas pela literatura para estiramentos NH.....	95
Tabela 25. Precisão das leituras espectrais (no mesmo dia).....	97
Tabela 26. Precisão intermediária	98
Tabela 27. Resultados (mg/dose) e percentual(%) de exatidão para os valores da validação externa.	99
Tabela 28. Avaliação da temperatura ambiente.	101
Tabela 29. Avaliação do tempo de leitura a partir do degelo (mesmo dia).	101
Tabela 30. Avaliação após troca de lâmpada	102
Tabela 31. Avaliação da variação da posição do frasco.	103
Tabela 32. Resultados para erros sistemáticos no modelo de calibração.	104
Tabela 33. Avaliação para incremento de 0,02 mg.dose ⁻¹ em nitrogênio proteico através de solução padrão de ovoalbumina.	104
Tabela 34. Teste para incremento de 0,01mg.dose ⁻¹ em nitrogênio proteico através de solução padrão de ovoalbumina.....	105
Tabela 35. Teste de recuperação de incremento de 0,02 mg.dose ⁻¹ em nitrogênio presente na composição do excipiente.	106
Tabela 36. Teste de recuperação de incremento de 0,02mg.dose ⁻¹ em nitrogênio presente na composição do excipiente.	107
Tabela 37. Precisão das amostras controles durante os meses e dias de estudo	108
Tabela 38. Previsão de lotes de rotina produzidos em 2016.	110
Tabela 39. Fontes de incerteza da fatoração do ácido	111
Tabela 40. Fontes de incerteza da titulação da amostra pós destilação.....	112
Tabela 41. Cálculo dos coeficientes de sensibilidade para a etapa de fatoração do ácido.....	112
Tabela 42. Cálculo dos coeficientes de sensibilidade para a etapa de titulação da amostra	113
Tabela 43. Estimativa da incerteza da fatoração do ácido.....	114
Tabela 44. Estimativa da incerteza da metodologia de referência (Kjeldahl).....	119
Tabela 45. Fontes de incerteza do método multivariado.	115
Tabela 46. Estimativa da incerteza para a calibração multivariada.	115
Tabela 47. Resumo comparativo entre os métodos.....	117

LISTA DE ABREVIATURAS

- ABNT**-Associação Brasileira de Normas Técnicas
- ANOVA**-Análise de variância
- ANVISA**-Agência Nacional de Vigilância Sanitária
- AOAC**-*International association of analytical communities*
- ASTM** - *American Society for Testing and Materials*
- BIO-MANGUINHOS** - Instituto de Tecnologia em Imunobiológicos.
- BPF**-Boas Práticas de Fabricação
- CDC**-Centros de Controle e Prevenção de Doenças (USA)
- CERF**-Fundo Central das Nações Unidas de Resposta de Emergência (*The Central Emergency Response Fund*)
- CLS**-*Classical Least Squares* (Quadrados Mínimo Clássicos)
- CM**-*Mean centering* (Método centrar na média)
- CQ**-Controle de qualidade
- CV**-Coeficiente de variação
- DOQCgcre-020** - Documento de Normatização e Qualidade Industrial do INMETRO
- DPR**-Desvio padrão relativo
- DRIFTS**-*Diffuse Reflection Infrared Fourier Transform Spectroscopy* (Espectroscopia por refletância difusa com transformada de Fourier)
- ECHO**- *European Community Humanitarian Aid Office* (Serviço Humanitário da Comunidade Europeia)
- FA** - Febre amarela
- FIOCRUZ** - Fundação Oswaldo Cruz
- FT-NIR**- *Fourier Transform Near Infrared* (Espectroscopia de infravermelho próximo com transformada de Fourier)
- GL**- Graus de liberdade
- GUM**- Guia para a Expressão da Incerteza de Medição
- HCA** - *Hierarchical Cluster Analysis* (Análise de Agrupamentos Hierárquicos)
- ICH** - *International Conference on Harmonization*
- INCQS** - Instituto Nacional de Controle de Qualidade em Saúde
- INMETRO** - Instituto Nacional de Metrologia, Qualidade e Tecnologia
- ISO/IEC 17025**-*International Standardization Organization / International Electrotechnical Commission*

IUPAC - *International Union of Pure and Applied Chemistry*

KBr-Brometo de potássio

LED - *Light Emitting Diode*

LOD - Limite de detecção

LOQ - Limite de quantificação

MAPA - Ministério da Agricultura, Pecuária e Abastecimento

MIR-*Mid-Infrared* (Espectroscopia no infravermelho médio)

M-média aritmética

MLR - *Multiple Linear Regression* (Regressão Linear Múltipla)

MSC- *Multiplicative Scatter Correction* (Correção Multiplicativa de Sinal)

MSEC- *Mean Square Error of Calibration*(Erro quadrático médio de calibração)

NBR - Norma Brasileira

NIR- *Near Infrared Spectroscopy* (Espectroscopia no infravermelho próximo)

OMS - Organização Mundial de Saúde

PBS- Solução para preparo de solução-padrão de ovoalbumina

PCA - *Principal Component Analysis* (Análise por componentes principais)

PCR - *Principle Component Regression* (Regressão por componentes principais)

PI - Programas de Imunização

PLS - *Partial Least Squares* (Regressão por mínimos quadrados parciais)

PNI - Programa Nacional de Imunização

POP - Procedimento operacional padrão

PRESS- Soma dos quadrados dos resíduos dos erros de previsão (*Prediction Residual Error Sum of Square*)

R(SEC)- Coeficiente de correlação atribuída aos dados de calibração

R(SECV)- Coeficiente de correlação atribuída aos dados de validação cruzada

R(SEP)- Coeficiente de correlação atribuída aos dados previstos

RDC 17/2010- Resolução da Diretoria Colegiada nº 17, de 16 de abril de 2010

RMSEC- Raiz quadrada do Erro quadrático médio para valores de calibração (*Root Means Squares Errors*)

RMSECV- Raiz quadrada do Erro quadrático médio para na validação cruzada (*Root Mean Square Error of Cross Validation*)

RMSEP- Raiz quadrada do Erro quadrático médio de previsão (*Root Means Squares Errors of Prediction*)

SDV - Desvio padrão dos erros de validação(*Standard Deviation of Validation*)

SG - Alisamento Savitsky-Golay.

SNV- Variação de padrão normal (*Standard normal variate*)

US-EPA - *United States-Environmental Protection Agency*

US-FDA - *United States-Food and Drug Administration*

VFC - Vacina de febre amarela de 10 doses

VIM - Vocabulário Internacional de Metrologia

VL - Variável Latente

WFI- *Water for injectables* (água para injetáveis)

SUMÁRIO

1 INTRODUÇÃO	22
1.1 A VACINA FEBRE AMARELA.....	22
1.1.1 A história na saúde pública: a trajetória de uma luta sem fim.	22
1.1.2 Aspectos clínicos e epidemiológicos.....	23
1.1.3 Produção da vacina	25
1.1.4 A importância do teste de nitrogênio proteico	28
1.2 VIGILÂNCIA SANITÁRIA E CONTROLE DE QUALIDADE.....	29
1.2.1 Estudo de estabilidade.....	30
1.3 MÉTODO CLÁSSICO KJELDAHL	31
1.4 ESPECTROSCOPIA DO INFRAVERMELHO PROXIMO (NIR).....	31
1.4.1 Instrumentação e acessórios	34
1.4.2 Quimiometria e a calibração multivariada	37
1.5 VALIDAÇÃO INTRALABORATORIAL DE MÉTODO	42
1.5.1 Faixa de trabalho	45
1.5.2 Curva analítica e linearidade.....	46
1.5.3 Seletividade.....	46
1.5.4 Sensibilidade.....	46
1.5.5 Efeito de matriz	46
1.5.6 Precisão	47
1.5.7 Exatidão	47
1.5.8 Robustez.....	47
1.5.9 Limites de detecção e quantificação	47
1.6 ESTIMATIVA DA INCERTEZA DA CALIBRAÇÃO MULTIVARIADA	48
1.7 JUSTIFICATIVA	49
2 OBJETIVOS.....	50
2.1 OBJETIVO GERAL	50
2.2 OBJETIVOS ESPECÍFICOS.....	50
3 MATERIAIS E MÉTODOS	51
3.1 MATERIAIS	51
3.1.1 Amostras.....	51
3.1.2 Vidrarias.....	51
3.1.3 Reagentes.....	51
3.1.4 Equipamentos	52
3.2 MÉTODOS	53
3.2.1 Aquisição dos espectros e Análise de agrupamentos hierárquicos (<i>Hierarchical Clusters Analysis – HCA</i>).....	53

3.2.2 Pré tratamento dos dados espectrais.....	53
3.2.3 Determinação de nitrogênio proteico (Kjeldahl)	54
3.2.4 Análise de valores discrepantes e Discriminação dos valores de referência	55
3.2.5 Seleção do set de validação e do set de calibração	55
3.2.6 Teste de normalidade	55
3.2.7 Calibração e análise quimiométrica	57
3.2.8 Otimização das regiões de trabalho.....	59
3.2.9 Validação do método de calibração multivariada	60
3.2.10 Outras avaliações	68
3.2.11 Estimativa da incerteza da calibração multivariada.....	69
4 RESULTADOS E DISCUSSÕES.....	74
4.1 MODELAGEM MULTIVARIADA.....	74
4.1.1 Análise de agrupamento hierárquico (<i>hierarchical clusters analysis</i> – HCA)	74
4.1.2 Pré tratamento nos dados espectrais coletados	76
4.1.3 Método de referência (kjeldahl), o teste de <i>grubbs</i> e análise discriminatória.	79
4.1.4 Seleção do set de calibração e do set de validação.	84
4.1.5 Avaliação da normalidade dos dados de calibração	85
4.1.6 Modelagem (avaliações dos métodos de calibração, parâmetros de pré- tratamentos e variáveis latentes).	87
4.1.7 Otimização da região de trabalho	91
4.2 VALIDAÇÃO DO MÉTODO DE CALIBRAÇÃO MULTIVARIADA	96
4.2.1 Precisão	96
4.2.2 Exatidão	98
4.2.3 Robustez.....	100
4.2.4 Teste de erros sistemáticos no modelo – viés (<i>bias</i>)	103
4.2.5 Recuperação.....	104
4.2.6 Avaliação da influência espectral dos excipientes da matriz.....	105
4.3 OUTRAS AVALIAÇÕES.....	108
4.3.1 Precisão (<i>inter meses</i>)	108
4.3.2 Avaliação da previsão do modelo com lotes de rotina	109
4.4 ESTIMATIVA DA INCERTEZA DA CALIBRAÇÃO MULTIVARIADA	110
4.4.1 Estimativa da incerteza dos resultados do método de referência	110
4.4.2 Estimativa da incerteza dos resultados da calibração multivariada	115
4.5 COMPARAÇÃO ENTRE OS MÉTODOS	116
5 CONCLUSÃO	118
6 PROPOSTAS FUTURAS	119
6.1 Monitoramento da calibração multivariada.....	119
6.2 Linearidade, sensibilidade e limites para calibração multivariada	119
REFERÊNCIAS.....	120

APÊNDICE A - PROCEDIMENTO OPERACIONAL PADRÃO PARA A DETERMINAÇÃO DE NITROGÊNIO PROTEICO EM VACINAS DE FEBRE AMARELA 10 DOSES POR NIR COM CALIBRAÇÃO MULTIVARIADA (POP).....	136
ANEXO A- ESPECTROS NIR.....	145
ANEXO B- ESTUDO DAS REGIÕES	150
ANEXO C- ESTUDO GRÁFICO DAS REGIÕES.....	167

1 INTRODUÇÃO

1.1 A VACINA FEBRE AMARELA

*“Jamais uma grande descoberta
foi feita, sem um palpite ousado”
Isaac Newton.*

1.1.1 A história na saúde pública: a trajetória de uma luta sem fim.

A existência da febre amarela (FA) na história é relatada, tanto no Brasil quanto em outros países do continente americano, entretanto com foco sanitário, apenas após meados do século XIX. Os primeiros registros mais notórios apresentam surto epidêmico em Yukatan (em 1648, México), na Filadélfia (em 1793, Estados Unidos) e no Mississippi, em 1878, nos Estados Unidos (PETTERSON, 1992; MONATH, 2013). No Brasil, esses relatos descrevem mais de 4 mil mortes, após a chegada de um navio negreiro ao Rio de Janeiro em 1849, talvez o registro mais grave da história no país (BENCHIMOL, 2001).

Em 1879, no Brasil, Domingos José Freire acreditava que o causador da doença seria um microrganismo (BENCHIMOL, 2001). Enquanto, Josiah Clark Nott do Alabama, em 1848, e Luis Daniel Beauperthuy da Venezuela, em 1854, já inferiam um inseto como vetor da doença. Somente a equipe chefiada por Walter Reed, em 1900, conseguiu demonstrar a hipótese publicada por Carlos Juan Finlay, pesquisador Cubano, que especificara o mosquito *Aedes aegypti* como o verdadeiro transmissor, em 1880 (BENCHIMOL, 2001; BRYAN, 2004; STAPLES; MONATH, 2008).

As descobertas que viabilizaram os avanços na epidemiologia, no diagnóstico e na prevenção da febre amarela, foram através dos estudos de Max Theiler, em 1930. Tais pesquisas resultaram em seu Prêmio Nobel, em 1951, através do qual apresentava um vírus atenuado com resposta imunitária protetora em macacos e em seres humanos, cuja cepa 17D, serviria como a base da vacina atual. Uma segunda vacina de vírus vivo atenuado, com passagens em cérebro de rato (neurotrópico francês – FNV), foi desenvolvida a partir de uma cepa diferente, isolada em Dakar em 1927 (STAPLES; MONATH, 2008).

Oswaldo Cruz, no Rio de Janeiro, através do Serviço de Profilaxia da Febre Amarela, em 1930, inicia a vigilância sanitária nas ruas, neutralizando depósitos de água com larvas do *Aedes aegypti*. Um conjunto de laboratórios em Manguinhos é construído por Oswaldo Cruz, com o propósito de novas linhas de pesquisa, de fabricação de mais soros e vacinas e ainda, do ensino da microbiologia, embasados nas vertentes do Instituto Pasteur de Paris (BENCHIMOL, 2001).

A Fundação Rockefeller norte americana inicia junto ao Brasil, um programa de erradicação da febre amarela, resultando na vitória contra o *Aedes aegypti*, em meados de 1950 até 1968. A partir deste momento, voltou a ter destaque na história do país. Desde então, pesquisadores de Manguinhos tentam várias metodologias para produzir uma vacina eficiente, contornando os efeitos adversos (BERTOLLI FILHO, 2002; LACERDA; MELLO, 2003; BRYAN, 2004). A Organização Mundial de Saúde (OMS), em 2002, estimou que a febre amarela ainda acometerá, cerca de 200.000 pessoas, com até 30.000 mortes por ano.

1.1.2 Aspectos clínicos e epidemiológicos

Sob o ponto de vista clínico, a febre amarela pode se apresentar de forma assintomática, moderada, grave e maligna. Pode apresentar febre alta, calafrios, cansaço, dor de cabeça, dor muscular, náuseas e vômitos por cerca de três dias. Na forma mais grave da doença, podem ocorrer insuficiências hepática e renal, icterícia (olhos e pele amarelados), manifestações hemorrágicas e cansaço intenso (BRASIL, 2004; SHEFF, 2005; BIO-MANGUINHOS, 2014a).

Epidemiologicamente, a febre amarela divide-se em duas formas, rural/silvestre e urbana, sendo a espécie *Aedes aegypti* a transmissora da febre amarela urbana, enquanto a silvestre é transmitida pelos gêneros *Haemagogus* e *Sabethes* (América Latina) e pela espécie *Haemagogus janthinomys* no Brasil (VASCONCELOS, 2002, 2003; GOMES et al, 2010). A forma urbana na América não ocorre desde 1954, porém, ainda hoje, ocorre na África. Geralmente, apresenta grande risco o grupo dos indivíduos jovens, do sexo masculino, que realizam atividades agropecuárias, bem como os viajantes (VASCONCELOS, 2003).

O entendimento detalhado da geografia da febre amarela é fundamental para o uso adequado da vacina. As áreas de maior risco de retomada da doença estão nas Américas e incluem a costa e as cidades do interior infestadas pelo *Aedes*

aegypti, onde a população não é vacinada. Europa, Oriente Médio, Ásia, Austrália e Oceania também correm o risco de introdução da febre amarela. Entretanto, a probabilidade de introdução e disseminação de vírus da febre amarela na Índia e outras partes da Ásia são pequenas. Isto ocorre, pois as cepas de *Aedes aegypti* têm baixa competência vetorial para vírus da febre amarela e a proteção cruzada, principalmente pela dengue, proporcionando imunidade à quase todas as pessoas que residem na Ásia (MONATH, 2013).

Em 2008, a Organização Mundial de Saúde (OMS) e o Centro de Controle e Prevenção de Doenças (CDC) realizaram uma avaliação abrangente e desenvolveram critérios para a classificação de risco endêmico (JENTES et al, 2011; MONATH, 2013; BEASLEY, 2015). Os trabalhos de Monath (2013) e Monath e Vasconcelos (2015) citam as regiões da África e do sul da América consideradas endêmicas. Da mesma forma, Jentes (2011) e Garske (2013) confirmaram a estimativa de casos com mortes nos países africanos, no período de 2006 a 2012.

Um surto de febre amarela foi detectada em Luanda, Angola, do final de dezembro de 2015 até junho de 2016. Angola relatou 2.893 casos suspeitos de febre amarela, com 325 mortes. Entre esses casos, 788 foram confirmados por laboratórios. Três países comunicaram casos confirmados de febre amarela importados de Angola: República Democrática do Congo (44 casos), Quênia (dois casos) e da República Popular da China (11 casos). Isso destaca o risco de propagação internacional através de viajantes (WHO, 2016).

Em 2014, o Ministério da Saúde foi notificado de 282 casos humanos suspeitos. A região Sudeste notificou a maior parte dos casos, 143 (51,3%), seguida das regiões Centro-Oeste, com 52 (18,6%), e Sul, com 49 (17,6%), embora nenhum caso tenha sido confirmado fora da região considerada endêmica (região amazônica). Os estados com maior frequência de notificação foram São Paulo, com 111 (39,8%), Goiás, com 34 (12,2%), Paraná, com 21 (7,5%), e Minas Gerais, com 19 (6,8%). Estas áreas foram afetadas recentemente, entre 2007 e 2009, durante a reemergência do vírus fora da região amazônica no Brasil (BRASIL, 2014).

O número de casos da doença vem aumentando no mundo, nas duas últimas décadas, devido a fatores de difícil controle, como desmatamento, urbanização, movimentos migratórios, redução da imunidade coletiva e, por último, por mudanças climáticas. Entre julho de 2014 e junho de 2016 no Brasil foram confirmados 11 casos humanos, Goiás (6), Pará (2), Mato Grosso do Sul (1), São Paulo (1) e no Rio

Grande do Norte (1). No Peru, fronteira com Acre e Amazonas, foram confirmados 37 casos (RAMOS FILHO, 2016; WHO, 2016). Contudo, a febre amarela (FA) é considerada uma doença emergente ou reemergente de grande importância (GARDNER; RYMAN, 2010).

No Brasil, segundo a Lei 8080/1990, compete à vigilância epidemiológica, a prevenção de surtos de FA em humanos, através de um conjunto de ações que proporciona o conhecimento, a detecção ou a prevenção nos fatores que atinjam a saúde pública (BRASIL, 1990).

O controle e a prevenção da FA são realizados mediante a contenção dos mosquitos vetores, através de medidas de eliminação da proliferação do inseto, como por exemplo, uso de inseticidas nos criadouros (WHO, 2014). O controle e a prevenção dessa doença, também incluem a vacinação que confere proteção individual e coletiva, criando imunidade e assim, evitando epidemias através da propagação geográfica da doença, sendo de suma importância a produção da vacina. Entretanto, uma vez instalada a doença, o tratamento é somente sintomático, pois não existe nenhum método específico e efetivo contra a doença (REISDÖRFER, 2011).

Entre 2007 e 2016, 14 países realizaram campanhas de vacinação preventivas da febre amarela: Benin, Burkina Faso, Camarões, República Centro Africano, da Costa do Marfim, Gana, Guiné, Libéria, Mali, Senegal, Serra Leoa e Togo, entre outros. A Iniciativa das campanhas de prevenção da Febre Amarela é apoiado financeiramente pela Vaccine Alliance Gavi, o Serviço Humanitário da Comunidade Europeia (*European Community Humanitarian Aid Office*, ECHO), Fundo Central das Nações Unidas de Resposta de Emergência (*The Central Emergency Response Fund*, CERF), Ministérios da Saúde e os parceiros em nível nacional (WHO, 2016).

1.1.3 Produção da vacina

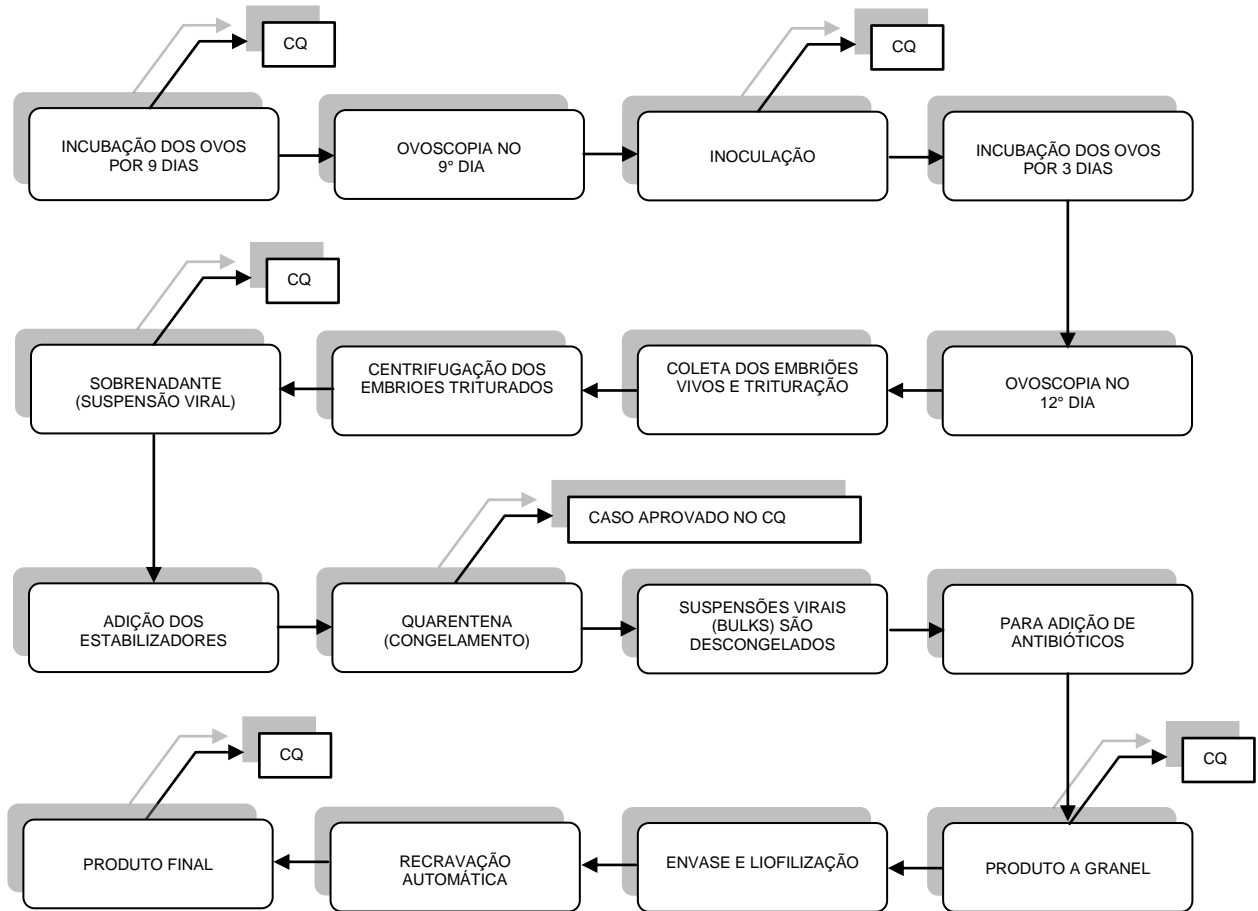
A vacina foi produzida a partir da cepa 17D, constituída do vírus Asibi extraído de um africano que morreu com a doença. Essa amostra foi trazida de Nova York por Hugh H. Smith, em 1937, e começou a ser produzida no Rio de Janeiro. É produzida hoje, por Bio-Manguinhos, com tecnologia 100% brasileira (BENCHIMOL, 2001). De 2001 a 2012, 131 milhões de doses foram produzidas para a exportação,

suprindo cerca de 74 países na América do Sul, África, Ásia, Oceania, América do Norte, Central e Caribe (HOMMA, 2013).

Os produtores da vacina no mundo são Bio-Manguinhos (Brasil), China National Biotech Group (China), Sanofi Pasteur (França), Chumakov Institute (Rússia), Instituto Pasteur (Senegal), Crucell Berna (Switzerland), Sanofi Pasteur (Estados Unidos). Entretanto, apenas Bio-manguinhos, Sanofi Pasteur, Chumakov Institute e Instituto Pasteur, no Senegal, são pré-qualificados pela OMS; o que significa que a vacina pode ser comprada pelo Fundo das Nações Unidas para a Infância (UNICEF) para Programas de Imunizações (PI) ou em casos de emergência. O número de fabricantes tem diminuído, desde a década de 1990, e instalações no Reino Unido, Holanda, Alemanha, Índia, Colômbia, Austrália, Nigéria e África do Sul não estão mais ativas (MONATH, 2013).

A produção da vacina (Figura 1) é baseada no sistema de lote semente a partir da cepa 17D, da qual, por meio de passagens em ovos embrionados de galinha livre de patógenos, se origina o lote semente secundário, cuja amostragem passa pelo controle de qualidade (CQ). Esses ovos são incubados e após o 9º dia ocorre a inoculação do vírus. Depois de mais três dias de incubação, ocorre a coleta dos embriões, trituração e centrifugação para separação da suspensão viral. Após a clarificação da suspensão viral para remoção de resíduos celulares, os estabilizadores são adicionados, resultando na suspensão viral estabilizada. A ela são adicionados outros estabilizadores, além de antibióticos. Antes do envase e liofilização, o produto é analisado principalmente, quanto à esterilidade e a concentração de vírus. Após esta etapa, ocorre o envase do produto e posterior liofilização, dando origem ao produto final, que também é controlado pelo CQ. Aprovação pelo CQ, os frascos são rotulados e embalados (BENCHIMOL, 2001; BRASIL, 2010; REISDORFER, 2011). A vacina é comercializada conforme as especificações e composição final apresentados nas tabelas 1 e 2, com destaque o teste de nitrogênio proteico, que é foco deste estudo.

Figura 1. Fluxograma de produção simplificada da vacina de febre amarela.



Fonte: adaptado (BENCHIMOL, 2001).

Tabela 1. Composição da vacina para 5 e 10 doses.

Virus 17D	1.000 DL ₅₀
Sacarose	0,8 mg/dose (C ₁₂ H ₂₂ O ₁₁ -MM 342,2965 g/mol)
Glutamato de sódio	4,0 mg/dose (C ₅ H ₈ NaNO ₄ -MM 169,11 g/mol)
Sorbitol	8,5 mg/dose (C ₆ H ₁₄ O ₆ -MM 182,17 g/mol)
Gelatina bovina hidrolisada	5,0 mg/dose (85% de proteína)
Eritromicina	1,5 ug/dose (C ₃₇ H ₆₇ NO ₁₃ -MM 733,93 g/mol)
Canamicina	5,0 ug /dose (C ₁₈ H ₃₆ N ₄ O ₁₁ -MM 484,44 g/mol)

Fonte: (BIOMANGUINHOS, 2014b).

Tabela 2. Especificações do produto final da vacina (atenuada).

Aspecto	Forma liofilizada que após a reconstituição pode apresentar coloração
Potência	1000 DL50 equivalente a 3,73 UFP/dose ou 3,00 UI/dose
Termoestabilidade	Após permanecer por 14 dias armazenada a 37°C, a vacina deve estar com o título de 1000 DL50 equivalente a 3,73 UFP/dose ou 3,00 UI/dose e não menos que 1 log 10 em relação ao título da potência
Umidade residual	No máximo 3%
Identidade	Positivo
Esterilidade	Ausência de crescimento bacteriano e fúngico
Endotoxinas bacterianas	No máximo 10 UE/mL
Ovoalbumina residual	Menor ou igual a 5 µg/dose
Nitrogênio proteico	Menor ou igual a 0,25 mg/dose

Fonte: (BIOMANGUINHOS, 2014b).

1.1.4 A importância do teste de nitrogênio proteico

Diversas proteínas da clara do ovo e seus peptídeos apresentam atividade imunomoduladora. Somente a ovoalbumina corresponde a 54% da proteína da clara (POSSAS, 2012). A gelatina, a sacarose, o sorbitol e o glutamato funcionam como estabilizantes da vacina. Entretanto, as bases proteicas são os principais alérgenos da vacina e a quantidade de proteínas do ovo, encontradas nela é suficiente para induzir reações graves de hipersensibilidade (ALPHA DIAGNOSTIC INTERNATIONAL, 2012 apud POSSAS, 2012). A contraindicação a pacientes alérgicos depende justamente da quantidade de proteína presente na vacina (SILVA; ALVES, 2011). Quando empregado na produção de vacinas, o conteúdo utilizado de ovoalbumina é um fator de influência na tolerância ao imunobiológico, cujo resíduo de produção deve ser monitorado (BRASIL, 2010), assim como a proteína total que constitui a vacina. A análise que determina indiretamente o teor de proteínas, além de ser um requisito para a liberação dos lotes, permite ao laboratório de controle de qualidade avaliar a consistência do processo de produção, através, por exemplo, do acompanhamento dos valores, utilizando gráficos de controle (NETTO et al, 2011). Tanto o conteúdo de ovoalbumina residual, quanto os antibióticos adicionados à vacina, devem ser determinados e devem estar dentro dos limites aprovados pelo

órgão regulador, neste caso, a Agência Nacional de Vigilância Sanitária, ANVISA (WHO, 2013).

1.2 VIGILÂNCIA SANITÁRIA E CONTROLE DE QUALIDADE

O Sistema Único de Saúde (SUS) foi fundamentado pela Lei nº8.080/90, que o define como um conjunto de serviços de saúde, concedidos por órgãos e instituições federais, estaduais e municipais (CARVALHO; MACHADO; TIMM, 2004). Construindo estes pilares, a Vigilância Sanitária está integrada ao SUS, compreendendo um conjunto de práticas que em suma regulamenta, monitora, fiscaliza e controla os processos e os produtos, com o propósito principal de prevenir de qualquer risco à saúde pública, dentro das premissas executiva e jurídica (CARVALHO; MACHADO; TIMM, 2004; BRASIL, 2010).

Neste contexto, vinculada ao Ministério da Saúde, foi criada pela Lei nº 9.782, de 26 de janeiro 1999, a Agência Nacional de Vigilância Sanitária (ANVISA), atribuindo-a como uma autarquia sob regime especial, que tem como principal função, atuar em todos os setores relacionados a produtos e serviços que possam afetar a saúde da população brasileira, através do controle destes produtos, sua comercialização, prestações de serviço, dentro da estrutura da administração pública federal (CARVALHO; MACHADO; TIMM, 2004; BRASIL, 2010).

A ANVISA, no uso das suas atribuições, estabelece na Resolução da Diretoria Colegiada nº 17, de 16 de abril de 2010 (RDC 17/2010), os requisitos mínimos das definições de Boas Práticas de Fabricação (BPF) de Medicamentos a serem cumpridos e verificados durante as auditorias sanitárias. A BPF está inserida na “Garantia da Qualidade”, que deve assegurar que os produtos são produzidos e controlados, com qualidade apropriada para sua principal finalidade e assim garantir o seu registro diante a ANVISA (BRASIL, 2010).

Segundo a RDC 17/2010, o controle de qualidade tem como funções, desenvolver, validar e implementar todos os procedimentos e protocolos, sempre de acordo com o monitoramento ambiental, com a principal finalidade da liberação dos lotes em conformidade com as especificações estabelecidas (BRASIL, 2010). Todo procedimento deve estar escrito, fornecendo instruções para a realização de operações na forma de Procedimento Operacional Padrão (POP) e conter um protocolo (ou plano) de validação, descrevendo as atividades a serem realizadas na

validação de um projeto específico, com os critérios de aceitação para a aprovação de um método analítico de uso na rotina (BRASIL, 2010).

Sob condições de BPF, a vacina contra Febre Amarela produzida em Bio-Manguinhos, a partir de ovos embrionados, é certificada pela ANVISA e pela Organização Mundial de Saúde (OMS) para suprir a demanda do Ministério da Saúde e das Agências das Nações Unidas.

A OMS recomenda que os testes de CQ, devem ser realizados na matéria-prima e nas diferentes etapas da produção, até o seu produto final, para garantir a segurança química e microbiológica da vacina, bem como estudos de estabilidade do produto (FERREIRA, 2007).

1.2.1 Estudo de estabilidade

A estabilidade das vacinas tem um grande impacto na eficiência dos programas de imunização. O produto deve manter suas propriedades químicas, físicas, biológicas e microbiológicas dentro das especificações, durante seu prazo de validade. Sendo assim, o estudo de estabilidade, é uma avaliação de garantia da qualidade e eficácia dos produtos (REISDORFER, 2011).

A RDC n° 17/2010 (BRASIL, 2010) e o guia para estabilidade, proposto pela *International Conference Harmonization* (ICH) (2003), estabelecem que, assim como o produto final, os produtos intermediários devem ser identificados pelo fabricante e envolvidos nas avaliações de estabilidade (ICH, 2003; BRASIL, 2010).

A ICH apresenta uma coleção de guias sobre estudos de estabilidade, com destaque o guia Q5C, específico para produtos biológicos, descrevendo as condições sob as quais os estudos de estabilidade devem ser realizados, considerando as características dos produtos biológicos, além de análise de dados estatísticos para estudos de estabilidade, publicado no guia Q1E (ICH, 1995;1996a; 1996b; 2003a; 2003b; 2003c).

Dessa forma, assim como, para o controle de qualidade do produto final da vacina, o teste de nitrogênio proteico também é requisito para os estudos de estabilidade. A determinação é feita no laboratório de Biomanguinhos, atualmente, através do método clássico de Kjeldahl.

1.3 MÉTODO CLÁSSICO KJELDAHL

Proposto por Johann Kjeldahl, em 1883 (BOEIRA; MAXIMILIANO, 2004), o método original tradicionalmente utilizado em laboratórios do mundo inteiro, para determinação de nitrogênio total, sofreu poucas alterações ao longo do tempo (CQAQUÍMICA, 2014). Esta técnica possibilita a determinação indireta de proteínas em vários tipos de amostras (GALVANI; GAERTNER, 2006). Trata-se de um processo farmacopeico, sugerido pela *United States Pharmacopoeia* (USP) para determinação de proteínas em amostras biológicas.

O método é baseado na decomposição da matéria orgânica através da digestão da amostra, em torno de 400°C com ácido sulfúrico concentrado, em presença de sulfato de cobre como catalisador, que acelera a oxidação da matéria orgânica. O nitrogênio presente na solução ácida resultante é determinado por destilação com arraste de vapor, seguida de titulação com ácido clorídrico diluído (GALVANI; GAERTNER, 2006; NARDO et al, 2011).

Este método é exato e preciso, porém apresenta desvantagens, que incluem o uso de reagentes químicos em grande quantidade, resultando na geração de resíduos, que devem ser direcionados adequadamente ao tratamento, de acordo com normas ambientais (MORGANO et al, 2005; NARDO et al, 2011). Perante aos agravos ocasionados pela metodologia atual, a espectroscopia do infravermelho próximo (NIR), associada à calibração multivariada, é uma técnica alternativa, sem a geração de resíduos e com tempo de análise relativamente reduzido (MORGANO et al, 2005; NARDO et al, 2011).

1.4 ESPECTROSCOPIA DO INFRAVERMELHO PRÓXIMO (NIR)

Desde a década de 80, a espectroscopia NIR vem se destacando e sendo utilizada como um método rápido, não destrutivo, em uma aplicação quase universal, com economia de reagentes e de tempo (LIMA et al, 2008). Por envolver moléculas com ligações CH, NH, SH ou OH (PASQUINI, 2003; REICH, 2005), apresenta aplicação em diversas áreas de produção e controle de qualidade.

É uma técnica muito utilizada para a determinação do teor de proteínas (BADDINI et al, 2010), de umidade, de carboidratos, de gorduras (LIMA, 2008; MORGANO et al, 2005); no setor agrícola e meio ambiente (GREDILLA et al, 2016)

especificamente em solos, plantas, madeira, sementes (BRAMBILLA et al, 2006; NARDO et al, 2011); no setor de alimentos, combustíveis, produtos petroquímicos (SILVA et al, 2015), polímeros (SANTOS et al, 2010), cosméticos, tintas; nas indústrias farmacêutica e têxtil (FIDENCIO, 2001; VIANA, 2008; SANTOS, 2011; DE SOUZA et al, 2013).

Pasquini (2003) relata em seu estudo, que muitas foram as tentativas, sem sucesso dos espectroscopistas, em selecionar um único comprimento de onda para análise quantitativa e que a lei de Beer fosse obedecida. Essa ideia foi descartada, segundo Ferreira Neto (2012), quando Karl Norris associou modelos para a utilização de mais de um comprimento de onda com a concentração do analito.

A região de infravermelho refere-se ao espectro eletromagnético compreendido entre a região do visível e do microondas, dividido em três regiões entre 700 e 250.000 nm: o infravermelho próximo, o médio e o longínquo (SKOOG et al, 2006; SAUER, 2007; SANTOS, 2011). A região limítrofe, definindo onde termina uma região e começa outra, não é comum a todos os autores do assunto, assim como Pasquini (2003) e Silverstein e Webster (2006) definem a região de NIR entre 750 a 2.500 nm, enquanto a IUPAC (2014) define a região NIR como sendo de 780 a 2.500 nm (SILVERSTEIN; WEBSTER, 2006; SANTOS, 2007).

Diferentemente da espectroscopia no infravermelho médio (MIR), onde apenas as transições de vibração fundamental são observadas, explicando os espectros com picos definidos e interpretação direta, a espectroscopia NIR gera espectros que apresentam sobreposições e bandas de combinações (ROGGO et al, 2007). Devido às regras da mecânica quântica que descrevem as vibrações harmônicas como menos prováveis de ocorrer do que as fundamentais, a absorção na região do NIR é mais fraca que no MIR (SCHÖNBRODT et al, 2006). Conferindo assim, no caso da espectroscopia NIR, espectros com menos informações, sendo necessário recorrer à calibração multivariada (ALMEIDA, 2009).

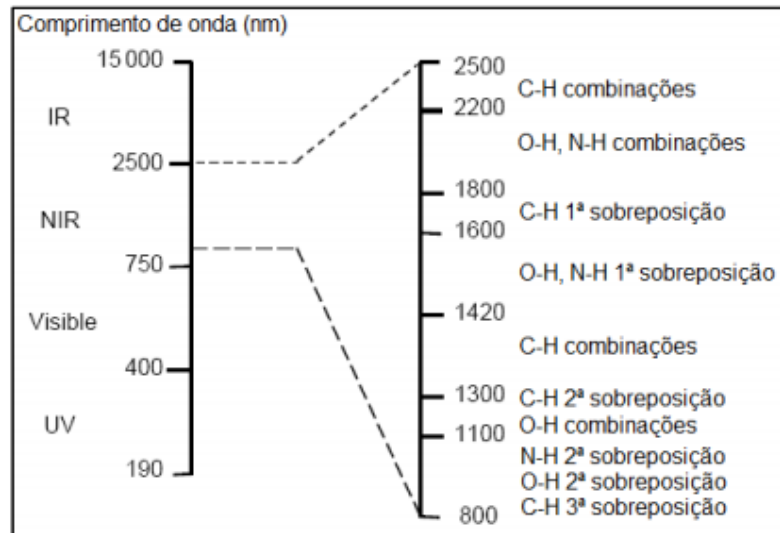
Composições de proteínas, amido, açúcares, lipídios, entre outros, possuem grupos funcionais (-OH, -NH) apresentados na tabela 3. Átomos de hidrogênio ligados covalentemente aos átomos mais pesados e eletronegativos (C, O, N), formam dipolos oscilantes que absorvem energia em regiões de alta frequência, isto é, na região NIR, entre 400 e 2.200 nm, demonstrada na figura 2 (ALMEIDA, 2009; SANTOS, 2011).

Tabela 3. Comprimentos de onda, para absorção de grupos funcionais na região do infravermelho próximo

<i>Grupo Funcional</i>	<i>Comprimento de onda (nm)</i>
C-H alifático	1212, 1414, 1725, 1760 e 2308
C-H aromático	1077, 1111 e 1615 a 1665
O-H água	1440 e 1940
O-H fenóis	1500 a 1800 e 2000 a 2200
N-H aminas	1534, 1000 e 2100
N-H amidas	1980, 2050 e 2180

Fonte: (CREASER; DAVIES, 1988; apud SANTOS, 2011)

Figura 2. Principais bandas de absorção em NIR e sua localização no espectro eletromagnético.



Fonte: (ALMEIDA, 2009).

A radiação eletromagnética pode ser considerada um feixe de fótons, cuja a energia é dada pela equação 1.

$$E = (h\nu) = \left(\frac{hc}{\lambda} \right) \quad (1)$$

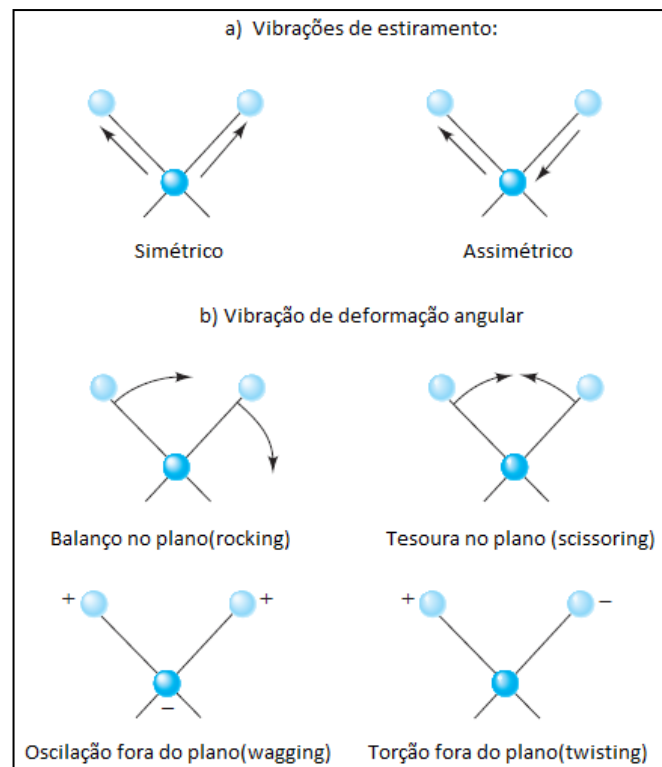
Onde E é a energia da radiação, h a constante de Planck, ν a frequência, c a velocidade da luz e λ o comprimento de onda (SANTOS, 2007).

A radiação infravermelha, incidente em uma molécula, provoca excitação dos modos vibracionais (SAUER, 2007; SANTOS, 2011), quando a energia radiante

corresponde à diferença entre dois níveis energéticos vibracionais (SANTOS, 2007). Caso a molécula apresente alteração no momento dipolar, devido à variação da distribuição eletrônica ao redor das ligações, pode ocorrer transição entre esses níveis vibracionais, sendo assim, detectáveis através do espectro (ALMEIDA, 2009).

As moléculas diatômicas heteronucleares, as moléculas poliatômicas com vibrações assimétricas e grupos funcionais ligados aos polímeros apresentam sinais no espectro. Deve-se ressaltar que as espécies atômicas e moléculas diatômicas homonucleares não são excitáveis (SANTOS, 2007; SANTOS, 2011). As vibrações moleculares são classificadas como: vibração de extensão ou estiramento (*stretching*) e vibração angular (*bending*), figura 3 (SILVERSTEIN; WEBSTER, 2006; SKOOG et al, 2006).

Figura 3. Modos vibracionais das moléculas pertencentes ao grupo AX₂.



Fonte: (SKOOG et al, 2006).

1.4.1 Instrumentação e acessórios

Geralmente um espectrômetro é constituído por uma fonte de luz, um sistema óptico, um detector e um suporte para a amostra. A energia radiante, a partir de

fonte controlada, é direcionada para a amostra e um detector mede a intensidade do feixe emergente. Pasquini (2003), em sua revisão, apresenta os diferentes equipamentos que estão disponíveis no mercado.

A fonte de luz utilizada pode ser lâmpadas de halogêneo, quartzo ou (*Light Emitting Diode*) (LED). Para baixos comprimentos de onda, são utilizados detectores de silício e para comprimentos de onda maiores, são utilizados os detectores de sulfeto de chumbo e índio-gálio-arsênio (ALMEIDA, 2009). Este último é mais sensível e tem maior velocidade de resposta. Associados às fontes de filamentos de tungstênio, geram uma alta razão de sinal-ruído, compensando as baixas intensidades das bandas do NIR (PASQUINI, 2003).

A reflectância difusa com transformada de Fourier, conhecida como DRIFTS (*Diffuse Reflectance Infrared Fourier Transform Spectrometry*), tem como princípio o interferômetro de Michelson, constituído por dois espelhos (um fixo e outro móvel) e por um divisor de feixe (*beam splitter*), geralmente composto de um cristal de brometo de potássio. Entretanto, a apresentação de um interferograma não é muito adequada, assim, esta informação é convertida num espectro, relacionando as intensidades de energia com as respectivas frequências (número de onda), através da transformada de Fourier (NASCIMENTO, 2011).

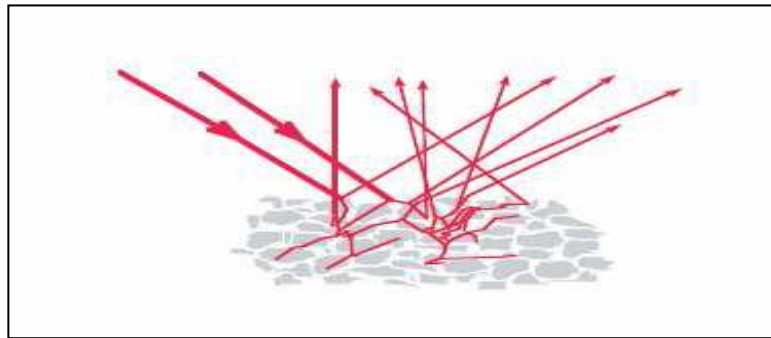
A reflectância (R) é um mecanismo ideal para amostras que são difíceis de analisar por métodos de transmitância, seja por ser pouco absorvente no comprimento incidente, ou que não podem ser destruídas (FIDENCIO, 2001) ou ainda, quando a variação do comprimento de onda, está entre 1.100 e 2.500 nm (VIANA, 2008).

Os métodos de reflectância podem ser classificados como: reflectância interna, que envolve uma célula de reflectância total atenuada (ATR) em contato com a amostra, e a reflectância externa que envolve um feixe de infravermelhos refletidos diretamente da superfície da amostra (Figura 4). E esta, pode ser refletida sob a forma de reflectância especular ou difusa (SANTOS, 2007). Essa ferramenta é fortemente afetada pelas características físicas da amostra, tais como a granulometria e o grau de homogeneidade, assim como a forma e características superficiais das partículas (SANTOS, 2007).

Na reflexão difusa, também será observada a reflexão especular, que é mais intensa na região em que a amostra apresenta forte absorção e neste caso podem ocorrer distorções no espectro, mas pouco perceptíveis em espectros do

infravermelho próximo. A ocorrência de distorções pelo efeito da reflexão especular, nos espectros obtidos com amostras diluídas com brometo de potássio, é similar aos obtidos com amostras puras, sendo que o processamento quimiométrico, na maioria dos casos, é capaz de modelar realçando as regiões onde não ocorre reflexão especular (SAUER, 2007).

Figura 4. Representação esquemática de reflectância difusa num sólido.



Fonte: (ATR, 2004 apud SANTOS, 2007; PORTOLÉS, 2007)

A reflectância difusa mede a razão entre a intensidade da luz refletida e a intensidade da luz incidente.

Por aplicação da equação de Kubelka-Munk, apresentada por Santos (2007) e Portolés (2007), os valores de reflectância, R , são relacionados com o coeficiente de absorção, K , e com o coeficiente de dispersão da amostra num dado comprimento de onda, S , assim como, com a concentração, C (Equação 2).

$$F(R) = \left(\frac{1-R}{2R} \right) = \left(\frac{K}{S} \right) = \left(\frac{kC}{S} \right) \quad (2)$$

Onde k é o coeficiente de absorvância.

A transmitância e a reflectância relacionam-se com a absorvância (A), sendo normalmente representadas como A vs número de onda (Equação 3).

$$A = -\log\left(\frac{I}{I_0}\right) = -\log T = \left(\frac{1}{R}\right) \quad (3)$$

1.4.2 Quimiometria e a calibração multivariada

Segundo Oliveira (2006) e Santos (2011), a dificuldade em escolher a espectroscopia NIR para análise quantitativa, está em interpretar e processar um imenso conjunto de dados, que podem ser obtidos a partir de um espectro, além da multiplicidade de fatores, que perturbam as medidas espectrais. A quimiometria, então, é a área da química destinada ao tratamento desses dados, permitindo resolver um problema de calibração para a determinação de um analito, quando os modelos clássicos de quantificação, tal como o de regressão linear univariado, são insuficientes para calibrar sistemas complexos, em função da presença de várias espécies juntas (FERREIRA, 1999; OLIVEIRA, 2006). Utilizar as ferramentas estatísticas dos métodos multivariados, que considera a correlação entre muitas variáveis analisadas simultaneamente, proporciona a extração de uma quantidade muito maior de informação comparada à análise univariada (SENA, 1999).

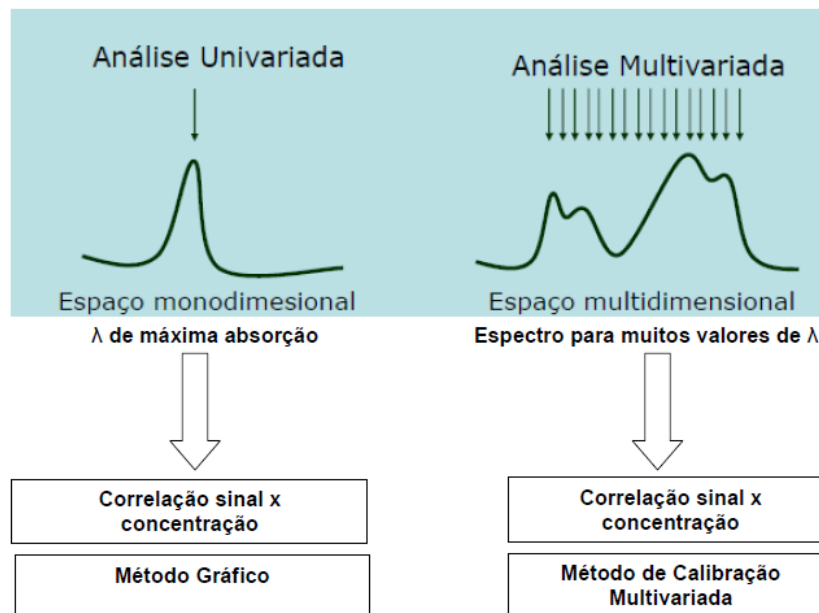
A análise multivariada tem como princípio a utilização das inúmeras variáveis x_1, x_2, \dots, x_n (como valores de absorvância a vários comprimentos de onda), conforme exemplificado na figura 5, para quantificar alguma outra variável de interesse y (como concentração, por exemplo). A quimiometria vem com a principal vantagem de substituir a separação física dos interferentes, pela separação matemática de seus sinais, de forma rápida e nenhuma manipulação da amostra (SANTOS, 2011).

A calibração é o procedimento para construir um modelo matemático a partir de um algoritmo que estabelecerá uma correlação entre a propriedade química, biológica ou física da amostra com seu espectro, seguindo uma extensão da lei de Beer, correlacionando absorvância e concentração, assim como nos modelos univariados (SAUER, 2007).

As técnicas de calibração multivariada mais utilizadas para construção de um modelo de calibração, baseado nos dados de NIR, são as técnicas de calibração lineares, a saber (MONTEIRO, 2011):

- ✓ Mínimos Quadrados Clássico (*Classical Least Squares*, CLS);
- ✓ Regressão Linear Múltipla (*Multiple Linear Regression*, MLR);
- ✓ Mínimos Quadrados Parciais (*Partial Least Squares*, PLS);
- ✓ Regressão por Componentes Principais (*Principle Component Regression*, PCR).

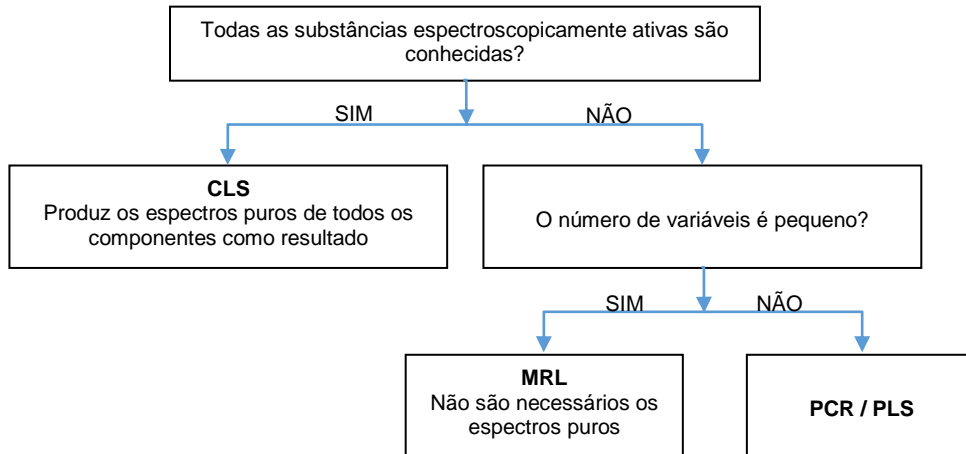
Figura 5. Ilustração gráfica dos procedimentos da análise univariada versus a análise multivariada.



Fonte: (MONTEIRO, 2011).

Cada método possui suas limitações e suas vantagens de aplicação aos problemas químicos. Ferreira e colaboradores (1999) apresentaram um esquema para auxiliar nesse entendimento e direcionar para o melhor uso dos modelos quimiométricos (Figura 6), concluindo que os modelos PCR e PLS são considerados os mais robustos, por não terem seus parâmetros alterados quando se pretende inserir mais amostras no conjunto de calibração. Geralmente, são aplicados quando o usuário não define ou não pode definir um conjunto de comprimentos de onda que devem ser utilizados para a calibração. Enquanto que a técnica de MLR é, geralmente, empregada quando as alterações mais significativas do espectro estão concentradas em regiões espectrais estreitas (SANTOS et al, 2005). O desempenho de um modelo vai depender da estabilidade do equipamento e da qualidade e diversidade dos dados espectrais (SANTOS, 2007). No trabalho de Roggo e colaboradores (2007), são citadas referências para determinação de umidade, de fármacos, de gorduras, em diferentes matrizes, aplicando diferentes modelos de regressão que apresentaram os melhores resultados. Ou seja, a determinação de qual modelo ótimo a ser utilizado na calibração, deve ser investigada em função do contexto analítico.

Figura 6. Fluxograma do método de calibração.

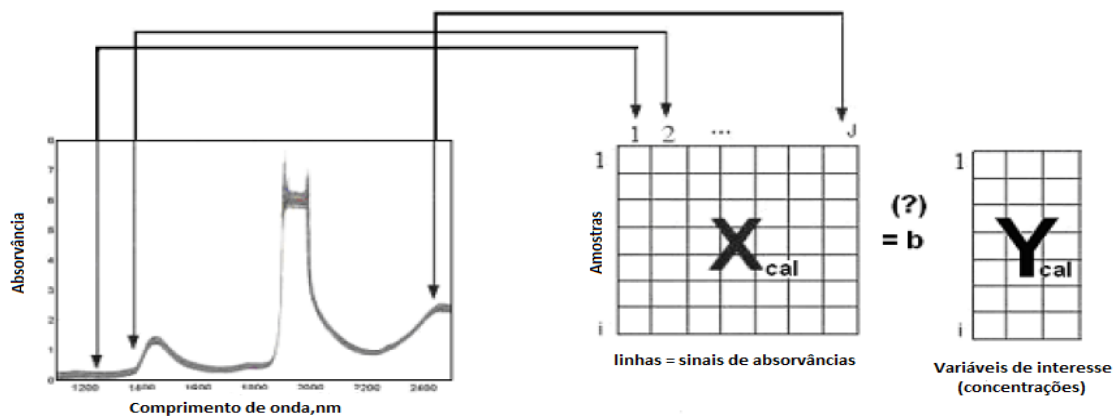


Fonte: (MONTEIRO, 2011, apud FERREIRA et al,1999).

A maioria dos métodos multivariados que são discutidos na literatura, o PLS é talvez o menos entendido. Entretanto, ele é amplamente usado como método de regressão multivariada (TEÓFILO, 2013).

Os dados a serem processados para uma abordagem multivariada em espectroquímica, devem ser organizados sob a forma de uma matriz X , que corresponde ao valor da absorvância no comprimento de onda para cada amostra. Ou seja, o conjunto de variáveis que representam a propriedade de interesse (concentração de referência) será utilizado nas derivações das equações de correlação com os dados da matriz X , estabelecendo uma relação matemática para obter a melhor descrição da variável de interesse (FERREIRA NETO, 2012), conforme representado na figura 7.

Figura 7. Organização de dados para calibração multivariada em espectrometria NIR.



Fonte:(adaptado de FERREIRA NETO, 2012).

1.4.2.1 Pré-processamento dos dados

O pré-processamento de dados é a fase após a construção das matrizes e que antecede os processamentos das análises de PCR, PLS ou outro método de calibração multivariada.

A análise de agrupamentos hierárquico (*cluster analysis*) é a primeira análise multivariada com objetivo classificatório, que interliga as amostras por sua similaridade, a partir de um conjunto de espectros de NIR (PEIQIANG YU, 2005; NONATO et al, 2007), onde as amostras semelhantes, segundo as variáveis escolhidas, são agrupadas entre si. A similaridade entre duas amostras pode ser expressa, como uma função da distância entre os dois pontos representativos destas amostras no espaço n-dimensional. A maneira mais usual de calcular a distância entre dois pontos no espaço n-dimensional, é conhecida por distância euclidiana. A suposição básica de sua interpretação, é quanto menor a distância entre os pontos, maior a semelhança entre as amostras (MOITA NETO; MOITA, 2007). O algoritmo começa com k *clusters* aleatórios movendo depois os objetos entre esses *clusters* com o objetivo de minimizar a variabilidade nos clusters e maximizar a variabilidade entre *clusters* (SANTOS, 2007).

De um modo geral, o pré processamento, consiste na aplicação de métodos corretivos de interferências ocasionadas por fontes instrumentais (ruído instrumental), por condições ambientais (ruído ambiente) ou por fatores inerentes à natureza da amostra que proporcionam variações na aquisição dos espectros, de modo a extrair-se a informação útil dos dados (PORTOLÉS, 2007; SANTOS, 2007; MALUF et al, 2010). Os mais conhecidos e utilizados são: Dados centrados na média, os filtros de Savitzky-Golay, a correção multiplicativa de sinal (MSC), a variação de padrão normal (SNV) e as derivadas.

Geralmente os métodos quimiométricos, apresentados neste texto, assumem que as colunas das matrizes são centradas na média, o que equivale a subtrair a resposta de cada variável sobre a resposta média dessa variável nos dados em todas as amostras (SANTOS, 2007; TEÓFILO, 2013).

O alisamento espectral é um procedimento que utilizando o algoritmo Savitzky-Golay (SG) para reduzir o ruído espectral, vulgarmente conhecido pela sua expressão em Inglês *smooth* (PORTOLÉS, 2007).

A Correção Multiplicativa de Sinal (ou *Multiplicative Scatter Correction*, MSC) usa uma função matemática, para compensar diferenças na trajetória ótica causada pelo tamanho diferente das partículas da amostra, ocasionando dispersão do feixe de radiação. Estas variações ocorrem em maior extensão, nas medições de reflectância. Para aplicar este tipo de correção, a relação entre a contribuição do sinal de absorvância deve ser perfeitamente linear (PORTOLÉS, 2007; SAUER, 2007).

A variação de padrão normal (ou *Standard Normal Variate Method*, SNV) é um pré processamento que realiza um ajuste aditivo e multiplicativo. Para cada espectro da amostra o desvio ajustado é simplesmente a média de todos os valores sobre todas as variáveis e o ajuste multiplicativo é o desvio padrão dos valores de todas as variáveis (SANTOS, 2007).

As Derivadas, como pré tratamento, apresentam a vantagem de remover efeitos de linha de base dos espectros que variam, dependendo das condições de coleta de dados, tal como a temperatura (FERREIRA NETO, 2012).

1.4.2.2 *Análise Multivariada de valores discrepantes*

Os valores discrepantes (outliers) existentes no conjunto de dados, podem afetar os resultados nas análises estatísticas, quando aplicados aos métodos multivariados. Por isso é fundamental que seja feita uma análise exploratória para identificar esses valores. Nos dados multidimensionais, uma observação pode ser considerada discrepante e é possível que não seja, no ponto de vista univariado, pois os métodos verificam os valores extremos em diversas variáveis e não apenas em uma (SARTÓRIO, 2008).

O uso do algoritmo para cálculo da distância de Mahalanobis é sugerido por Oliveira (2010), entretanto Sartório (2008) *apud* Penny (1996) ressalta não ser um método apropriado, para testar pequenos dados amostrais, mas salienta que por meio de métodos como este, o pesquisador consegue selecionar os dados que demonstram verdadeira peculiaridade em comparação ao restante do conjunto amostral.

1.4.2.3. Sinal Analítico Líquido (NAS)

Os modelos de calibração multivariada podem usar sinais que não são seletivos para uma propriedade de interesse (e.g. concentração de certo analito). O sinal analítico líquido (NAS) descreve, justamente a parte do sinal modelado que está quantitativamente relacionada à propriedade a ser prevista (TEÓFILO, 2013). Desse modo, quando se deseja validar utilizando a calibração multivariada, o NAS exerce uma importante função na determinação de figuras de mérito. Para calibração inversa (PLS, PCR, MLR), o NAS é a parte do sinal instrumental que é ortogonal às contribuições de outros possíveis constituintes presentes na amostra (BRAGA; POPPI, 2004).

1.5 VALIDAÇÃO INTRALABORATORIAL DE MÉTODO

Após o desenvolvimento de um método analítico, a validação do método, muito bem definida e documentada, deve ser adotada para evidenciar, que nessas condições, o método tem a capacidade de gerar os resultados esperados com credibilidade, precisão e exatidão, visando diminuir fatores que levam à variabilidade da amostra. Ou seja, determinam os parâmetros de desempenho e limitações do método, indicando os fatores que podem modificá-lo e em qual proporção isso acontece (SWARTZ; KRULL, 1998; LANÇAS, 2004; VALENTINI; SOMMER; MATIOLI, 2007; EURACHEM, 2014).

Em termos gerais, validar significa confirmar através de demonstrações estatísticas, que o seu propósito foi atendido, devendo assim, qualquer desenvolvimento de um método estar devidamente completo, quando se termina a demonstração de performance analítica aceitável. Apesar de alguns autores considerarem, esta última, uma pré-validação (condição intralaboratorial, *in house*), e a validação propriamente dita, seria àquela nos moldes interlaboratoriais (SOUZA, 2007). No ponto de vista sanitário, a validação do método de análise de um produto biológico, como a vacina FA, está diretamente relacionada à segurança da saúde pública.

A RDC nº17/2010 (BRASIL, 2010) preconiza que os métodos analíticos devem ser validados, de acordo com os protocolos escritos, através do qual a

validação é uma parte essencial de Boas Práticas de Fabricação (BPF). Sendo assim, um elemento da garantia da qualidade associado a um produto, padronizando a verificação do cumprimento da BPF, durante as inspeções sanitárias dos laboratórios farmacêuticos (BRASIL, 2010). A Associação Brasileira de Normas Técnicas (ABNT, 2005), também define que os procedimentos dos métodos desenvolvidos pelo laboratório, a ser acreditado, devem estar assegurados para sua finalidade (ABNT, 2005).

O procedimento para a validação de um método pode variar de acordo com o objetivo da análise. Existem diferentes sistemas sugeridos para validação, não há um procedimento único, como por exemplo, a *United States-Environmental Protection Agency* (US-EPA), difere em vários aspectos daqueles indicados pela *United States-Food and Drug Administration* (US-FDA), assim como os critérios da ANVISA, diferenciam-se dos adotados pelo Ministério da Agricultura, Pecuária e do Abastecimento (MAPA) e do documento descrito pelo Instituto Nacional de Metrologia (INMETRO) em 2011. Cada país adota seu próprio programa de qualidade em função da área e sempre em concordância com *International Standardization Organization / International Electrotechnical Commission* (ISO/IEC) (LANÇAS, 2004). A FDA adota as especificações na edição vigente da USP (SWARTZ; KRULL 1998).

A ISO/IEC, *International Union Pure and Applied Chemistry* (IUPAC) e *Internacional association of analytical communities* (AOAC) *international* apresentaram um guia harmonizado para validação intralaboratorial, publicado por Thompson, Ellison e Wood (2002). Semelhantemente, a EURACHEM (2014) disponibiliza um guia que aborda assuntos relacionados com a validação intralaboratorial de métodos quantitativos e qualitativos.

A resolução n° 899 da ANVISA também define um guia para validação de métodos analíticos e bioanalíticos, com base nas premissas da ICH (1996). Entretanto, a diretoria colegiada da agência ANVISA disponibilizou em consulta pública n° 129, de 12 de fevereiro de 2016, uma proposta de validação de métodos analíticos, com exigências mais estatísticas (BRASIL, 2016).

Segundo a resolução n°899 (BRASIL, 2003), qualquer método analítico descrito em farmacopeias ou documentos oficiais, reconhecidos pela ANVISA, é considerado validado. Caso contrário, somente será considerado validado, quando

os parâmetros forem verificados e aceitos, conforme as especificações (Tabelas 4 e 5) referentes a cada categoria desta resolução (BRASIL, 2003).

Tabela 4. Classificação dos testes, segundo sua finalidade.

Categoria	Finalidade do teste
I	Testes quantitativos para a determinação do princípio ativo em produtos farmacêuticos ou matérias-primas
II	Testes quantitativos ou ensaio limite para a determinação de impurezas e produtos de degradação em produtos farmacêuticos e matérias-primas
III	Testes de performance (por exemplo: dissolução, liberação do ativo)
IV	Testes de identificação

Fonte: (BRASIL, 2003).

Tabela 5. Ensaio necessários para a validação do método analítico, segundo sua finalidade.

Parâmetro	Categoria I	Categoria II Quantitativo	Categoria II Ensaio limite	Categoria III	Categoria IV
Especificidade/Seletividade	Sim	Sim	Sim	*	Sim
Linearidade	Sim	Sim	Não	*	Não
Faixa de trabalho(intervalo)	Sim	Sim	*	*	Não
Precisão-repetibilidade	Sim	Sim	Não	Sim	Não
Precisão-intermediária	**	**	Não	**	Não
Limite de detecção	Não	Não	Sim	*	Não
Limite de quantificação	Não	Sim	Não	*	Não
Exatidão	Sim	Sim	*	*	Não
Robustez	Sim	Sim	Sim	Não	Não

* pode ser necessário, dependendo da natureza do teste específico.

** se houver comprovação da reprodutibilidade não é necessária a comprovação da Precisão Intermediária

Fonte: (BRASIL, 2003).

Dessa forma, os parâmetros geralmente a serem verificados durante uma validação são: linearidade e faixa de trabalho, precisão (repetibilidade e precisão intermediária), efeito de matriz, robustez e limites de quantificação e detecção. Esses parâmetros, de um modo geral, são de senso comum a vários autores e instituições de validação (LANÇAS, 2004; VALENTINI, 2007; INMETRO, 2011). Para efeito de uma validação intralaboratorial, a resolução n°899, não exige a apresentação da avaliação da reprodutibilidade (BRASIL, 2003).

Entretanto, todas essas figuras de mérito são desenvolvidas para uma avaliação em métodos analíticos com calibrações univariadas, principalmente na avaliação da linearidade, cujas premissas estatísticas são desenvolvidas para o

método dos mínimos quadrados (univariado). Poucos são os trabalhos que discutem uma avaliação para métodos envolvendo calibrações multivariadas, e sua aceitação ainda é restrita devido às exigências regulamentadoras. A validação de métodos baseados em espectroscopia no infravermelho próximo, é orientada pela norma E1655-00 da *American Society for Testing and Materials* (ASTM), segundo Valderrama (2005). A IUPAC apresentou um documento elaborado por Olivieri e colaboradores (2006), com uma proposta para estimação da incerteza e figuras de mérito em calibrações multivariadas. Os trabalhos de Teófilo (2013) e de Sena e seus colaboradores (2007), descrevem os cálculos para parâmetros de sensibilidade, seletividade e limites de detecção (LOD), seguindo os últimos avanços para a avaliação de incertezas em calibração multivariada, e usando o conceito de sinal analítico líquido (NAS).

A exatidão, bem difundida e como de praxe, é apresentada na maioria dos estudos analíticos que utilizam calibração multivariada, através da validação externa ao processamento dos dados. Braga e Poppi (2004) declaram que a avaliação da linearidade não é uma tarefa simples em calibração multivariada e não demonstram propostas para essa figura de mérito, uma vez que não há curvas de calibração para modelos multivariados (BOTELHO, 2014), e todo o desenho para a linearidade está relacionado aos resíduos provenientes da curva de calibração. O estudo de Almeida (2009), de uma forma qualitativa, utiliza a curva (não a de calibração) resultante da correlação, entre os valores de referência e os valores previstos pelo modelo multivariado, para verificar o comportamento linear.

1.5.1 Faixa de trabalho

Faixa de concentração mínima e máxima do analito com comprovada linearidade, precisão e exatidão (THOMPSON, ELLISON E WOOD, 2002; INMETRO, 2011). A extremidade inferior da faixa de trabalho é limitada pelo limite de quantificação, LOQ (EURACHEM, 2014).

De forma geral, o VIM (2012) define:

...conjunto de valores de grandezas da mesma natureza que pode ser medido por um dado instrumento de medição ou sistema de medição

com incerteza de medição instrumental especificada, sob condições determinadas (VIM, 2012).

1.5.2 Curva analítica e linearidade

A linearidade de um procedimento analítico representa sua capacidade de gerar resultados diretamente proporcionais à concentração do analito na faixa de trabalho definida (AOAC, 1998b; BRITAIN, 1998; INMETRO, 2011; EURACHEM, 2014). A curva analítica é definida como a representação gráfica do sinal ou da resposta medida, em função da concentração do analito (AOAC, 1998b).

1.5.3 Seletividade

É a confiabilidade nos resultados do analito, mesmo em presença de interferências de outros componentes de comportamento similar (EURACHEM, 2014). Sendo a especificidade considerada, quando há 100% da seletividade (SOUZA, 2007). O guia do INMETRO (2011) diferencia a especificidade como o parâmetro que identifica que o equipamento produz resposta apenas para um analito, enquanto a seletividade produz resposta para vários analitos, porém consegue distingui-los.

1.5.4 Sensibilidade

Parâmetro que demonstra a variação da resposta em função da concentração (INMETRO, 2011). Evidenciada no grau de mudança da resposta do instrumento, quando alterados os níveis de concentração para o analito (EURACHEM, 2014).

1.5.5 Efeito de matriz

Como os interferentes de uma matriz podem afetar a inclinação de uma curva analítica, diferentemente do que seria apenas o analito (EURACHEM, 2014), deve-se avaliar se há diferença significativa entre uma curva de calibração usual com o gráfico de adição de matriz (THOMPSON; ELLISON; WOOD, 2002). Nesta etapa de

um processo de validação, segundo o INMETRO (2011), trata-se também a seletividade (INMETRO, 2011).

1.5.6 Precisão

Dispersão entre os resultados (THOMPSON; ELLISON; WOOD, 2002) medidos independentemente em condições definidas, podendo expressá-las através da repetibilidade, delimitadas nas mesmas condições de ensaio em um curto espaço de tempo (EURACHEM, 2014). A reprodutibilidade e precisão intermediária são definidas com o mesmo método, porém alterando, por exemplo, laboratório para avaliação da reprodutibilidade e o operador, para o caso da precisão intermediária (INMETRO, 2011).

1.5.7 Exatidão

Concordância entre o resultado e o valor adotado como referência (INMETRO, 2011; EURACHEM, 2014).

1.5.8 Robustez

Mede a sensibilidade do método frente a variações pequenas de alguns parâmetros definidos (INMETRO, 2011), ou seja, mede a resistência do método em presença de pequenos desvios (THOMPSON; ELLISON; WOOD, 2002) provocados para verificar os efeitos no desempenho do método, e ou equipamento durante o uso normal (EURACHEM, 2014).

1.5.9 Limites de detecção e quantificação

O limite de detecção é a concentração mínima do analito que produz um sinal 3 a 5 vezes a razão sinal/ruído, declarada com 95% ou 99% de confiança, ser maior que zero (THOMPSON; ELLISON; WOOD, 2002; INMETRO, 2011). Já o Limite de quantificação é a menor concentração que pode ser determinada com precisão e exatidão, sendo cinco, seis ou dez vezes o desvio padrão (INMETRO, 2011).

1.6 ESTIMATIVA DA INCERTEZA DA CALIBRAÇÃO MULTIVARIADA

Embora a ANVISA não exija o cálculo da incerteza dos resultados de um método, durante os procedimentos de validação, o conhecimento da incerteza serve como ferramenta de avaliação de controle de qualidade de laboratórios e o resultado de uma medição estará completo, somente quando acompanhado de sua incerteza. Com o propósito de garantir qualidade e confiabilidade das medições, a ABNT NBR ISO/IEC 17025.2005, orienta que os laboratórios devem estimar a incerteza dos resultados (MAGG; ALBANO; MESQUITA, 2013).

A incerteza é definida como um parâmetro não negativo, associado ao resultado de uma medição que expressa a dispersão dos valores, como por exemplo, um desvio padrão (ou um múltiplo dele), ou a metade de um intervalo correspondente a um nível de confiança estabelecido (INMETRO; ABNT, 2003; EURACHEM, 2012).

Para se estimar a incerteza, devem-se conhecer os pontos críticos do método, cujas fontes de incerteza deverão ser relevantes no cálculo, ou seja, fontes que deverão oferecer impacto no resultado final (BAZILIO, 2012). Por isso a necessidade de se tratar cada fonte separadamente, para se ter a contribuição de cada uma, quando se deseja estimar a incerteza total (EURACHEM, 2012). De forma geral, a ISO GUM (2008) em sua última edição, traduzida por um grupo de pesquisadores do INMETRO, conclui que a incerteza reflete a falta de conhecimento exato do valor da sua medida (INMETRO, 2008).

A estimativa da incerteza na predição quimiométrica é um assunto recente e pouco discutido nos trabalhos analíticos, voltados para calibração multivariada. Uma proposta para a estimação da incerteza, ou também chamada de Erro amostral específico de predição (EAEP), em trabalhos multivariados, é a utilização de equações com base em seu *leverage* (h) e na incerteza dos valores padrões obtidos através do método de referência, foram utilizadas por Botelho (2014) apud Martens e de Vries (1995) e Faber e Kowalski (1996) e por Olivieri(2006).

1.7 JUSTIFICATIVA

Dentro do escopo da vigilância sanitária exposto neste trabalho, a vacina de febre amarela (atenuada) é, ainda, o meio mais eficaz para prevenir e controlar a doença. E segundo registros recentes, apresentados neste trabalho, preocupa sua possível reemergência, refletindo no aumento da solicitação tanto pelo Ministério da Saúde para áreas nacionais, quanto para exportação, determinando a importância de um estoque estratégico deste produto em Bio-Manguinhos.

Fortalecendo tal demanda, estudos de estabilidade inseridos no contexto do controle de qualidade das vacinas produzidas em Bio-Manguinhos, são realizados de acordo com os procedimentos operacionais padronizados (POPs) estabelecidos. Desta forma, destaca-se o foco deste estudo, o teste de nitrogênio proteico, atualmente realizado através do método de Kjeldahl. Este método é exato e preciso, entretanto, apresenta inconveniências relativas aos reagentes e aos resíduos ambientalmente inadequados, que devem ser tratados, resultando em custos significativos para a instituição.

O contexto atual envolve uma política (GALVANI, 2006), com o objetivo de evitar a geração de resíduos na fonte, com a aplicação de tecnologias que utilizem menos matérias-primas e recursos naturais em geral, sendo imprescindível, ações e estratégias para se adequar a esta nova realidade. Sendo assim, alternativamente à metodologia do Kjeldahl, a análise por espectroscopia no infravermelho próximo (NIR, do inglês, *Near-Infrared*), possibilita a determinação de diversos compostos químicos simultaneamente, contornando os problemas pontuados.

Diante da importância atribuída ao teste de nitrogênio proteico nos estudos de estabilidade e no controle do produto final da vacina, com o propósito de cumprimento da legislação e em contratempos ocasionados pela técnica atual de arraste por vapor, é clara a motivação para o desenvolvimento e validação deste trabalho, que apresenta a espectroscopia do infravermelho próximo (NIR) com reflectância difusa, associada à calibração multivariada.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Aplicar a espectroscopia no infravermelho próximo por refletância difusa (DRIFTS-NIR) associada à calibração multivariada, nas análises de nitrogênio proteico como alternativa à metodologia convencional de análise laboratorial, a fim de atender as necessidades do controle e dos estudos de estabilidade, garantindo qualidade e especificidade para a matriz vacinal, conseqüentemente à saúde pública.

2.2 OBJETIVOS ESPECÍFICOS

- ✓ Selecionar um conjunto de amostras para a etapa de calibração, a partir de uma análise de “cluster” dos espectros dessas amostras;
- ✓ Construir um banco de dados de espectros das amostras com valores conhecidos de nitrogênio proteico obtidos pelo método de referência;
- ✓ Realizar a calibração do equipamento; em triplicatas de respostas, por diferentes modelos quimiométricos;
- ✓ Avaliar se o modelo multivariado proposto é confiável e atende às especificações, através das estatísticas e conceitos de validação de método.
- ✓ Verificar o potencial do modelo para a previsão de novas amostras, avaliando outro grupo de amostras, não pertencentes ao banco de dados, mediante a comparação dos resultados obtidos pelo método realizado no laboratório e os resultados previstos pelo NIR;
- ✓ Avaliar a influência dos excipientes constituintes da vacina, na resposta espectral;
- ✓ Avaliar as regiões espectrais de trabalho, verificando o comportamento dos modelos;
- ✓ Avaliar os parâmetros possíveis para validação intralaboratorial do método proposto;
- ✓ Estimar a incerteza dos resultados da previsão da calibração multivariada.

3 MATERIAIS E MÉTODOS

3.1 MATERIAIS

3.1.1 Amostras

O lote de vacina de febre amarela de 10 doses, com total de 36.888 frascos, distribuídos na linha de produção em 212 caixas (com 174 frascos cada), foi amostrado conforme o documento interno de BIOMANGUINHOS (2014b), comumente realizado pelo controle de qualidade, durante o processo de inspeção/revisão dos lotes. As caixas foram numeradas de 1 a 212, das quais foram retiradas 20 caixas selecionadas, num intervalo de 10 caixas, permanecendo a primeira e a última caixa. Foram retirados 25 frascos de cada caixa selecionada, de forma aleatória, totalizando 500 amostras, direcionadas às etapas de calibração e validação. Para fins de outros possíveis estudos, foram amostrados mais 200 frascos. Todos os frascos foram armazenados em câmara fria de 2 a 8°C.

3.1.2 Vidrarias

Para a determinação de nitrogênio proteico pelo método Kjeldahl, foram utilizados: Espátula de aço, balão volumétrico 100 e 1000 mL, bécher 50 e 500 mL, tubos de centrífuga, frasco erlenmeyer de 50 mL, micropipetas de 1000 e 5000 uL e tubos de amostra Buchi de 300 mL.

3.1.3 Reagentes

Para a determinação de nitrogênio proteico através do método primário Kjeldahl, utilizando a aparelhagem de destilação manual da Marconi, foram utilizados: Solução aquosa de ácido bórico (Merck) a 4%, solução aquosa de hidróxido de sódio (Merck) a 40%, solução aquosa de hidróxido de sódio (Merck) a 2%, solução aquosa de ácido clorídrico (Merck) 0,014N, indicador misto preparado a partir de 0,1g de verde de bromocresol (Merck), com 0,02g de vermelho metila (Merck) e 100 mL de etanol (Merck), mistura selênica (Merck), ácido sulfúrico

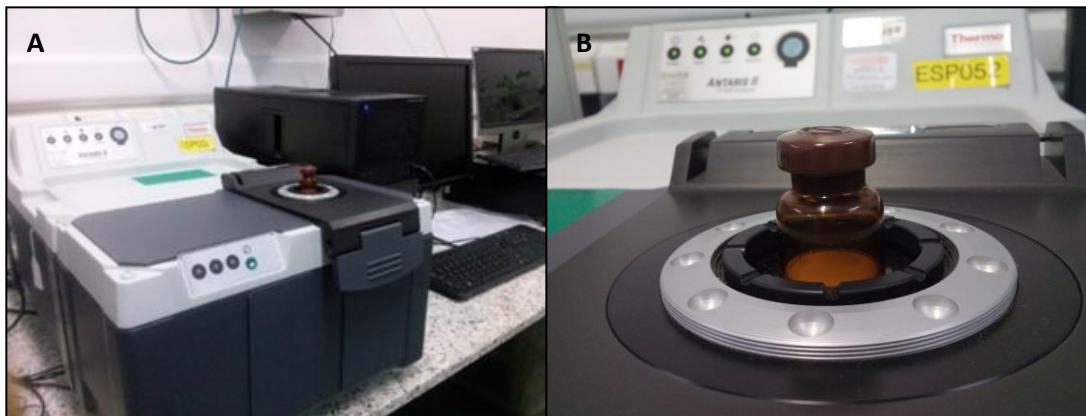
concentrado P.A. (Merck) e ácido tricloroacético (Merck) a 10%. Todas as soluções foram preparadas com água ultrapura (MILLIQ).

Para a preparo da solução padrão de ovoalbumina (Sigma), utilizou-se uma solução aquosa de PBS (pH=7,2), composta por: cloreto de sódio (Merck) a 0,8%, cloreto de potássio (Merck) a 0,02%, hidrogenofosfato dissódico dihidratado (Merck) a 0,14% e dihidrogenofosfato de potássio (Merck) a 0,02%.

3.1.4 Equipamentos

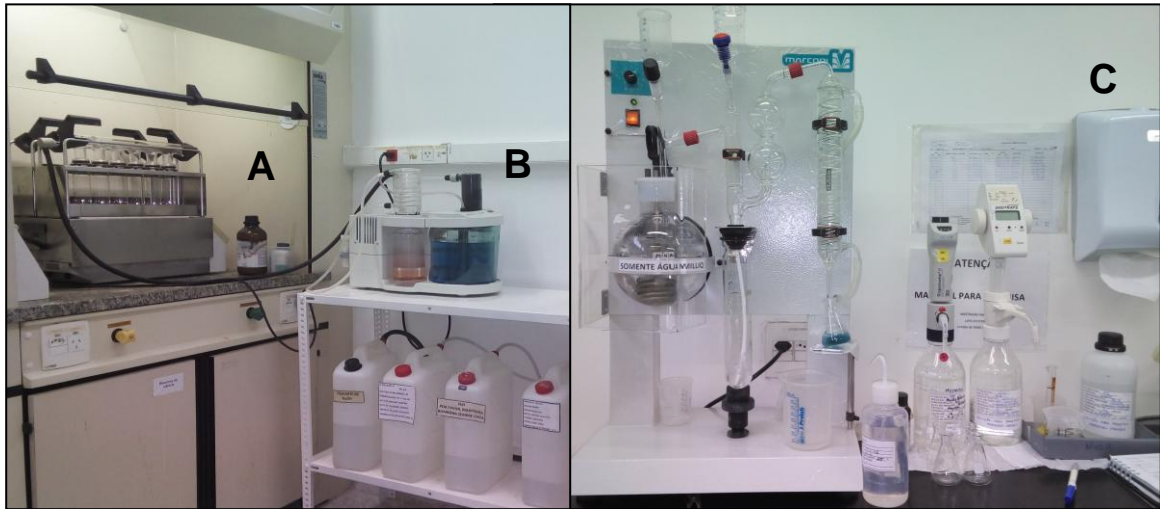
As análises espectroscópicas das amostras foram realizadas em espectrofotômetro da série Antaris II FT-NIR Analyser da Thermo Scientific (Figura 8), equipado com fonte de halogênio, acessório de refletância difusa, com detector de alta sensibilidade e estabilidade índio-gálio-arsênio pareados e interferômetro de Michelson. Com faixa espectral de 12000 a 3800 cm^{-1} (833 a 2630 nm), 32 scans, resolução de 4 cm^{-1} e temperatura ambiente controlada em 22°C .

Figura 8. A) Espectrometro Antaris II FT-NIR Analyser Thermo Scientific, associado ao sistema de microcomputador. B) Compartimento de leitura, com o frasco de VFC.



Para a determinação de nitrogênio proteico pelo método Kjeldahl, foram utilizados: purificador de gases Buchi Scrubber B-414, destilador manual da Marconi, um digestor Buchi Speed digester K-436 (Figura 9), centrífuga Fanem Excelsa II e agitador do tipo vortex.

Figura 9. A) Digestor Buchi Speed digester K-436. B) Purificador de gases Buchi Scrubber B-414. C) Destilador Marconi.



3.2 MÉTODOS

3.2.1 Aquisição dos espectros e Análise de agrupamentos hierárquicos (*Hierarchical Clusters Analysis – HCA*)

A varredura espectral na faixa do NIR foi realizada em triplicata, no primeiro momento, em todos os 700 frascos amostrados, totalizando 2100 espectros. Os 500 espectros médios (reservados para a etapa de calibração) referente a cada frasco selecionado, conforme procedimento estabelecido pelo controle de qualidade de Bio-manguinhos, foram organizados em forma de matriz. Em seguida submetida ao método de classificação por similaridade espectral, *Clusters Analysis* (HCA), utilizando o programa *Unscrambler versão 10.4*, e as amostras foram separadas em 10 grupos. A partir dessas classes, foram selecionadas amostras representativas de cada grupo, para determinação dos respectivos valores de nitrogênio proteico, através do método de referência.

3.2.2 Pré-tratamento dos dados espectrais

Todos os espectros coletados direcionados à modelagem (do conjunto de 500 amostras) foram submetidos aos pré tratamentos: correção multiplicativa de sinal

(MSC), variação de padrão normal (SNV), e as derivadas (primeira e segunda respectivamente) acompanhadas do filtro de alisamento Savitzky-Golay, com ajuste de 5 pontos e polinômio de 2ª ordem, utilizando o programa Unscrambler@ versão 10.4.

3.2.3 Determinação de nitrogênio proteico (Kjeldahl)

O conteúdo de nitrogênio proteico dos respectivos frascos classificados, foi determinado empregando-se o método de Kjeldahl, conforme procedimento descrito em DI 0060, revisão 4.

Para cada frasco selecionado, o líofilo foi previamente reconstituído com água para injetáveis, *water for injections* (WFI), no momento da análise. Para cada frasco, determinou-se o conteúdo de nitrogênio proteico em triplicata. Foram adicionados a um tubo de centrífuga, 1,0 mL da vacina reconstituída, 2,0 mL de água destilada e 0,5 mL de solução de ácido tricloroacético a 10%. Deixou-se em repouso por 1 hora e centrifugou-se por 30 minutos a 2700 rpm. Desprezou-se o sobrenadante e reconstituiu-se o precipitado com 10 mL de hidróxido de sódio 2%, transferindo-o para um tubo de amostra Buchi (300 mL), adicionando em seguida, 2 mL de ácido sulfúrico concentrado e uma pitada de mistura selênica. Efetuou-se o ensaio do branco em triplicata, utilizando 10 mL de hidróxido de sódio 2%, 2 mL de ácido sulfúrico concentrado e uma pitada mistura selênica, em tubo de amostra Buchi.

As amostras e os brancos foram mineralizados durante 30 minutos, em módulo Speed digester k-436. Ao término da digestão e após o esfriamento da amostra mineralizada, destilou-se em equipamento de destilação Marconi, utilizando 10 mL de solução de hidróxido de sódio a 40%, para a liberação da amônia. Recolhendo-se a amônia liberada em um frasco erlenmeyer de 50 mL, contendo 2 mL de ácido bórico a 4% e 2-3 gotas de indicador misto. O recolhimento foi interrompido, quando o volume atingiu cerca de 30 mL e em seguida procedendo-se a titulação, utilizando solução de ácido clorídrico 0,014N, previamente padronizado. O conteúdo de nitrogênio para as vacinas de febre amarela, foi calculado através da equação 4.

$$N_p \left(\frac{\text{mg}}{\text{dose}} \right) = \frac{(V_a - V_b)(n)(f)(0,014)(1000)}{(2V_{am})} \quad (4)$$

Onde, N_p é o nitrogênio proteico em mg/dose, V_a , o volume de ácido gasto na titulação da amostra, V_b , o volume de ácido gasto na titulação do branco, V_{am} , o volume de amostra utilizada, n , a normalidade do ácido utilizado e f , o fator de correção.

3.2.4 Análise de valores discrepantes e discriminação dos valores de referência

Ao conjunto de resultados (convertidos em mg.dose^{-1}) obtidos através do método clássico de Kjeldahl, foi aplicado o teste de Grubbs, conforme orientação do INMETRO (2011), para avaliação de valores discrepantes. Somente o conjunto de dados que será direcionado ao set de validação externa e ao set de calibração foi avaliado quanto a veracidade dos dados de referência.

A fim de avaliar o comportamento das concentrações de referência em consonância aos seus respectivos espectros, foi aplicado o método discriminatório que utiliza o cálculo estatístico da distância de Mahalanobis. Para a análise de discriminação dos valores de referência foi utilizado a ferramenta estatística disponível no pacote Result Software Suite, versão 3 sp7, aplicado ao Antaris II.

3.2.5 Seleção do set de validação e do set de calibração

A separação das amostras para o conjunto de calibração e para o conjunto da validação externa foi obtida a partir da distribuição das mesmas, utilizando uma distância entre as concentrações, com o propósito de formar 10 grupos de concentrações em um histograma. De cada grupo, foram retiradas 50% das amostras para cada set de modelagem.

3.2.6 Teste de normalidade

A normalidade dos dados da regressão foi verificada pelo teste Ryan-Joiner (SOUZA, 2007; INCQS, 2012). Os dados do set de calibração foram primeiramente organizados em ordem crescente. Em seguida, foi construído um gráfico dos valores dos dados ordenados e_i versus os valores dos percentis estimados para uma distribuição normal reduzida q_i (quantis normais).

Os quantis normais foram obtidos pela equação 5, e complementada através das equações 6 a 12 para conclusão do teste.

$$q_i = \varphi^{-1} \left\{ \frac{\left(i - \frac{3}{8} \right)}{\left(n + \frac{1}{4} \right)} \right\}, i = 1, \dots, n \quad (5)$$

$$R_{eq} = \frac{S_{eq}}{\sqrt{(S_{ee})(S_{qq})}} \quad (6)$$

$$S_{eq} = \sum_{i=1}^n (e_i - \bar{e})(q_i - \bar{q}_i) = \sum_{i=1}^n (e_i)(q_i) - \frac{\sum_{i=1}^n e_i \sum_{i=1}^n q_i}{n} \quad (7)$$

$$S_{ee} = \sum_{i=1}^n (e_i - \bar{e})^2 = \sum_{i=1}^n (e_i)^2 - \frac{\left(\sum_{i=1}^n e_i \right)^2}{n} \quad (8)$$

$$S_{qq} = \sum_{i=1}^n (q_i - \bar{q})^2 = \sum_{i=1}^n (q_i)^2 - \frac{\left(\sum_{i=1}^n q_i \right)^2}{n} \quad (9)$$

$$\bar{e} = \frac{\left(\sum_{i=1}^n e_i \right)}{n} \quad (10)$$

$$\bar{q} = \frac{\left(\sum_{i=1}^n q_i \right)}{n} \quad (11)$$

$$R_{crit(n)} \approx 1,0063 - \left(\frac{0,6118}{n} \right) + \left(\frac{1,3505}{n^2} \right) - \left(\frac{0,1288}{\sqrt{n}} \right) \quad (12)$$

Onde, q_i é o valor normal esperado, φ^{-1} é o inverso de uma função de distribuição normal padrão. O coeficiente de correlação R_{eq} deve ser maior que o valor de $R_{crit(n)}$, para um nível de significância α de 0,05 (SOUZA, 2007).

$R_{eq} > R_{crit(n)}$, indica normalidade.

$R_{eq} < R_{crit(n)}$, indica não normalidade.

3.2.7 Calibração e análise quimiométrica

Os modelos multivariados de calibração foram desenvolvidos por uma rotina de regressões, seguida da análise estatística e montagem de matrizes de dados, através do qual se utilizou o pacote *Result Software Suite*, versão 3 sp7, aplicado ao Antaris II.

Os dados espectrais foram utilizados sem pré-processamento e pré processados (tratamento que antecede os processamentos de modelagem multivariada). As ferramentas de pré tratamento espectrais, disponíveis no programa de processamento aplicado ao Antaris II e avaliadas neste trabalho, foram: os dados centrados na média (CM), a correção multiplicativa de sinal (MSC), a variação de padrão normal (SNV), a primeira e segunda derivadas, e os filtros de Savitzky-Golay.

O processamento dos dados consistiu em aplicar os métodos: mínimos quadrados clássicos (CLS), regressão linear múltipla (Multiple Linear Regression, MLR), mínimos quadrados parciais (Partial Least Squares, PLS) e regressão por componentes principais (Principle Component Regression, PCR), em matrizes

contendo os espectros NIR (X) e matrizes contendo os resultados de análises químicas (Y) oriundos da determinação pelo método de referência Kjeldahl.

A exatidão dos modelos, ou seja, o poder de previsão dos modelos foi verificada calculando-se o grau de concordância entre o valor de referência e o valor estimado pelo modelo, valores estes expressos na forma de raiz quadrada do erro quadrático médio de calibração (*RMSEC* - “Root Mean Squares Error of Calibration”), de validação cruzada (*RMSECV* – “Root Mean Squares Error of Leave One Out Cross-Validation” com “n” interações), de previsão (*RMSEP* – “Root Mean Squares Error of Prediction”), conforme as equações a seguir (BADDINI et al, 2010; SANTOS, 2007; VIANNA, 2008).

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (13)$$

Onde, y_i é o valor de referência, \hat{y}_i o valor estimado e o n amostras utilizadas na calibração (Equação 13).

A validação cruzada (validação interna) foi usada para determinação do modelo e do respectivo número de variáveis latentes, conforme a equação 14 (SANTOS, 2007).

$$RMSECV = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}} \quad (14)$$

Onde, n é o número de amostras do conjunto de calibração, y_i o valor medido da amostra i e \hat{y}_i o valor previsto para a amostra i , quando o modelo é construído sem ela (Equação 14).

A fase final do processo de avaliação do poder de predição ou seja desempenho final do modelo de calibração foi verificada a capacidade preditiva dos modelos, através do processo de validação externa (SAUER, 2007), calculando-se o *RMSEP* a partir do grau de concordância entre valor de referência (y_i) e o valor estimado (\hat{y}_i), e o n amostras utilizadas validação externa, ou seja, do conjunto de teste (Equação 15).

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (15)$$

O número apropriado de fatores, ou variáveis latentes (*VL*), na calibração PLS foi determinado utilizando como critério o menor RMSECV e com o menor PRESS (predicted residual error sum of squares), que é a soma dos erros quadrados para os valores estimados em relação aos valores reais, conforme a equação 16 (FERRAO et al, 2004; MORGANO et al, 2005).

$$PRESS = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

Onde, y_i é o valor de referência para a i ésima amostra, e \hat{y}_i é a previsão para o valor desta amostra. Um decréscimo no PRESS indica um modelo “robusto”, ou seja, com habilidade de previsão. O modelo otimizado foi considerado a partir do número mínimo de fatores, com um PRESS de erro padrão mínimo.

3.2.8 Otimização das regiões de trabalho

Através da matriz dos dados (variáveis, valores de referência e seus respectivos sinais de absorvância), foram calculados os coeficientes de correlação entre cada vetor das variáveis, em função dos valores de referências. O gráfico resultante da correlação entre esses coeficientes e os números de onda (cm^{-1}), foi dividido em regiões espectrais limitados pela inversão de sinal dos coeficientes de correlação. Os parâmetros estatísticos RMSEC e RMSEP, foram avaliados para as possíveis combinações dessas regiões.

As regiões sugeridas pela literatura (STUART, 2014) 4545-5000 cm^{-1} , 6250-7143 cm^{-1} e 9091-9999 cm^{-1} , para estiramentos N-H, foram também avaliadas segundo a representação gráfica.

Através do programa *Unscrambler@ versão x10.4*, foram calculados os coeficientes de regressão do modelo PLS.MSC.CM e então plotados em função das regiões espectrais do NIR.

3.2.9 Validação do método de calibração multivariada

A validação do método foi realizada conforme o procedimento operacional padrão - POP nº 65.3120.126 (INCQS, 2012). O procedimento segue como referência as recomendações de Horwitz e Albert (2006), do documento DOQ-CGCRE-008 (INMETRO, 2011), de Souza (2007) e de Thompson, Ellison e Wood (2002). Dentro deste escopo, serão também contempladas algumas figuras de mérito da respectiva categoria II, conforme a resolução nº899, de 29 de maio de 2003, da Agência Nacional de Vigilância Sanitária (ANVISA), para ensaio limite.

No contexto multidimensional, foram exploradas as sugestões para avaliação das figuras de mérito proposto por Valderrama e Poppi (2009), através do qual certificam que o modelo multivariado proposto é confiável e atende às especificações impostas pela indústria e órgãos de fiscalização nacionais e internacionais como *United States Pharmacopoeia (USP)*, *International Conference on Harmonisation's (ICH)*, *American Society for Testing and Materials (ASTM)*, Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO) e a Agência Nacional de Vigilância Sanitária (ANVISA).

3.2.9.1 Avaliação da repetibilidade do modelo de calibração

Conforme a resolução nº899, para a avaliação da repetibilidade, foi verificada a concordância entre os resultados, num intervalo curto de tempo, com o mesmo analista e mesmo equipamento, através de, 10 (dez) determinações, dentro da faixa de trabalho do método, respeitando os 3 (três) níveis de concentrações, baixa, média e alta (BRASIL, 2003).

Como critérios de aceitação da repetibilidade da previsão do método, foi avaliado por meio da razão de Horwitz ($HorRat_{rep}$), calculada para condições de reprodutibilidade (HORWITZ; ALBERT, 2006). Na avaliação da repetibilidade, utilizando-se o coeficiente de variação da repetibilidade, $CV_{rep.exp.}$ (%), muito utilizado pela literatura como desvio padrão relativo, $DPR_{rep.exp.}$ (Equação 17), calculado

através de dados experimentais e o desvio padrão relativo estimado (previsto-teórico) $DPR_{teórico}$ (%), ambos em condições de repetibilidade (INCQS, 2012).

$$DPR_{rep.exp} = \left(\frac{s}{\bar{x}} \right) \times 100 \quad (17)$$

Onde, \bar{x} é a média das concentrações medidas para as repetições, s é o desvio padrão das concentrações medidas para as repetições.

Os valores de $DPR_{rep.teórico}$ deverão ser menores que um meio, partindo de uma premissa menos permissível, ou menores que dois terços do valor do desvio padrão relativo de reprodutibilidade (teórico-previsto) , $DPR_{Repro.Teórico}$, conforme a Equação 18 (HORWITZ; ALBERT, 2006). Sendo este calculado através da equação de Horwitz (Equação 18), utilizada para concentrações maiores que $0,1 \text{ mg.kg}^{-1}$ e menores que $138.000 \text{ mg.kg}^{-1}$ (0,138 fração de massa), faixa em que se encontra a concentração encontrada de $N_{proteico}$ para a vacina de febre amarela ($< 3.500 \text{ mg.kg}^{-1}$). Para o cálculo do $DPR_{rep.teórico}$, neste trabalho foi utilizada a razão máxima de dois terços (Equação 19).

$$DPR_{\substack{\text{reprodutibilidade} \\ \text{Teórico}}} = \left(2(C^{-0,1505}) \right) \quad (18)$$

$$DPR_{\substack{\text{repetibilidade} \\ \text{Teórico}}} = \left(\frac{2}{3} \right) DPR_{\substack{\text{reprodutibilidade} \\ \text{teórico}}} \quad (19)$$

Onde, C é a concentração do analito em fração de massa (considerada como 0,0035). O método é considerado repetitivo com um valor de $HorRat_{rep}$ (Equação 20) menor ou igual a 2.

$$HorRat_{rep} = \left(\frac{DPR_{\substack{\text{repetibilidade} \\ \text{experimental}}}}{DPR_{\substack{\text{repetibilidade} \\ \text{teórico}}}} \right) \quad (20)$$

3.2.9.2 Precisão Intermediária

A resolução nº899 da ANVISA, exige apenas que, para a avaliação da concordância entre os resultados do mesmo laboratório, (precisão intermediária), seja em dias diferentes, com analistas diferentes e/ou equipamentos diferentes, sugerindo um mínimo de dois dias diferentes com analistas diferentes. Dessa forma, o procedimento foi realizado em seis dias diferentes, com três repetições de uma mesma amostra. Foram utilizadas três mostras diferentes, caracterizadas como “ amostras controles “ para todo o estudo.

O teste foi avaliado através da determinação do desvio padrão de precisão intermediária relativo ao tempo Si_{tempo} (Equação 21) e do desvio padrão relativo de precisão intermediária $DPR_{prec.int}$ (Equação 22), conforme orientação do INMETRO(2011).

$$Si_{tempo} = \sqrt{\left(\frac{1}{t(n-1)} \sum_{j=1}^n \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2 \right)} \quad (21)$$

$$DPR_{prec.int} = \frac{Si_{tempo}}{\bar{x}} 100 \quad (22)$$

Onde, t são os dias de análise. n é o número de repetições por dia, x é a concentração medida para a repetição e \bar{x} é a média das concentrações medidas para as repetições.

A razão $HorRat_R$ (Equação 23) foi calculada pela divisão entre o desvio padrão relativo de precisão intermediária, $DPR_{prec.int}$ e o $DPRP_R$. Valores de $HorRat_R$ menores ou iguais a 2, indicarão precisão intermediária adequada (HORWITZ; ALBERT, 2006; INCQS, 2012).

$$HorRat_R = \frac{DPR_{prec.int}}{DPRP_R} \leq 2 \quad (23)$$

3.2.9.3 Exatidão do método de calibração

A resolução nº899 aceita a comparação dos resultados obtidos, com um segundo método analítico validado, neste caso, os valores de referência obtidos através do método Kjeldahl para amostra, comparado com os valores estimados pelo NIR calibrado.

A exatidão foi calculada através da equação 24, onde, $C_{\text{método}}$ é a concentração média, determinada experimentalmente através do método em processo de validação e $C_{\text{referência}}$ é a concentração teórica obtida através do método do kjeldahl, adotada como referência (BRASIL, 2003). Como critério de aceitação, será considerada a recuperação de 95 a 105%, para a concentração deste trabalho na faixa de 0,2 a 0,35%, conforme a tabela 6 (AOAC,2002).

$$\text{Exatidão (\%)} = \frac{(C_{\text{método}})100}{C_{\text{referência}}} \quad (24)$$

Tabela 6. Faixa de recuperação correspondente a concentração limite

CONCENTRAÇÃO	RECUPERAÇÃO (%)
100%	98-102
10%	98-102
1%	97-103
0,1%	95-105
100 mg/kg	90-107
10 mg/kg	80-110
1 mg/kg	80-110
100 µg/kg	80-110
10 µg/kg	60-115
1 µg/kg	40-120

Fonte: (AOAC, 2002).

Como de praxe nos trabalhos analíticos com calibração multivariada (ADAMAPOULOS; GOULA, 2004; QU et al., 2008; FILGUEIRAS et al., 2012; DE SOUZA et al., 2013; ALVES; POPPI, 2016; GUO; NI; KOKOT, 2016; GUÉVÉLOU; ALLEN Jr, 2016; LI; QU, 2016; LI et al., 2016) a exatidão também foi avaliada através da raiz quadrada do erro médio quadrático de previsão (RMSEP, *Root Mean Squares Error of Prediction*), que é parâmetro para a indicação de boa previsão do

modelo, através do qual, foi considerada todas as amostras pertencentes ao set de validação externa.

3.2.9.4 Robustez

A avaliação da robustez foi realizada segundo o método de comparação de médias (teste t). As condições modificadas foram temperatura ambiente, tempo de leitura após retiradas as amostras da refrigeração, a partir de 1 minuto, leituras anteriores à troca de acessório instrumental (lâmpada) e a variação da posição do frasco sobre a lente de leitura, conforme tabela 7. Em cada condição, foi verificada a homogeneidade entre as variâncias da amostra nas condições com a variação e da amostra nas condições normais do método, para um nível de confiança de 95%, através de teste F para duas amostras (Equação 25), onde S^2 é a variância da amostra, ou através da análise de variâncias, ANOVA (INCQS, 2012).

$$F_{cal} = \frac{S_{maior}^2}{S_{menor}^2} \quad (25)$$

Tabela 7. Variações avaliadas na verificação da robustez

Variação	Condição normal	Condições variadas
Temperatura ambiente	23°C	22°, 24°, 25° e 27°C
Tempo de degelo	10 minutos	1min, 5 min, 1 h, 2 h, 4h e 6h
Posição do frasco	Referencial	90°/180°/270°
Troca de acessório (Lâmpada)	Data anterior à troca	Data pós troca

Quando comprovada a homoscedasticidade entre as variâncias da amostra avaliada e da amostra nas condições normais, as médias das concentrações foram comparadas através do teste t para amostras com variâncias homogêneas. O método foi considerado robusto para a variação quando o p – valor calculado para

estatística $t_{calculado}$ do teste (Equação 26) for maior que a significância $\alpha/2$ de 0,025, para um nível de confiança de 95%, com $n - 1$ graus de liberdade.

$$t_{calculado} = \frac{|\bar{d}|}{\sqrt{\frac{s^2}{n}}} \quad (26)$$

$$\bar{d} = \frac{\sum_{i=1}^n (x_{i1} - x_{i2})}{n} \quad (27)$$

Onde, x_{i1} é a concentração da i -ésima replicata da amostra ensaiada nas condições normais do método, x_{i2} é a concentração da i -ésima replicata da amostra ensaiada para a variação, S_2 é a variância das diferenças e n é o número de pares de replicatas comparadas.

Quando comprovada a heterocedasticidade, as médias das concentrações das amostras foram comparadas através do teste t com variâncias amostrais não homogêneas. O método foi considerado robusto para a variação, quando o p -valor calculado para estatística $t_{calculado}$ (Equação 28) for maior que a significância $\alpha/2$ de 0,025, para um nível de confiança de 95%, com ν graus de liberdade, definidos pela equação 29.

$$t'_{calculado} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (28)$$

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 2} + \frac{(s_2^2/n_2)^2}{n_2 - 2}} \quad (29)$$

3.2.9.5 Teste de erros sistemáticos no modelo-Viés (Bias)

O viés, ou *bias*, é uma figura de mérito (FOM) que avalia a presença de erros sistemáticos no modelo. A existência desse tipo de erro afeta a precisão, exatidão e a determinação dos intervalos de confiança. A norma E1655-00 da ASTM, segundo Valderrama (2005) e Botelho(2014), aborda essa investigação. Para este teste, o bias médio para o conjunto de validação foi calculado pela equação 30.

$$bias = \frac{\sum_{i=1}^n (x_i - \hat{x}_i)}{n} \quad (30)$$

Onde x_i são os valores de referência e \hat{x}_i os preditos para a propriedade de interesse, respectivamente, e n é o número de amostras no conjunto de validação.

O desvio padrão dos erros de validação (SDV) também foi estimado (Equação 31) e através de um teste t de Student (Equação 32), para determinar se as estimativas de validação apresentam um viés (bias) estatisticamente significativo. O valor t_{bias} é então comparado com o valor $t_{crítico}$ para n graus de liberdade.

$$SDV = \sqrt{\frac{\sum_{i=1}^n [(x_i - \hat{x}_i) - bias]^2}{n-1}} \quad (31)$$

$$t_{bias} = \frac{|bias|\sqrt{n}}{SDV} \quad (32)$$

Quando o valor de t_{bias} apresenta um resultado maior do que o valor de $t_{crítico}$ para $n-1$ graus de liberdade, onde n é o número de amostras da validação, com 95% de confiança, isso é uma evidência de que erros sistemáticos presentes no modelo multivariado são significativos. No entanto, quando o valor de t_{bias} calculado apresentar valor menor do que o valor crítico, então, o erro sistemático incluído no modelo pode ser considerado insignificante e desprezado.

3.2.9.6 Recuperação

Foi preparada uma solução de padrão de ovoalbumina a $3,0 \text{ mg.mL}^{-1}$, medindo-se a massa em balança analítica calibrada e transferida quantitativamente, para um balão volumétrico calibrado, e este foi avolumado à marca com solução de PBS, previamente preparada. As concentrações dessas soluções, em termos de nitrogênio total e proteico foram obtidas através do método clássico de kjeldahl.

Foram selecionadas 15 amostras e reservados seus respectivos espectros, nomeados originais (ou seja, antes da manipulação da amostras). A partir do conhecimento da concentração da solução de ovoalbumina, foram retiradas alíquotas e transferidas aos liófilos de cada amostra vacinal em estudo. Essas novas amostras foram encaminhadas à liofilização e recravadas após o ciclo de liofilização. Na tabela 8, estão descritos os dados dos preparos das soluções.

As alíquotas foram retiradas com o objetivo de obter um adicional de nitrogênio proteico, em dois níveis possíveis de concentrações.

Tabela 8. Procedimento para avaliação da recuperação de padrão de ovoalbumina nas vacinas liofilizadas.

Solução	Preparo da solução			Análise do <i>KJ</i>		Alíquotas (mL)	Incremento na vacina (mg.dose ⁻¹)
	<i>m</i> (mg)	<i>V_d</i> (mL)	<i>C_f</i> (mg.mL ⁻¹)	<i>N_{total}</i> (mg.mL ⁻¹)	<i>N_{proteico}</i> (mg.mL ⁻¹)		
Ovoalbumina (em PBS)	150	50	3,00	0,40	0,42	0,500	0,02
						0,250	0,01

m, massa; *V_d*, volume de diluente(qsf); *C_f*, concentração final da solução padrão de ovoalbumina; *N_{total}*, concentração de nitrogênio total em mg.mL⁻¹; *N_{proteico}*, concentração de nitrogênio proteico em mg.mL⁻¹; *Kj*, referencial ao método clássico de Kjeldahl.

3.2.9.7 Avaliação da influência espectral dos excipientes da matriz

Foram preparadas soluções dos excipientes da vacina de febre amarela 10 doses, conforme apresentada na tabela 9, medindo-se a massa em balança analítica calibrada e transferida quantitativamente, para um balão volumétrico calibrado, e este foi avolumado à marca com água WFI (*Water for injectables*), água para injetáveis, com exceção do preparado para solução de eritromicina feita em

etanol P.A. As concentrações dessas soluções, em termos de nitrogênio total e proteico foram obtidas através do método clássico de Kjeldahl.

Foram selecionadas 8 amostras (para cada excipiente) e reservados seus respectivos espectros, nomeados originais (ou seja, antes da manipulação das amostras). Conhecendo-se as concentrações das soluções de cada excipiente, foram retiradas alíquotas e transferidas aos líofilos de cada amostra vacinal em estudo. Essas novas amostras foram encaminhadas à reiofilização, recravadas após o ciclo de liofilização e obtidos seus respectivos espectros. Na tabela 9 estão descritos os dados dos preparos das soluções.

As alíquotas foram retiradas com o objetivo de obter um adicional de nitrogênio proteico, em dois níveis possíveis de concentrações.

Tabela 9. Procedimento para avaliação da influência espectral dos componentes da matriz nas vacinas liofilizadas.

Solução	Preparo da solução			Análise do <i>KJ</i>		Alíquotas (mL)	Incremento de N na vacina (mg.dose ⁻¹)
	<i>m</i> (mg)	<i>V_d</i> (mL)	<i>C_f</i> (mg.mL)	<i>N_{total}</i> (mg.mL)	<i>N_{proteico}</i> (mg.mL)		
Gelatina	104,4	50	2,1	0,4	n.e	0,5	0,02
Canamicina	125,0	50	2,5	0,2	n.e	1	0,02
Sorbitol	4000,0	50	80,0	n.e	n.e	1	n.a
Sacarose	400,0	50	8,0	n.e	n.e	1	n.a
Glutamato	200,0	50	4,0	0,18	n.e	1	0,02
Eritromicina (em etanol P.A)	524,2	50	10,5	0,22	n.e	1	0,02

m, massa; *V_d*, volume de diluente(qsf); *C_f*, concentração final da solução padrão de cada excipiente; *N_{total}*, concentração de nitrogênio total em mg.mL⁻¹; *N_{proteico}*, concentração de nitrogênio proteico em mg.mL⁻¹; *K_j*, referencial ao método clássico de Kjeldahl; n.e., não encontrado; n.a., não aplicável; N, nitrogênio.

3.2.10 Outras avaliações

3.2.10.1 Precisão (inter meses)

Três frascos, nomeados de amostras controles, foram reservados à avaliação da precisão dos resultados, durante todos os meses de estudo. A variação da aquisição espectral, foi avaliada através do desvio padrão relativo.

Como critérios de aceitação da precisão da previsão do método, foi avaliado por meio da razão de Horwitz ($HorRat_{rep}$), calculada para condições de reprodutibilidade (HORWITZ; ALBERT, 2006). A razão $HorRat_R$ (Equação 23) foi calculada pela divisão entre o desvio padrão relativo de precisão, $DPR_{precisao}$ e o $DPRP_R$. Valores de $HorRat_R$ menores ou iguais a 2, indicam precisão adequada (HORWITZ; ALBERT, 2006; INCQS, 2012).

3.2.10.2 Avaliação da previsão do modelo com lotes de rotina

Dez lotes, produzidos em 2016, foram analisados e comparados estatisticamente com os respectivos valores de referência. A estatística F do teste foi calculada através da equação 25, para o valor previsto (NIR) de cada lote produzido em 2016, em relação ao respectivo valor fornecido pelo método de referência (KJ). Quando comprovada a homoscedasticidade entre as variâncias da amostra avaliada e da amostra nas condições normais, as médias das concentrações foram comparadas através do teste t para amostras com variâncias homogêneas (Equação 26). Quando comprovada a heterocedasticidade, as médias das concentrações das amostras foram comparadas através do teste t com variâncias amostrais não homogêneas (Equação 28).

3.2.11 Estimativa da incerteza da calibração multivariada

O desenho experimental deste trabalho para o cálculo da incerteza, foi baseado a partir dos guias apresentados por IPAC(2007), ISOGUM/INMETRO (2008) e EURACHEM/CITAC (2012), ordenados a seguir:

- ✓ Foram definidas as fontes de incerteza, obtidas através do método de referência e do método de calibração.
- ✓ Foram definidos os modelos matemáticos utilizados nos cálculos e transformações matemáticas, como a fatoração do ácido titulante e o cálculo da concentração de nitrogênio proteico (com a conversão);

- ✓ Foram definidas as componentes de incerteza (sugeridos pelos erros de cada fonte);
- ✓ Foram calculados os coeficientes de sensibilidade;
- ✓ Foram estimadas as incertezas padrão (somente aquelas componentes de incerteza representadas nas equações);
- ✓ Foi calculada a incerteza combinada.
- ✓ Por fim , calculada a incerteza expandida.

3.2.11.1 A identificação das fontes de incerteza e estimativa da incerteza padrão

Através de uma análise detalhada da expressão matemática utilizada para os cálculos envolvidos e do conhecimento dos pontos críticos do método de referência (Kjeldahl), foram identificadas as fontes de incerteza que pudessem impactar no resultado. Ou seja, (R_f) a fatoraçoão do ácido titulante e o (R_c) cálculo da concentração de nitrogênio proteico.

A partir de n observações repetidas de forma independentes, foi determinada a incerteza padrão (Equações 33 a 35) relacionada a cada fonte de incerteza identificada (R_f e R_c).

$$s = \sqrt{\left[\left(\frac{1}{n-1} \right) \sum_{i=1}^n (x_i - \bar{x})^2 \right]} \quad (33)$$

$$S(X) = \frac{s}{\sqrt{n}} \quad (34)$$

$$u = S(X) \quad (35)$$

Onde, x_i é a grandeza independente identificada, n é o número de repetições, \bar{x} é a média das repetições, s é o desvio padrão dos resultados das fontes identificadas, $S(X)$ é o desvio padrão experimental da média, u é a incerteza

padrão de uma grandeza identificada (INMETRO, 2008). E então dividida pela distribuição de probabilidade.

As grandezas para o cálculo das incertezas padrão de outras fontes identificadas foram obtidas a partir dos certificados de calibração ou do fabricante.

3.2.11.2 Cálculo da incerteza padrão combinada $u_c(y)$.

As incertezas padrão, relacionadas a cada variável do modelo matemático da medição, foram propagadas para estimar a incerteza padrão combinada $u_c(y)$. A incerteza padrão combinada $u_c(y)$ (Equação 36) é a incerteza padrão quando o resultado é obtido por meio dos valores de várias outras grandezas, sendo igual à raiz quadrada positiva de uma soma de termos (REDE METROLOGICA, 2013).

$$u_c(y) = \sqrt{\sum_{i=1}^n c_i^2 u^2(x_i)} = \sqrt{\sum_{i=1}^n u_i^2(y)} \quad (36)$$

Onde, $u_c(y)$ é a incerteza padrão combinada, $u(x_i)$ é a incerteza padrão para cada i -ésima fonte de incerteza x_i e n é o número de fontes de incerteza atribuídas à avaliação. Cada coeficiente de sensibilidade, c_i é o quanto a estimativa de saída y é influenciada por variações da estimativa de entrada x_i (REDE METROLOGICA, 2013).

Os coeficientes de sensibilidade (Equação 37) foram calculadas geometricamente, a partir das equações identificadas críticas no procedimento (Equações 38 e 39):

$$c_i = \frac{\Delta R_i}{\Delta v} \quad (37)$$

$$R_f = f = \frac{m}{(v_{a.f} k)} \quad (38)$$

$$R_c = C_{A(mg/dose)} = \frac{(v_{a.t} - v_{br}) \cdot f \cdot v_{rec}}{(v_{ali})} k \quad (39)$$

Onde, C_i , é o coeficiente de sensibilidade para cada componente detectada, ΔRi é a diferença entre o resultado com a variável e sem a variável, Δv é a diferença entre as incertezas das variáveis (a variável adotada), f é o fator do ácido titulante, m é a massa do padrão primário (g), v_{af} é o volume gasto de ácido na fatoraçoão (mL), R_f é o resultado do fator, $v_{a.t}$ é o volume gasto de ácido na titulaçoão da amostra (mL), v_{br} é o volume do gasto com o branco na titulaçoão da amostra (mL), v_{ali} é o volume da alíquota digerida (mL), v_{rec} é o volume de reconstituiçoão da amostra, C_A é a concentraçoão de nitrogênio proteico na amostra, R_c é o resultado da concentraçoão da amostra e k é a constante de conversão.

3.2.11.3 Estimativa da incerteza expandida (U)

A incerteza expandida foi dada pelo produto entre a incerteza padrão combinada estimada e um fator de abrangência k para um nível de confiança desejável de 95,45% (INMETRO, 2008) conforme a Equação 40.

$$U = ku_c(y) \quad (40)$$

Onde, k é o fator de abrangência.

Através da fórmula de Welch-Stterthwaite (Equação 41), é possível obter o número efetivo de graus de liberdade da incerteza padrão combinada (INMETRO, 2008; HATANAKA, 2015):

$$v_{eff} = \frac{u_c^4(y)}{\sum_{i=1}^n \frac{u_i^4(y)}{v_i}} \quad (41)$$

Onde, $u_c(y)$ é a incerteza padrão combinada, $u_i(y)$ é a incerteza padrão da fonte de entrada, na unidade do mensurando, v_i é o número de graus de liberdade das fontes de entrada.

O número de graus de liberdade é um número inteiro, dessa forma, quando o valor calculado de v não resultou em um número inteiro, considerou-se somente a parte inteira do número. Para uma incerteza padrão do tipo B, considerou-se o

número de graus de liberdade infinito (INMETRO, 2008). Nos casos em que a incerteza $u_c(y)$ é aproximadamente normal ou os graus de liberdade efetivos ν_{eff} , são de tamanho significativo, utiliza-se $k=2$ para um nível de confiança de aproximadamente 95,45% (INMETRO, 2008). Para quase todos os fins, devido à natureza aproximativa do processo de estimativa da incerteza de medição, pode-se adotar $k=2$ (INMETRO 2008; EURACHEM, 2012). Contudo, também foi possível determinar o valor de k , quando o processo de estimativa de incerteza não possuiu graus de liberdade efetivos suficientemente grandes, através da estatística t-student (BAZILIO, 2014).

4 RESULTADOS E DISCUSSÕES

4.1 MODELAGEM MULTIVARIADA

4.1.1 Análise de agrupamento hierárquico (*hierarchical clusters analysis – HCA*)

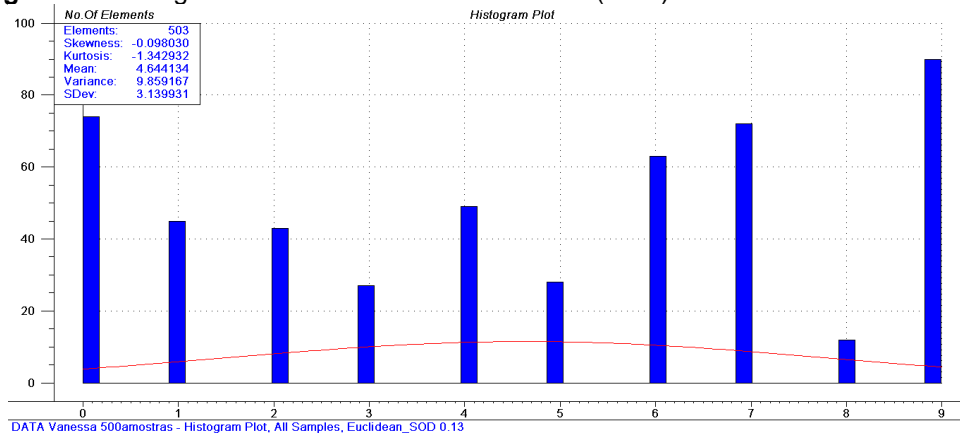
Os dados espectrais organizados em forma de matriz são apresentados na figura 10. E a matriz dos dados, submetida à análise de cluster (HCA), resultou no agrupamento das amostras semelhantes em dez conjuntos, conforme a apresentação do histograma na figura 11.

Figura 10. Matriz dos dados espectrais executada em Excel.

Amostras	3999,64	4001,569	4003,497	4005,426	4007,354	4009,282	4011,211	4013,139	4015,068	4016,996	4018,925	4020,853	4022,782	4024,71	4026,639	4028,567	4030,496	4032,424
6	0,935048	0,934872	0,934532	0,93423	0,933106	0,93306	0,933004	0,932431	0,931862	0,931149	0,930168	0,929709	0,928896	0,927744	0,927352	0,926782	0,925843	0,925158
7	0,888844	0,888707	0,888436	0,888091	0,887053	0,886723	0,886589	0,886075	0,885475	0,884731	0,883732	0,883207	0,882383	0,881095	0,880311	0,879547	0,878362	0,877426
8	0,834172	0,83418	0,833914	0,833555	0,83246	0,832046	0,83205	0,831831	0,831395	0,830866	0,829946	0,829386	0,828677	0,827576	0,827034	0,826396	0,825547	0,824806
9	0,849516	0,849421	0,849198	0,848856	0,847923	0,847748	0,847648	0,847249	0,846662	0,845907	0,84516	0,844681	0,843953	0,842816	0,8424	0,841808	0,840709	0,839956
10	0,967924	0,967676	0,96715	0,96676	0,965521	0,965182	0,96505	0,964584	0,963785	0,962826	0,961721	0,961183	0,960274	0,958893	0,958064	0,957066	0,955671	0,954555
11	0,874294	0,874256	0,873925	0,873601	0,872645	0,872231	0,87221	0,871691	0,871199	0,870601	0,869709	0,869222	0,868855	0,867711	0,866931	0,866286	0,865253	0,864459
12	0,879013	0,878907	0,878699	0,878344	0,87718	0,876989	0,876798	0,875893	0,875037	0,874021	0,873631	0,872933	0,871769	0,870953	0,870175	0,869007	0,868184	0,867406
13	0,891699	0,891745	0,891558	0,891194	0,890085	0,889881	0,889893	0,889346	0,888749	0,888101	0,887288	0,886797	0,885978	0,884678	0,883946	0,883367	0,882382	0,881556
14	0,899162	0,899184	0,898969	0,898657	0,897507	0,897307	0,897225	0,896878	0,896159	0,89529	0,894365	0,893777	0,893144	0,892248	0,89157	0,890853	0,889807	0,889027
15	0,903122	0,903051	0,902886	0,902469	0,901275	0,901126	0,901107	0,900596	0,900007	0,899436	0,898481	0,898	0,897314	0,896139	0,895442	0,894654	0,893637	0,892781
16	0,875018	0,875067	0,874992	0,874585	0,87352	0,87324	0,873133	0,87273	0,872415	0,871759	0,870678	0,870094	0,869502	0,868424	0,867833	0,866705	0,866082	0,865315
17	0,924138	0,924081	0,923907	0,923482	0,922349	0,921877	0,921861	0,921567	0,921208	0,92048	0,919525	0,918871	0,918121	0,916932	0,916223	0,915454	0,914269	0,913342
18	0,819213	0,819141	0,81891	0,818444	0,817418	0,817222	0,81718	0,816737	0,816143	0,815485	0,814698	0,814317	0,813633	0,812582	0,811905	0,811219	0,810143	0,809357
19	0,864225	0,864232	0,864015	0,863554	0,862565	0,862359	0,862371	0,862104	0,861399	0,860508	0,85959	0,859066	0,858142	0,856991	0,856175	0,855287	0,854	0,853056
20	0,838851	0,838839	0,838483	0,838169	0,837214	0,836911	0,836901	0,836609	0,836125	0,835465	0,834462	0,833992	0,833275	0,832329	0,831765	0,831115	0,830031	0,829253
21	0,949517	0,949447	0,949293	0,94889	0,947577	0,947648	0,947518	0,947147	0,946818	0,946314	0,94525	0,944547	0,943756	0,942525	0,941954	0,941179	0,940156	0,939459
22	0,850881	0,850885	0,85065	0,85034	0,849266	0,849133	0,849075	0,848501	0,847862	0,847096	0,846167	0,845701	0,844957	0,84388	0,843372	0,842695	0,841799	0,840915
23	0,904013	0,903986	0,90362	0,903269	0,902343	0,902129	0,901912	0,901459	0,90085	0,900242	0,899283	0,898622	0,897867	0,896766	0,89609	0,895482	0,89441	0,893493
24	0,983058	0,982895	0,982668	0,982456	0,981426	0,980996	0,980909	0,98046	0,979617	0,978735	0,977806	0,976978	0,975989	0,974741	0,973789	0,972917	0,971684	0,970647
25	0,828291	0,828139	0,827915	0,827622	0,826648	0,826449	0,826598	0,8262	0,825577	0,824929	0,824103	0,823757	0,823169	0,82207	0,821474	0,820888	0,819922	0,81912
26	0,908428	0,908303	0,907954	0,907792	0,906748	0,906321	0,906235	0,905878	0,905146	0,904122	0,903039	0,902386	0,901658	0,900371	0,899503	0,898641	0,89738	0,896455
27	0,874256	0,874079	0,873828	0,873547	0,872573	0,872391	0,872421	0,872048	0,871581	0,870955	0,870018	0,86953	0,868834	0,867791	0,867243	0,866466	0,865533	0,864782
28	0,880703	0,880651	0,880468	0,880184	0,878967	0,878716	0,878621	0,877809	0,8771	0,876087	0,875556	0,874692	0,873436	0,872913	0,872241	0,871174	0,870347	0,86947
29	0,880032	0,879846	0,879606	0,879183	0,878018	0,877842	0,877858	0,877443	0,876786	0,875947	0,874947	0,874395	0,873667	0,872408	0,871686	0,870883	0,869742	0,868854
30	0,974579	0,974518	0,974642	0,974355	0,973014	0,972777	0,972735	0,971748	0,970808	0,969636	0,969195	0,968534	0,967333	0,9664	0,965471	0,964019	0,963112	0,96219

A linha 2B apresenta 3112 valores de números de onda (cm^{-1}). Na coluna A estão dispostas as 500 amostras, cujas linhas, apresentam seus sinais espectrais (em absorvância).

Figura 11. Histograma resultante da análise cluster (HCA).



As amostras em negrito organizadas nas tabelas 10 e 11, em seus

respectivos grupos classificados pelo HCA, foram selecionadas para a determinação do seu conteúdo de nitrogênio proteico, através do método kjeldahl.

Tabela 10. Classificação hierárquica (HCA) dos espectros obtidos no NIR, para as amostras de vacinas de febre amarela.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10
11	6	19	10	31	18	7	24	43	13
15	21	22	41	38	27	8	32	166	14
16	30	26	53	45	36	9	40	263	17
28	51	34	71	59	42	12	46	349	23
33	52	44	82	69	47	20	55	386	29
35	77	50	98	97	68	25	56	405	48
37	85	54	103	99	80	63	57	418	58
39	93	67	125	116	94	73	72	431	60
49	104	70	126	123	108	79	74	451	62
61	107	87	186	138	112	84	75	458	64
65	122	109	200	161	128	102	86	460	76
66	134	115	211	170	133	111	88	464	83
78	149	132	232	173	148	114	106		89
81	150	160	233	181	156	118	113		90
91	157	171	276	182	162	124	119		92
101	190	172	284	183	163	129	127		95
110	194	174	286	206	193	131	143		96
117	227	176	306	213	212	140	152		100
120	244	177	319	215	285	147	153		105
121	252	195	326	235	291	158	198		130
135	253	196	331	236	304	164	222		139
136	269	217	378	239	310	167	256		142
137	277	224	419	241	361	168	259		151
141	283	228	449	242	382	175	265		154
144	287	240	479	268	407	178	270		165
145	299	327	485	278	482	184	272		179
146	309	351	496	288	490	188	274		185
155	328	376		293	508	192	281		187
159	347	381		334		204	282		189
169	358	384		348		208	290		191
180	372	396		357		210	295		199
197	390	441		369		214	298		201
203	408	445		370		218	301		202

Com o objetivo de avaliar a frequência de amostras, que apresentavam características semelhantes, o método classificatório multivariado HCA, quando aplicado às matrizes espectrais sem pré-tratamento algum, se mostrou um

planejamento adequada para selecionar as amostras. Essa estratégia fornece uma variabilidade ao set de calibração e de validação, e conseqüentemente, torna o modelo de calibração, mais robusto possível, através do aumento da extensão dos níveis que participarão da curva de calibração.

Tabela 11. Classificação hierárquica dos espectros obtidos no NIR para as amostras de vacinas de febre amarela (continuação da tabela 10).

Cluster 1	Cluster 1	Cluster 2	Cluster 3	Cluster 5	Cluster 7	Cluster 8	Cluster 8	Cluster 10	Cluster 10
207	363	422	454	377	221	302	421	205	356
219	374	434	456	380	223	305	423	209	359
220	394	439	468	392	231	312	424	216	360
237	400	442	469	393	238	315	425	225	368
245	401	452	481	410	246	316	427	226	373
249	402	455	486	416	248	324	436	229	383
250	409	470	494	428	251	325	440	230	389
254	420	473	498	432	255	329	443	234	391
258	429	483	505	437	260	330	444	243	397
262	438	488	509	462	267	336	453	257	399
289	471	499		465	275	339	457	261	404
292	478	504		472	280	340	467	264	413
296	480			476	294	342	474	266	426
300	489			477	314	350	484	271	430
303	497			500	318	365		273	433
307	503			502	320	371		279	435
308	507				322	379		297	446
317					337	395		311	448
332					341	403		313	450
333					354	406		321	459
338					364	411		323	461
346					366	412		335	463
353					367	414		343	475
362					375	415		344	487
					385	417		345	491
					387			352	492
					388			355	493
					466				506

4.1.2 Pré tratamento nos dados espectrais coletados

Os espectros sem tratamento apresentaram deslocamento nos sinais (Figura 12) entre as amostras. Isso foi ocasionado pelo fenômeno de espalhamento da

radiação, associado à técnica de reflectância difusa, sendo necessário o tratamento MSC (Correção Multiplicativa de Sinal), que tem por finalidade, reduzir tais variações espectrais ocasionadas por fatores instrumentais (Figura 13). A correção por SNV (Variação de padrão normal) também foi aplicada (Figura 14), uma vez que, este método está destinado à medir diferentes características de dispersão do sinal espectral, relacionado com o caminho óptico, semelhantemente ao efeito de correção por MSC (PORTOLÉS, 2007).

Figura 12. Espectros médios das amostras sem tratamento

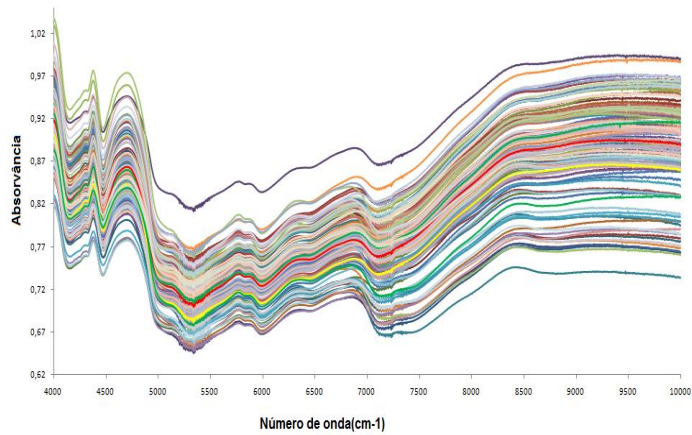


Figura 13. Espectros médios das amostras com tratamento MSC

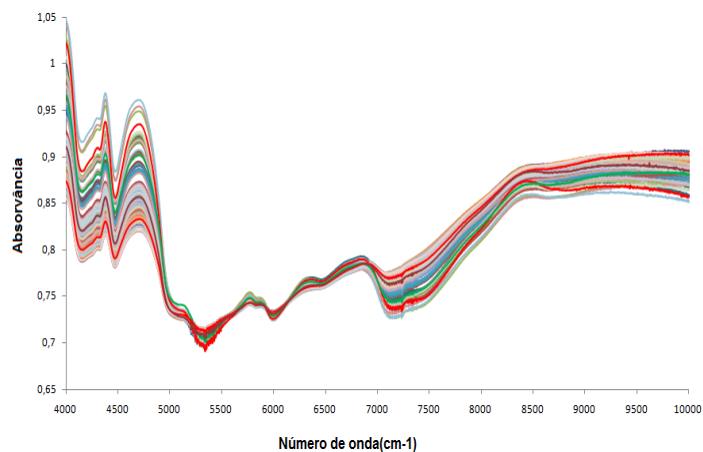
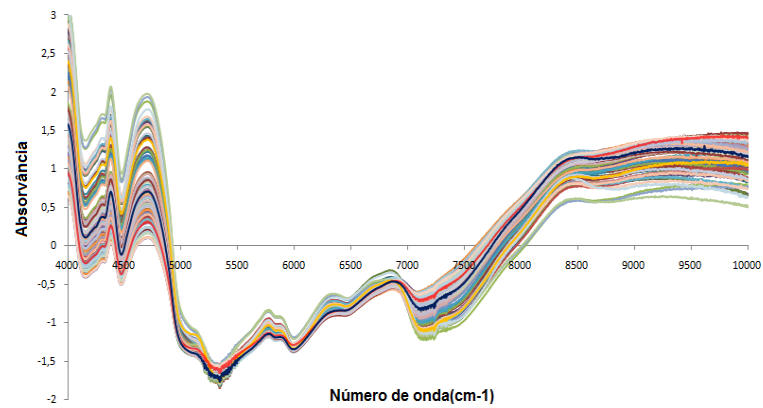


Figura 14. Espectros médios das amostras com tratamento SNV.



As derivadas intensificaram o ruído, sendo essa característica, mais evidente na segunda derivada. Dessa forma, o seu uso acompanhado do alisamento pelo filtro de Savitzky-Golay (SG), com janela de 5 pontos e polinomial de segunda ordem (5p e 2ª Ord.), permitiu reduzir flutuações originadas pelo ruído instrumental, mais aparentes nas regiões 5100 a 5500 cm^{-1} , 7000-7500 cm^{-1} e 9500-9600 cm^{-1} , como foi possível observar nas figuras 15 e 16.

Figura 15. Espectros médios das amostras aplicada a primeira Derivada, com alisamento Savitzky-Golay (5p e 2ª Ord).

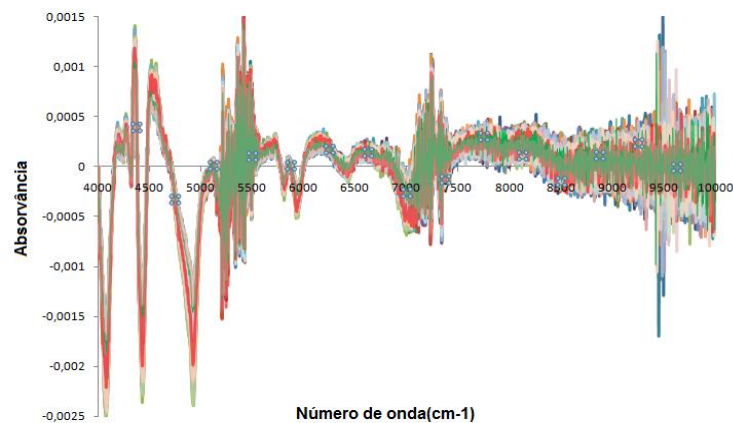
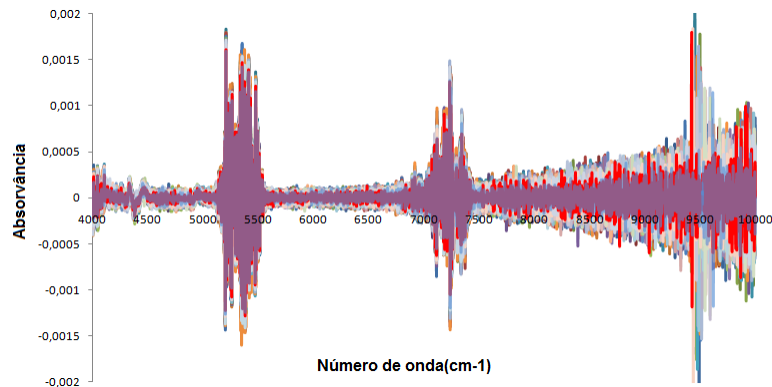


Figura 16. Espectros médios das amostras aplicada a segunda derivada com alisamento Savitsky-Golay(5p e 2ª Ord.).



4.1.3 Método de referência (kjeldahl), o teste de *grubbs* e análise discriminatória.

Os 250 valores médios de referência resultantes do método de Kjeldahl foram tratados em 3 avaliações de valores extremos. Utilizando o teste de *Grubbs*, a tabela 12 apresenta o número final de amostragem para a modelagem.

Tabela 12. Tratamento de valores extremos.

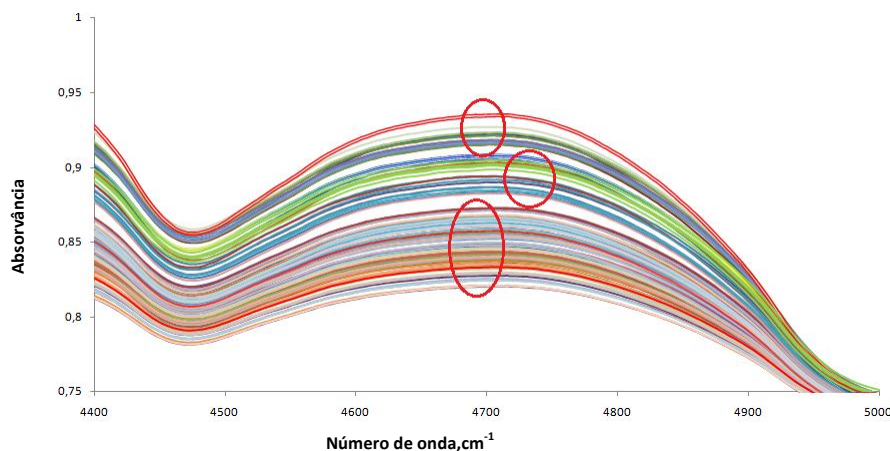
	Avaliação pelo Teste de Grubbs		
	1ª avaliação	2ª avaliação	3ª avaliação
Nº amostras avaliadas	250	246	245
Concentração encontrada (mg/dose)	0,029 a 0,087	0,029 a 0,074	0,029 a 0,068
Média	0,049	0,049	0,049
Coefficiente de Variação	0,17	0,14	0,14
Nº amostras extremas encontradas	4	1	0

O método de referência Kjeldahl envolve várias etapas do procedimento que podem influenciar significativamente nos resultados, esses valores podem ser avaliados pelo teste de Grubbs para valores extremos. No entanto, o teste somente avalia os resultados referentes à cada amostra, diante da homogeneidade do conjunto amostral, mas não avalia se os valores, incluindo os extremos, possuem alguma relação real, com seus espectros gerados pela técnica espectroscópica na região do infravermelho próximo.

Considerando que a etapa de escolha dos valores de referência é a fase mais importante de uma modelagem e a confiança que os mesmos, devem proporcionar à calibração multivariada, foi necessária a aplicação da análise discriminante entre os 250 valores de referência. A partir das informações espectrais de cada amostra e utilizando o cálculo matemático de Mahalanobis, a análise discriminante resulta em valores que não correspondem aos grupos indicados, a fim de reavaliar a veracidade dos valores de referência versus o comportamento espectral desse conjunto.

Os 250 espectros tratados com MSC e CM (dados centrados na média), na região sugerida pela literatura para estiramentos NH em 5000 e 4545 cm^{-1} (STUART, 2004), estão representados na figura 17. As 250 amostras foram separadas em três níveis de valores (alto, médio e baixo) conforme a marcação representada, e submetidas à análise de discriminação.

Figura 17. Espectros médios tratados com MSC e CM na região proposta pela literatura.



Para a região 4545 a 5000 cm^{-1} , sugerida pela literatura (STUART, 2004), foram observados 52 espectros fora da classificação inicialmente definida (valores de nível 3 alto, nível 2 médio e nível 1 baixo), segundo a discriminação. A distância entre os níveis adjacentes, N1 com N2 (Figura 18 A) e N2 com N3 (Figura 18B), não foi tão notória, podendo ser atribuído aos valores muito próximos dos dados, não resultando em grande diferença entre as extremidades das classes. Diferentemente das classes N1 e N3 (Figura 19), que foi possível observar uma distância entre seus grupos.

Figura 18. Análise discriminante na região 4545-5000cm⁻¹, com tratamento MSC e CM. A) Distâncias entre os níveis 2 e 1. B) Distâncias entre os níveis N1 e N3.

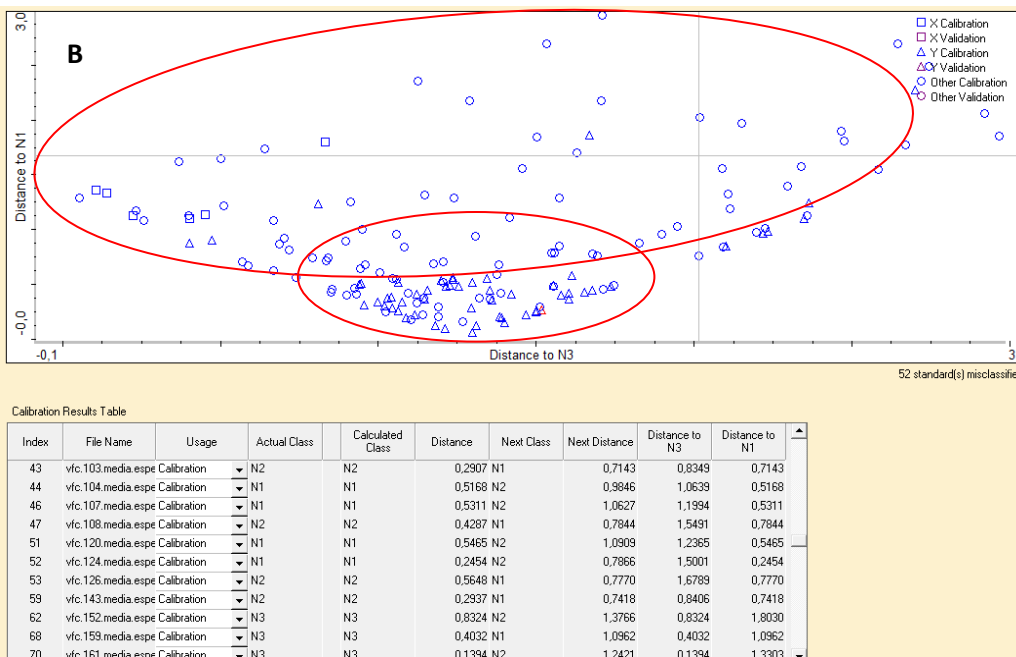
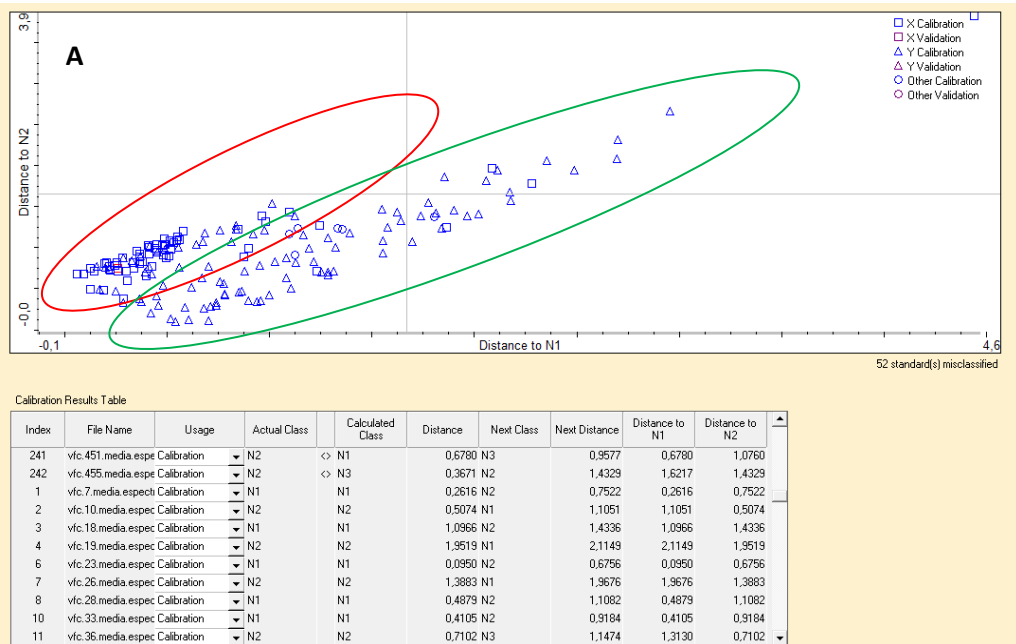
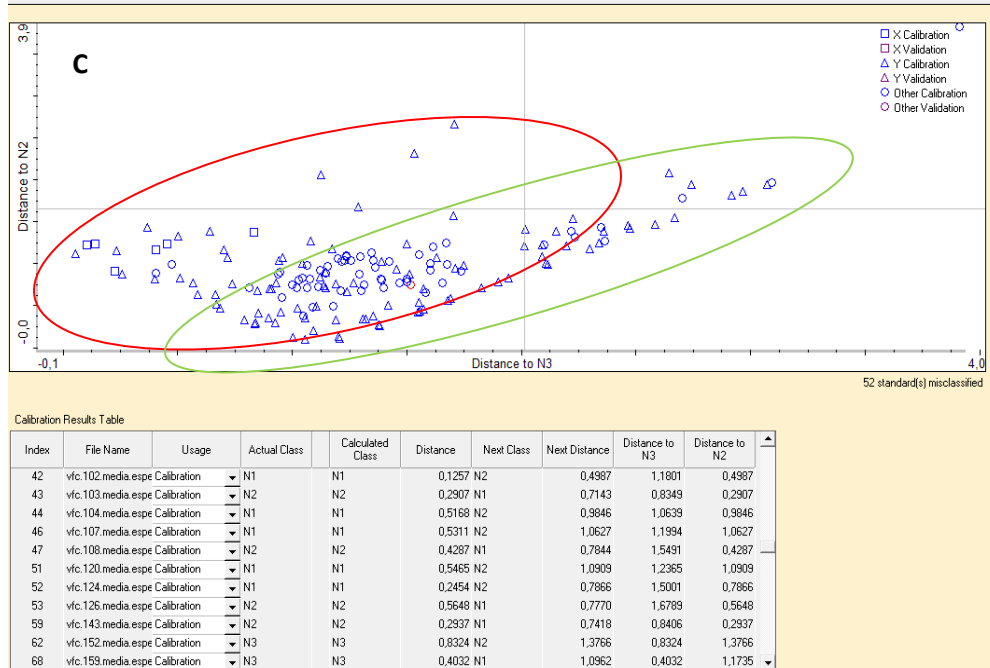


Figura 19. Análise discriminante na região $4545\text{-}5000\text{cm}^{-1}$, com tratamento MSC e CM. C) Distâncias entre os níveis 2 e 3.



A análise também foi aplicada para a região de $6250\text{ e }7145\text{ cm}^{-1}$, sugerida pela literatura (STUART, 2004), com tratamento MSC (Figura 20). Observou-se que a partir do número de onda em torno 6900cm^{-1} , a correlação dos sinais dos espectros é invertida, coerentemente, podendo influenciar na classificação para essa região. Reduzindo a janela dessa região, a fim de retirar a influência da inversão de sinais, houve uma diminuição dos valores desclassificados, resumidos na tabela 13.

Figura 20. Espectros médios tratados com MSC e CM na região $6250\text{ e }7145\text{ cm}^{-1}$, proposta pela literatura.

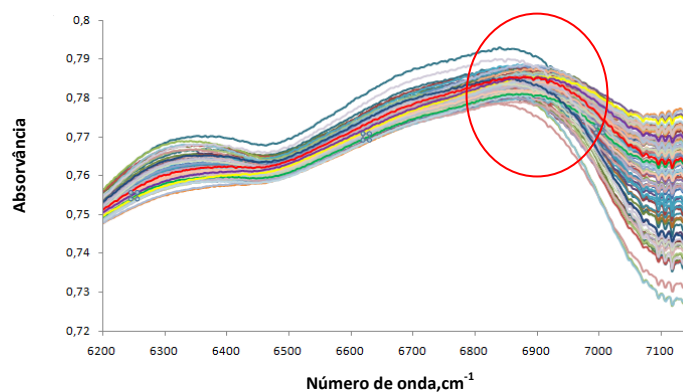
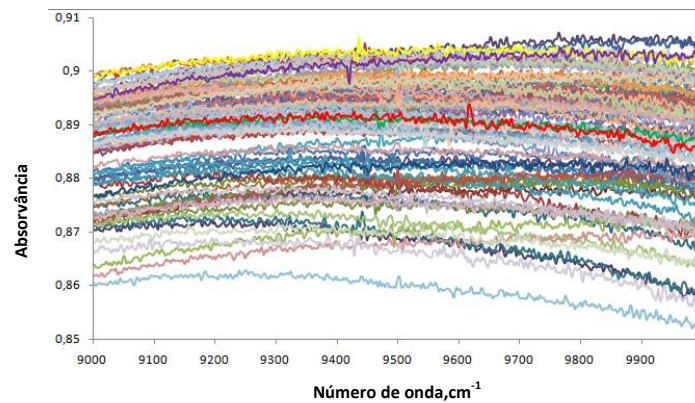


Tabela 13. Análise discriminante, utilizando o cálculo de Mahalanobis.

Tratamento	Região	Número de amostras desclassificadas
MSC/CM	4545-5000 cm^{-1}	52
MSC/CM	6250-7143 cm^{-1}	74
MSC/CM	6250-6800 cm^{-1}	58
MSC/CM	9091-9999 cm^{-1}	76
SNV/CM	4545-5000 cm^{-1}	53
SNV/CM	6250-7143 cm^{-1}	88
SNV/CM	6250-6800 cm^{-1}	57
SNV/CM	9091-9999 cm^{-1}	77

Figura 21. Espectros médios tratados com MSC e CM na região 9091-9999 cm^{-1} proposta pela literatura

A análise foi feita para espectros tratados com o método MSC e SNV, todos centrados na média (CM) e para as três regiões sugeridas, conforme resumo na tabela 13. Nos gráficos espectrais das figuras 17, 20 e 21, somente foi possível observar uma separação dos níveis de classificação dos sinais, no gráfico referente a região 4545 a 5000 cm^{-1} (Figura 17). Tal característica, resultou no menor número de espectros desclassificados, diferentemente das outras regiões estudadas que apresentaram um número maior de amostras que se misturavam em classes diferentes. Isso pode ser explicado pela aproximação dos sinais de absorvância para essas regiões. Dessa forma, 52 amostras foram excluídas do conjunto para modelagem, já que essa avaliação infere dúvida aos valores de referência desclassificados.

Após o teste de Grubbs e a análise de discriminação, o conjunto de resultados do Kjeldahl apresentou 198 valores de referência, para suas respectivas médias espectrais (Tabela 14). A principal vantagem observada, foi evitar a suposição sobre a distribuição dos dados, com exclusões de discrepantes, mesmo baseado em métodos estatísticos.

Tabela 14. Faixa de resultados dos valores de Kjeldahl.

Faixa inicial dos resultados do Kjeldahl	
0,028-0,089 (mg/dose)	n=250
Faixa final dos resultados do Kjeldahl	
0,039-0,062 (mg/dose)	n=198

4.1.4 Seleção do set de calibração e do set de validação.

A separação das amostras para o conjunto de calibração e para o conjunto da validação externa foi obtida a partir da distribuição das 198 amostras, com uma distância de 0,0029 entre as concentrações, formando 8 grupos de concentrações em um histograma (Tabela 15 e Figura 22). De cada grupo, foram retiradas 50% das amostras para cada set de modelagem, resultando em 99 amostras para a calibração e 99 amostras para a validação externa (Tabela 16). Esse procedimento proporcionou uma melhor distribuição de níveis de concentrações presentes em cada set, de forma a avaliar o poder de previsão do modelo na validação externa, em toda a faixa de concentração da curva de calibração.

Tabela 15. Faixas de concentrações (mg/dose) vs frequência.

Faixa (8 grupos)	Frequência
0,039-0,042	12
0,042-0,045	40
0,045-0,048	48
0,048-0,051	29
0,051-0,054	25
0,054-0,056	16
0,056-0,059	19
0,059-0,062	9

Figura 22. Histograma com as faixas de teores de nitrogênio proteico obtidas pelo método de referência Kjeldahl.

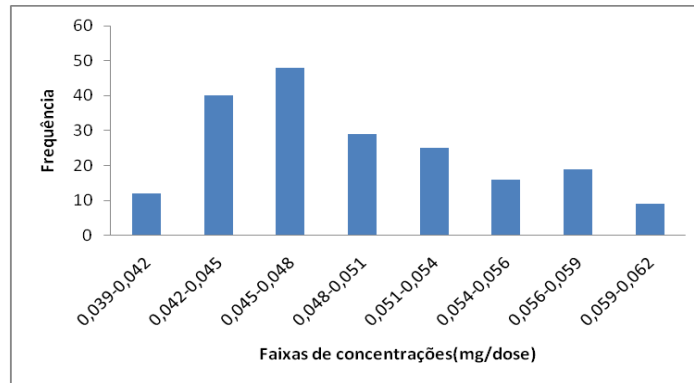


Tabela 16. Faixa de resultados dos valores de selecionados para cada conjunto.

Faixa dos resultados do Kjeldahl selecionados para o set de calibração	
0,039-0,062 (mg/dose)	n=99
Faixa dos resultados do Kjeldahl selecionados para o set de validação externa	
0,039-0,062 (mg/dose)	n=99

4.1.5 Avaliação da normalidade dos dados de calibração

As técnicas de PLS e PCR fazem considerações que impõem limitações ao seu uso, admitindo que as variáveis devem estar relacionadas linearmente e seguem uma distribuição normal (MONTEIRO, 2011). Apesar de Liu et al (2008) considerar em seu estudo, através de monitoramento estatístico, que a estrutura dos dados pode ser representada por componentes Gaussianos e não Gaussianos, quando se pretende usar esses modelos. Por outro lado, a não normalidade também é impeditiva, para quaisquer inferências baseadas nos testes t, Student, e F, Snedecor (SOUZA, 2005), frequentemente utilizadas nas estatísticas deste estudo.

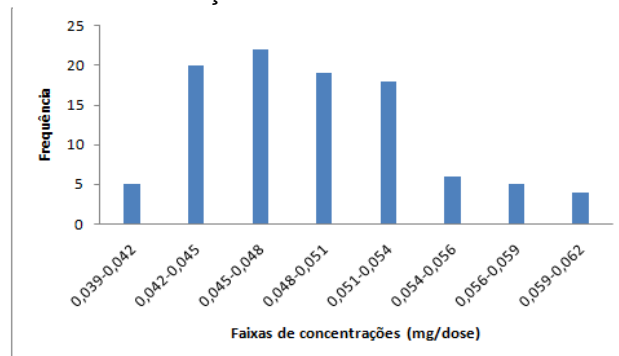
Essa suposição de normalidade para os dados amostrais de referência deste trabalho, condição exigida para regressão multivariada por mínimos quadrados parciais (RPLS), foi comprovada estatisticamente através do teste de Ryan Joiner. O gráfico de frequência (Figura 23), resultante da análise das faixas das concentrações (Tabela 17) para os dados de calibração, não apresentou uma

representação simétrica gaussiana, entretanto, a literatura reforça que dados que apresentam distribuição normal podem aparentar uma assimetria.

Tabela 17. Faixas de concentrações do set de calibração

Faixa (8 grupos)	Frequência
0,039-0,042	5
0,042-0,045	20
0,045-0,048	22
0,048-0,051	19
0,051-0,054	18
0,054-0,056	6
0,056-0,059	5
0,059-0,062	4

Figura 23. Histograma da distribuição das amostras do set de calibração



A partir das premissas H_0 (o conjunto provém de uma população com distribuição normal) e H_1 (o conjunto não provém de uma população com distribuição normal), o método de Ryan-Joiner (Tabela 18) foi aplicado para testar a distribuição normal dos dados de calibração. Utilizando o nível de confiança $\alpha=0,05$ para um $n=99$, resultou em Req (0,9880) maior que o $Rcrit$ (0,9833), dessa forma, não rejeitamos H_0 em um nível de significância de 5%, ou seja, não rejeitamos que a amostra provém de uma população com distribuição normal. Essa hipótese confere com a preocupação inicial deste estudo, no momento da amostragem das 500 frascos destinados à modelagem.

Tabela 18. Coeficientes calculados pelo teste de Ryan Joiner para normalidade dos dados da regressão.

<i>n</i>	99	<i>Seq</i>	0,4990
<i>qmed</i>	3,589E-17	<i>See</i>	0,002660
<i>emed</i>	0,04876	<i>Sqq</i>	95,88
<i>Req</i>=	0,9880	<i>Rcrit</i> (α: 0,05) =	0,9833

n, número amostral, *Seq*, *See*, *Sqq*, *qmed*, *emed*, coeficientes de cálculos; *Req*, coeficiente de correlação entre e_i e q_i ; *Rcrit*, coeficiente de correlação de Ryan Joiner crítico.

4.1.6 Modelagem (avaliações dos métodos de calibração, parâmetros de pré-tratamentos e variáveis latentes).

A qualidade do ajuste de previsão de um modelo multivariado, geralmente é avaliada de acordo com os baixos valores de RMSEC (amostras de calibração) e RMSEP (validação externa), com os seus altos coeficientes de correlação (R), respectivamente. A unidade dos erros é a mesma da propriedade de interesse deste estudo (mg/dose de nitrogênio proteico) e assim permitiu uma avaliação do erro médio associado à previsão. Na tabela 19, os valores de RMSEC (amostras de calibração) e RMSEP (validação externa) mais baixos, aparecem para os modelos PLS/MSC/CM (1 variável latente), PLS/SNV/CM, MRL/MSC/CM (1 fator), PCR/SNV/CM. As amostras encontram-se todas próximas à reta de correlação e os valores dos coeficientes de correlação para a calibração e a validação foram os mais altos, indicando a melhor concordância entre os valores esperados (obtidos pelo método de referência) e àqueles previstos pelos modelos e tratamentos avaliados. Por outro lado, o maior percentual de amostras com exatidão, no intervalo de 95 a 105%, conforme preconiza os parâmetros tabelados para a avaliação da exatidão, foram encontrados para o modelo PLS, justificando sua escolha.

Os altos valores para RMSEP e RMSEC, assim como o baixo valor de amostras recuperadas para os tratamentos com as derivadas, são atribuídos à maior concentração do ruído, nas regiões propostas pela literatura (5000-4545, 7143-6250 e 9091-9999 cm^{-1}) referentes aos estiramentos N-H, que provavelmente influenciaram no poder de previsão do modelo. Sendo o método MSC e SNV, os pré-tratamentos que foram mais satisfatórios para esse conjunto de dados, com erros muito próximos, visto que os gráficos correspondentes (Figura 24, Figura 25)

de correlação entre os valores de referência (kjeldahl) e os valores calculados pelos modelos, apresentaram comportamentos similares.

Tabela 19. Modelos avaliados com 99 valores para calibração e 99 para a validação, utilizando as 3 regiões sugeridas pela literatura.

Modelos/ pré-tratamento	RMSEC	R(SEC)	RMSEP	R(SEP)	Proporção de amostras com exatidão entre 95 a 105%
PLS (1 VL)	0,00625	0,37420	0,00717	0,30640	0,52
PLS/CM (1 VL)	0,00394	0,64960	0,00528	0,54520	0,58
PLS/MS (1 VL)	0,00515	0,51720	0,00607	0,36820	0,45
PLS/MS/CM (1 VL)	0,00088	0,98550	0,00334	0,86570	0,90
PLS/SNV	0,00780	0,79500	0,00918	0,71350	0,40
PLS/SNV/CM	0,00090	0,98480	0,00332	0,86370	0,86
PLS/1DER/SG(5pontos/2Ordem)	0,00451	0,79560	0,00566	0,76610	0,54
PLS/1DER/SG(5pontos/2Ordem)/CM	0,00273	0,85000	0,00379	0,79300	0,72
PLS/2DER/SG(5pontos/2Ordem)	0,00671	0,02590	0,00849	0,10400	0,39
PLS/2DER/SG(5pontos/2Ordem)/CM	0,00317	0,79120	0,00655	0,00180	0,54
CLS	0,00627	0,36805	0,00711	0,30050	0,54
CLS /MS	0,00520	0,5888	0,00603	0,44110	0,60
CLS/SNV	0,00642	0,60585	0,00769	0,56740	0,61
CLS/1DER/SG (5pontos/2Ordem)	0,00397	0,76471	0,00488	0,74680	0,65
CLS/2DER/SG (5pontos/2Ordem)	0,00538	0,11190	0,00614	0,11530	0,38
MRL(1 fator)	0,00164	0,95140	0,00372	0,84610	0,85
MRL/MS(1 fator)	0,00084	0,98690	0,00337	0,86430	0,80
MRL/SNV	0,00062	0,99400	0,00353	0,85070	0,81
MRL/1DER/SG(5pontos/2Ordem)	0,00209	0,91700	0,00404	0,78520	0,77
MRL/2DER/SG(5pontos/2Ordem)	0,00369	0,70270	0,00533	0,50220	0,72
PCR(1CP)	0,00622	0,3688	0,00713	0,3011	0,55
PCR/CM	0,00468	0,4294	0,00578	0,3540	0,55
PCR/MS	0,00517	0,5371	0,00609	0,3883	0,56
PCR /SNV/CM	0,00085	0,9865	0,00333	0,8630	0,82
PCR/1DER/SG(5pontos/2Ordem)/CM	0,00307	0,8064	0,00392	0,7748	0,78
PCR/2DER/SG(5pontos/2Ordem)/CM	0,00518	0,04208	0,0061	0,0644	0,15

RMSEC-raiz quadrada do Erro quadrático médio para valores de calibração (*Root Means Squares Errors*). RMSEP-raiz quadrada do Erro quadrático médio de previsão (*Root Means Squares Errors of Prediction*). R(SEC)-coeficiente de correlação atribuída aos dados de calibração. R(SEP)-coeficiente de correlação atribuída aos dados previstos. PCR - *Principle Component Regression* (Regressão por componentes principais). PLS - *Partial Least Squares* (Regressão por mínimos quadrados parciais). CLS-*Classical Least Squares* (Quadrados Mínimo Clássicos). MLR - *Multiple Linear Regression* (Regressão Linear Múltipla). CM-*Mean centering* (Método centrar na média). 2DER-Segunda derivada. 1DER-Primeira derivada. SG - Alisamento Savitsky-Golay. SNV- Variação de padrão normal (*Standard normal variate*). MSC-*Multiplicative Scatter Correction* (Correção Multiplicativa de Sinal).

Figura 24. Representação gráfica (A) de correlação entre valores de referência e valores previstos pelo modelo PLS.MSC.CM (1 fator) e (B) resíduos dos valores previstos e valores de referência.

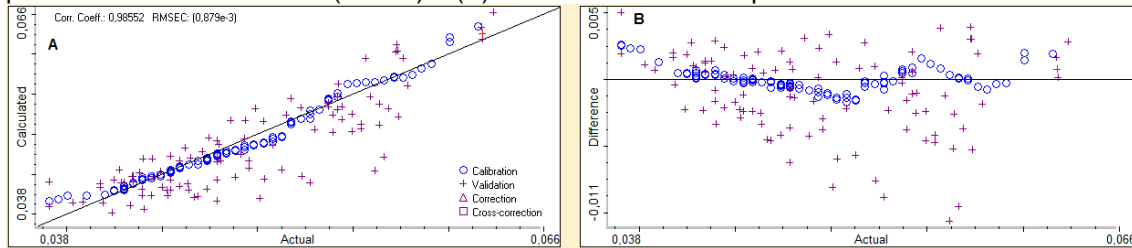
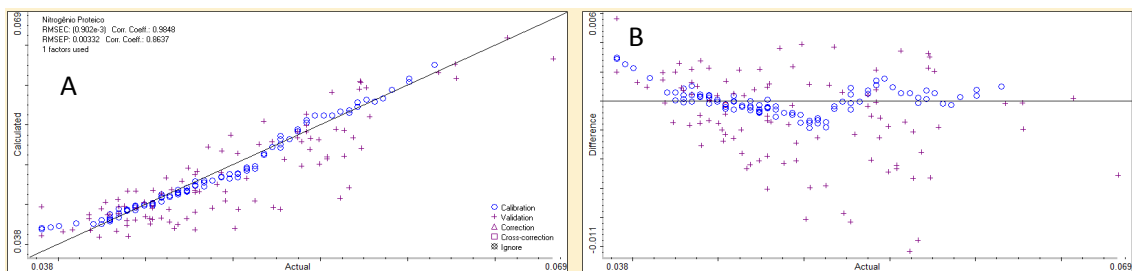


Figura 25. Representação gráfica (A) de correlação entre valores de referência e valores previstos pelo modelo PLS.SNV.CM (1 fator) e (B) resíduos dos valores previstos e valores de referência.



A tabela 20 apresenta a avaliação, quanto ao número adequado de variáveis latentes (VL) para o modelo PLS. Um número pequeno de fatores pode não ser satisfatório, ao passo que nem todas as informações, estão sendo totalmente exploradas pelo modelo. E quanto maior o número de fatores no modelo (variáveis latentes para o modelo PLS), menor será o desvio da reta de correlação, aumentando assim o coeficiente de correlação entre os valores de referência e os previstos pelo modelo multivariado, como é observado na tabela 19. Porém, o aumento demasiado e desnecessário de fatores, também aumenta o ruído e os erros de modelagem, ocasionando uma super estimação dos dados de calibração (VIANA, 2008), comprometendo a validação, como é observado na redução dos resultados de amostras recuperadas na validação externa.

Para que um modelo seja eficiente, é necessário que ele descreva da melhor forma possível a situação real (através dos dados de validação), levando-se em conta o maior número possível de informações que são encontradas através dos fatores contemplados. Então a determinação do número apropriado de fatores a serem utilizados, foi realizada pela análise do RMSEP (erro na validação externa), através do qual, o modelo mais apropriado foi alcançado para o método PLS (MSC e CM), utilizando um número mínimo de VL (1 variável latente utilizada), com menor

valor de RMSEP (0,00334), associado ao maior valor de percentual (90%) de amostras com exatidão na faixa de recuperação de 95 a 105% (Tabela 20).

Tabela 20. Avaliação do número ideal de variáveis latentes (VL) para o modelo PLS.

Modelos/pré-tratamentos	RMSEC	R(SEC)	RMSEP	R(SEP)	Proporção de amostras com exatidão entre 95 a 105%	RMSECV	R(SECV)
PLS /MSC/CM (1VL)	0,00088	0,9855	0,00334	0,8657	0,90	9,21E-04	0,9841
PLS /MSC/CM (2VL)	0,00087	0,9860	0,00332	0,8675	0,86	0,00093	0,9839
PLS /MSC/CM (5VL)	0,00077	0,9888	0,00334	0,8709	0,88	0,00106	0,9793
PLS /MSC/CM (8VL)	2,26E-04	0,999	0,00348	0,8612	0,88	0,00107	0,9786
PLS /MSC/CM(10VL)	1,30E-04	0,9997	0,00351	0,8601	0,87	0,00109	0,9779
PLS /SNV/CM (1VL)	0,00090	0,9848	0,00332	0,8637	0,86	9,39E-04	0,9835
PLS /SNV/CM (2VL)	6,87E-04	0,9912	0,00342	0,8597	0,83	7,39E-04	0,9898
PLS /SNV/CM (5VL)	6,16E-04	0,9929	0,00354	0,8517	0,79	7,19E-04	0,9903
PLS /SNV/CM (8VL)	2,73E-04	0,9986	0,0036	0,8473	0,75	8,86E-04	0,9853
PLS /SNV/CM (10VL)	1,19E-04	0,9997	0,00365	0,8433	0,74	9,30E-04	0,9838

RMSEC-raiz quadrada do Erro quadrático médio para valores de calibração (*Root Means Squares Errors*). RMSEP-raiz quadrada do Erro quadrático médio de previsão (*Root Means Squares Errors of Prediction*). RMSECV-raiz quadrada do Erro quadrático médio da validação cruzada (*Root Mean Square Error of Cross Validation*). R(SEC)-coeficiente de correlação atribuída aos dados de calibração. R(SEP)-coeficiente de correlação atribuída aos dados previstos. R(SECV)-coeficiente de correlação atribuída aos dados de validação cruzada.

Outros parâmetros importantes, para a avaliação do número ideal de VL, são o RMSECV (validação cruzada) e o PRESS, utilizados nas publicações de Li e Qu (2016) e Ferrão et al (2004). Observa-se que o número ideal foi alcançado para menores valores de RMSECV (0,00092). E o PRESS (Figura 26) da validação interna aumenta com o aumento da complexidade exigida para o modelo, ou seja, com o aumento do número de variáveis latentes, no modelo PLS com tratamento MSC (Tabela 21).

Para números de variáveis latentes entre 1 e 5, os erros da validação interna diminuem drasticamente, e aumenta para VL superiores a 5 (Figura 27 e Tabela 22). Esta tendência apresentada no PRESS de validação interna, não é observada com relação ao erro encontrado para esse tratamento (SNV) na validação externa (RMSEP), apresentada anteriormente na Tabela 20. Assim, o PRESS e o RMSECV são mais expressivos que RMSEP para avaliar VL.

Figura 26. Representação gráfica para modelo PLS.MSC.CM

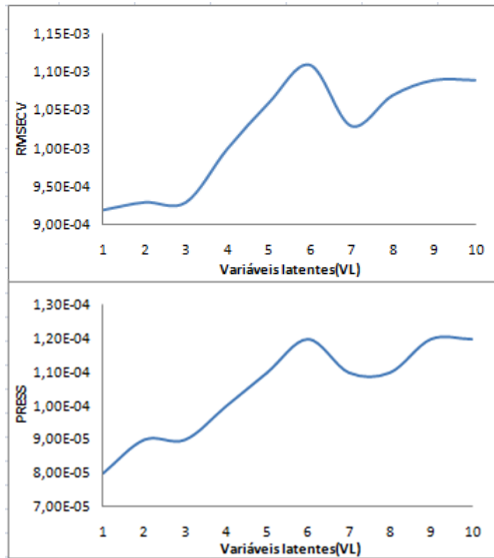


Tabela 21. Parâmetros estatísticos para PLS.MSC.CM.

VL	PRESS	RMSECV
1	8,00E-05	9,20E-04
2	9,00E-05	9,30E-04
3	9,00E-05	9,30E-04
4	1,00E-04	1,00E-03
5	1,10E-04	1,06E-03
6	1,20E-04	1,11E-03
7	1,10E-04	1,03E-03
8	1,10E-04	1,07E-03
9	1,20E-04	1,09E-03
10	1,20E-04	1,09E-03

PRESS-soma dos quadrados dos erros de previsão. VL-variável latente. RMSECV-raiz quadrada do Erro quadrático médio para na validação cruzada.

Os valores de PRESS e RMSECV foram menores para o modelo PLS com tratamento SNV a três variáveis latentes, entretanto esse número ideal (VL) deve ser coerente a quantidade de espécies a serem determinadas neste estudo, sendo apenas um componente (Nitrogênio proteico).

Figura 27. Representação gráfica para modelo PLS.SNV.CM

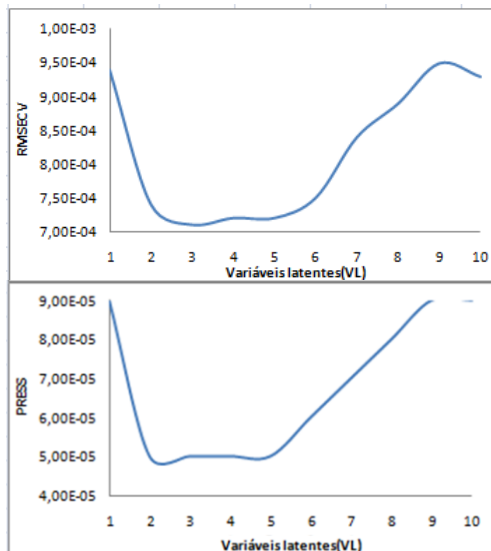


Tabela 22. Parâmetros estatísticos para PLS.SNV.CM.

VL	PRESS	RMSECV
1	9,00E-05	9,40E-04
2	5,00E-05	7,40E-04
3	5,00E-05	7,10E-04
4	5,00E-05	7,20E-04
5	5,00E-05	7,20E-04
6	6,00E-05	7,50E-04
7	7,00E-05	8,40E-04
8	8,00E-05	8,90E-04
9	9,00E-05	9,50E-04
10	9,00E-05	9,30E-04

PRESS-soma dos quadrados dos erros de previsão. VL-variável latente. RMSECV-raiz quadrada do Erro quadrático médio para na validação cruzada.

Os resultados baixos dos erros de validação interna com valores de VL diferentes de 1, indicam que o modelo PLS com SNV pode estar considerando

outros componentes na modelagem. Dessa forma o tratamento MSC viabiliza o número adequado de variáveis latentes contraposto aos dados estatísticos.

4.1.7 Otimização da região de trabalho

A figura 28 representa uma pequena parte da matriz utilizada para o cálculo dos coeficientes de correlação (R), entre os valores de referência do Kjeldahl (set de calibração) e os sinais de absorvância para seus respectivos números de onda (cm⁻¹), por sinais tratados com MSC.

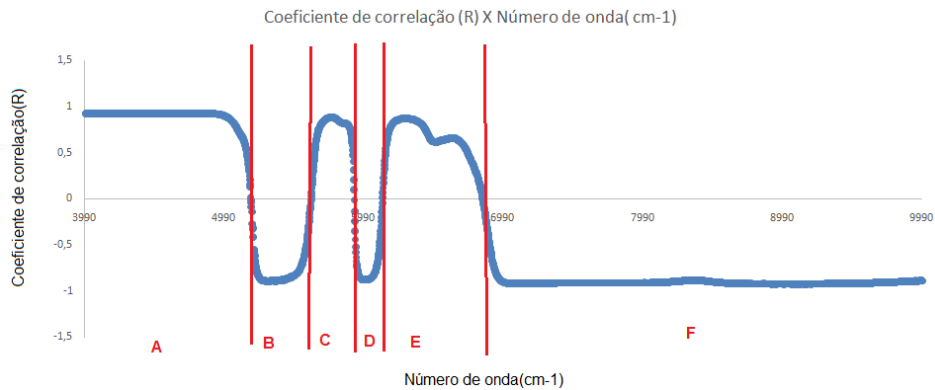
Figura 28. Representação da matriz para cálculo dos coeficientes de correlação.

	A	B	C	D	E	F	G	H	I	J	K	L	M
	numero	amostras de calibração	Kjeldahl mg/dose	Número de onda (cm-1)									
1				3999,64	4001,569	4003,497	4005,426	4007,354	4009,282	4011,211	4013,139	4015,068	4016,996
78	77	218	0,0419997	0,864247	0,864223	0,863972	0,863627	0,862798	0,862645	0,86252	0,862097	0,861659	0,860995
79	78	225	0,0444415	0,880954	0,880931	0,880659	0,880226	0,879241	0,879006	0,878923	0,878548	0,878059	0,877341
80	79	227	0,0429764	0,870487	0,870353	0,870107	0,869723	0,868723	0,868508	0,868477	0,868066	0,86747	0,866718
81	80	230	0,0507903	0,922982	0,922929	0,922632	0,922326	0,921429	0,921087	0,920919	0,920348	0,919793	0,918905
82	81	231	0,0493252	0,917644	0,917685	0,917327	0,917047	0,916122	0,915755	0,915597	0,914911	0,914187	0,913593
83	82	234	0,0454183	0,891985	0,891772	0,891515	0,891365	0,890425	0,890015	0,889869	0,889443	0,888883	0,888143
84	83	238	0,0468834	0,905953	0,905824	0,905615	0,905324	0,904115	0,90398	0,903867	0,903256	0,902655	0,901909
85	84	239	0,0566507	0,986378	0,986364	0,985918	0,985448	0,984349	0,984075	0,983634	0,982945	0,982408	0,981712
86	85	244	0,0459066	0,890972	0,890954	0,890704	0,890483	0,889696	0,889502	0,889184	0,888613	0,888189	0,887507
87	86	246	0,0468834	0,900181	0,900033	0,899819	0,899506	0,898555	0,898324	0,898028	0,897544	0,897136	0,896431
88	87	249	0,0400462	0,861258	0,86117	0,860887	0,860479	0,85957	0,859463	0,859417	0,859056	0,858528	0,857977
89	88	251	0,0507903	0,920356	0,920299	0,920122	0,919824	0,918732	0,918364	0,918214	0,917794	0,917315	0,916584
90	89	255	0,0429764	0,868952	0,868824	0,868686	0,868341	0,867295	0,867118	0,866956	0,866427	0,866094	0,865547
91	90	256	0,0586042	1,007719	1,00759	1,007112	1,006648	1,005244	1,004902	1,00474	1,004278	1,003415	1,002541
92	91	259	0,0542089	0,975997	0,976316	0,975937	0,975269	0,97407	0,973798	0,973424	0,972634	0,971832	0,971064
93	92	260	0,0498136	0,923059	0,923043	0,922757	0,922336	0,921259	0,92097	0,920771	0,920291	0,919525	0,918778
94	93	261	0,0483485	0,912165	0,911878	0,911356	0,911116	0,910317	0,910049	0,909822	0,909228	0,908628	0,908047
95	94	279	0,0488368	0,913485	0,913502	0,913247	0,912936	0,911943	0,911613	0,911317	0,910843	0,910334	0,909574
96	95	381	0,0590926	1,022208	1,021716	1,02092	1,020531	1,018998	1,018937	1,018914	1,018009	1,016981	1,015833
97	96	401	0,0390695	0,855528	0,855493	0,85518	0,854924	0,854034	0,853868	0,853905	0,853467	0,852947	0,852296
98	97	432	0,0537205	0,966478	0,966507	0,966196	0,965761	0,964704	0,964295	0,963957	0,963444	0,962853	0,961882
99	98	461	0,0429764	0,873655	0,873491	0,873128	0,872854	0,871924	0,871652	0,871554	0,871234	0,870752	0,870092
100	99	506	0,0478601	0,911274	0,911106	0,910786	0,910444	0,909299	0,909027	0,90885	0,908341	0,907832	0,906923
101													
102			Coefficiente de correlação:	0,98319	0,983196	0,983232	0,98331	0,983458	0,983371	0,983336	0,983366	0,983469	0,983533

Nas linhas estão dispostos os sinais (absorvância), para cada número de onda (cm⁻¹), organizados na linha superior. Na coluna principal, estão os respectivos valores de referência.

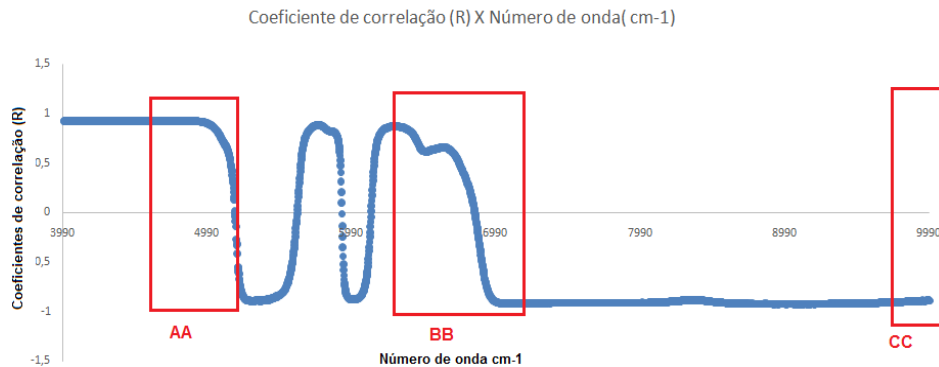
A representação gráfica da figura 29, apresenta a divisão das regiões de estudo, limitadas pela inversão de sinal dos coeficientes. O Gráfico referente à figura 30, apresenta as regiões delimitadas pela literatura (STUART, 2004).

Figura 29. Correlação entre os coeficientes de correlação com números de onda. **Regiões delimitadas para estudo.**



Regiões: A (3990-5184 cm^{-1}), B (5185-5616 cm^{-1}), C (5617-5928 cm^{-1}), D (5929-6129 cm^{-1}), E (6130-6852 cm^{-1}), F (6853-9999 cm^{-1}), utilizando a modelo PLS/MSC/CM, 99 valores de calibração e 99 valores de validação para uma variável latente.

Figura 30. Regiões da literatura (STUART,2004) que representam estiramentos de NH.



Regiões: AA(4545-5000 cm^{-1}), BB(6250-7143 cm^{-1}), CC(9091-10000 cm^{-1}).

Ao analisar os resultados estatísticos dispostos na tabela 23, observa-se a região A (4000-5184 cm^{-1}), com erros (RMSEC e RMSEP) mais satisfatórios que os resultantes das combinações com as demais regiões. Baddini et al (2010) enfatiza que para a determinação do conteúdo de proteína total, normalmente as bandas de absorção relacionadas, são encontrados ao redor 4587 cm^{-1} , e para seu estudo concluiu a região de trabalho em 4634-4527 cm^{-1} , para determinar proteína em sorohiperimune. Lima, Poppi e Andrade (2008) utilizaram a região em torno de 5053-6623 cm^{-1} para determinar nitrogênio total em amostras de plantas. Cox et al (2000) encontrou melhores resultados para a determinação de proteínas nos números de onda 4329, 4529 e 4675 cm^{-1} . Shenk et al (2008) ressalta que regiões

em torno de 4398cm^{-1} estão relacionadas às ligações C-N-C, referente às estruturas proteicas. Esses estudos concordam com os melhores resultados de recuperação encontrados neste trabalho, para a região A e AA (Tabelas 23 e 24).

Tabela 23. Parâmetros estatísticos para as regiões limitadas pelos coeficientes de correlação.

Regiões (Figura1)	RMSEC	R(SEC)	RMSEP	R(SEP)	Proporção de amostras com exatidão entre 95 a 105% (validação externa)
Todo o espectro	0,00089	0,9852	0,00334	0,8651	0,89
A	0,00088	0,9855	0,00330	0,8673	0,91
B	0,00232	0,8943	0,00386	0,7891	0,74
C	0,00280	0,8415	0,00433	0,7140	0,67
D	0,00284	0,8365	0,00383	0,8027	0,77
E	0,00345	0,7470	0,00536	0,5062	0,58
F	0,00096	0,9826	0,00340	0,8576	0,87
AB	8,79E-04	0,9855	0,00333	0,86730	0,91
AC	8,79E-04	0,9855	0,00333	0,86730	0,91
AD	8,79E-04	0,9855	0,00333	0,86740	0,91
AE	8,79E-04	0,9855	0,00333	0,86710	0,91
AF	8,89E-04	0,9852	0,00334	0,86530	0,90
BC	0,00177	0,9399	0,00370	0,8105	0,81
BD	2,27E-03	0,8986	0,00380	0,7974	0,77
BE	2,60E-03	0,8649	0,00432	0,7144	0,63
BF	9,61E-04	0,9827	0,00340	0,8575	0,87
CD	1,76E-03	0,9403	0,00365	0,8175	0,73
CE	2,99E-03	0,8171	0,00495	0,6041	0,71
CF	9,65E-04	0,9825	0,00340	0,8574	0,87
DE	3,04E-03	0,8097	0,00494	0,6023	0,75
DF	9,61E-04	0,9827	0,00340	0,8577	0,87
EF	9,69E-04	0,9824	0,00340	0,8568	0,87

RMSEC-raiz quadrada do Erro quadrático médio para valores de calibração (*Root Means Squares Errors*). RMSEP-raiz quadrada do Erro quadrático médio de previsão (*Root Means Squares Errors of Prediction*). R(SEC)-coeficiente de correlação atribuída aos dados de calibração. R(SEP)-coeficiente de correlação atribuída aos dados previstos. Regiões: A($3990-5184\text{ cm}^{-1}$), B($5185-5616\text{ cm}^{-1}$), C($5617-5928\text{ cm}^{-1}$), D($5929-6129\text{ cm}^{-1}$), E($6130-6852\text{ cm}^{-1}$), F($6853-9999\text{ cm}^{-1}$), utilizando a modelo PLS/MS/CM.

Explorando as combinações das regiões A (Tabela 23), nota-se a influência dessa região frente as demais, quando analisadas isoladamente. Os coeficientes de correlação negativos, observados nas regiões B, D e F, indicam contribuições de grupos que influenciam fortemente no modelo de regressão, e nessas regiões são encontradas combinações de O-H oriundas das bandas de água (STUART, 2004).

As regiões sugeridas por STUART (2004) e utilizadas desde o início deste estudo, apresentam combinações e sobretons de estiramentos de O-H e N-H, talvez relacionadas aos grupos dos excipientes, influenciando fortemente na previsão do modelo. Isso justifica em erros mais baixos e um percentual significativo de amostras com exatidão acima de 95%. Entretanto a região A ($4000-5184\text{ cm}^{-1}$), proposta pelos resultados correlacionados aos dados específicos deste estudo, e em concordância à região AA ($4545-5000\text{ cm}^{-1}$) inicialmente utilizada, evidencia a maior recuperação de valores previstos. Essa observação fortalece que a fonte da variação está relacionada com o conteúdo de proteína total, encontrada na região A.

Tabela 24. Parâmetros estatísticos para regiões sugeridas pela literatura para estiramentos NH.

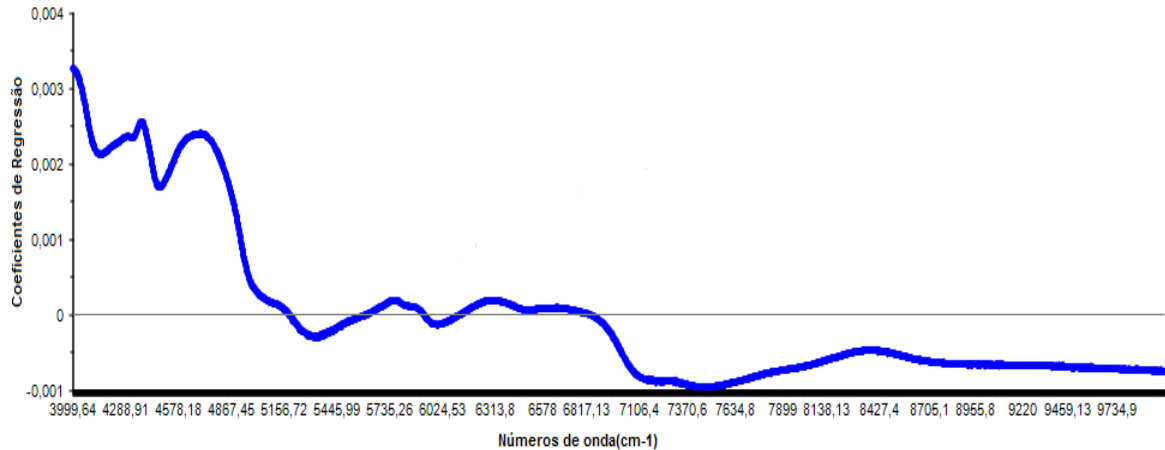
Regiões (Figura 2)	RMSEC	R(SEC)	RMSEP	R(SEP)	Proporção de amostras com exatidão entre 95 a 105% (validação externa)
Todo o espectro	0,00089	0,9852	0,00334	0,8651	0,89
AA	0,00088	0,9856	0,00333	0,8576	0,87
BB	0,00129	0,9684	0,00347	0,8433	0,81
CC	0,00139	0,9631	0,00347	0,8520	0,81
AA E BB	0,00089	0,9853	0,00333	0,8655	0,90
AA E CC	0,00088	0,9854	0,00334	0,8661	0,88
CC E BB	0,00121	0,9727	0,00342	0,8566	0,82

RMSEC-raiz quadrada do Erro quadrático médio para valores de calibração (*Root Means Squares Errors*). RMSEP-raiz quadrada do Erro quadrático médio de previsão (*Root Means Squares Errors of Prediction*). R(SEC)-coeficiente de correlação atribuída aos dados de calibração. R(SEP)-coeficiente de correlação atribuída aos dados previstos. Regiões: AA($4545-5000\text{ cm}^{-1}$), BB($6250-7143\text{ cm}^{-1}$), CC($9091-10000\text{ cm}^{-1}$).

Os coeficientes de regressão do modelo PLS, para a determinação de nitrogênio proteico, são apresentados na figura 31. É possível identificar as regiões espectrais que mais contribuiriam para a predição do nitrogênio. Como era esperado, os coeficientes mais significativos estão a cerca de $4000\text{ a }5000\text{ cm}^{-1}$, compatíveis

às vibrações ocasionadas por estiramentos das ligações N-H e coerentes aos resultados apresentados para a região A.

Figura 31. Representação gráfica entre os coeficientes de regressão do modelo PLS.CM e número de onda(cm^{-1}), calculados através do Unscrambler.



4.2 VALIDAÇÃO DO MÉTODO DE CALIBRAÇÃO MULTIVARIADA

4.2.1 Precisão

4.2.1.1 Repetibilidade

A tabela 25 apresenta os resultados obtidos em decuplicatas para as amostras, em três concentrações diferentes (mg.dose^{-1}), sendo os cálculos baseados no desvio padrão relativo. O método de calibração multivariado teve comportamento nas suas previsões, considerado repetitivo com um valor de $HorRat_{rep}$ menor que 2 (HORWITZ; ALBERT, 2006), e com desvio padrão relativo menor que 5% conforme exigência da resolução n°899 da Anvisa (2003). As sugestões da norma E1655-00 da ASTM (VALDERRAMA, 2005) foram atendidas, para as condições de repetibilidade, em um curto intervalo de tempo, com mínimo de três amostras em concentrações diferentes cobrindo a faixa útil do modelo de calibração e seis replicatas para cada nível de concentração.

Tabela 25. Precisão das leituras espectrais (no mesmo dia).

Frasco 571	(mg/dose)	Frasco 573	(mg/dose)	Frasco 581	(mg/dose)
1	0,04206	1	0,04707	1	0,06093
2	0,04208	2	0,04714	2	0,06099
3	0,04213	3	0,04655	3	0,06099
4	0,04212	4	0,04706	4	0,06081
5	0,04217	5	0,04709	5	0,06043
6	0,04238	6	0,04701	6	0,06094
7	0,04258	7	0,04712	7	0,06093
8	0,04168	8	0,04639	8	0,06091
9	0,04278	9	0,04743	9	0,06055
10	0,04149	10	0,04649	10	0,06089
<i>M</i> (<i>n</i> =10)	0,042		0,047		0,061
<i>s</i>	3,81E-04		3,38E-04		1,93E-03
<i>DPR rep.exp.</i> (%)	0,90		0,72		0,32
Cálculo do critério de aceitação por <i>HorRat</i> repetibilidade					
<i>FRASCO</i>	571	573		581	
<i>DPR Repto.teórico</i> (%)	4,68	4,68		4,68	
<i>DPR rep.teórico</i> (%)	3,12	3,12		3,12	
<i>Razão de HorRat rep.cal.</i>	0,29	0,23		0,10	

s, desvio padrão; *M*, média aritmética; *DPR_{rep.exp.}*(%), desvio padrão relativo (em condições de repetibilidade); *DPR_{Repto.teórico}*, desvio padrão relativo em condições de reprodutibilidade, teórico; *DPR_{rep.teórico}*, desvio padrão relativo teórico (em condições de repetibilidade); *HorRat_{rep}*, razão de Horwitz (para repetibilidade). Critério considerado pela Anvisa (2003), para repetibilidade, valores de *DPR_{rep.exp.}*(%) menores que 5%; critério considerando a razão de *HorRat_{rep.}*, menor ou igual a 2.

4.2.1.2 Precisão intermediária

As condições para a avaliação da precisão intermediária foram atendidas, tanto para a norma E1655-00 da ASTM, quanto para a Resolução n°899 da ANVISA (2003). Os critérios de aceitação foram satisfatórios, indicando precisão intermediária adequada, conforme tabela 26.

Tabela 26. Precisão intermediária.

Datas	Frasco 2 (mg/dose)		Frasco 3 (mg/dose)		Frasco 5 (mg/dose)	
	$M(n=3)$	$D.P$	$M(n=3)$	$D.P$	$M(n=3)$	$D.P$
04/08/15	0,05146	5,9E-05	0,04242	4,5E-05	0,04752	7,0E-05
05/08/15	0,05155	1,9E-04	0,04215	1,6E-04	0,04841	2,7E-04
06/08/15	0,05120	1,0E-03	0,04238	1,4E-04	0,04742	2,0E-05
07/08/15	0,05242	9,5E-05	0,04311	9,5E-05	0,04760	7,0E-05
08/08/15	0,05111	1,0E-04	0,04242	7,6E-05	0,04742	8,9E-05
09/08/15	0,05126	7,2E-05	0,04197	1,7E-04	0,04684	1,5E-05
$M (n=6)$	0,05150		0,04241		0,04753	
$s(n=6)$	4,81E-04		3,87E-04		5,06E-04	
$DPR \%_{(ANVISA)}$	0,93		0,91		1,06	
$DPR \%_{(INMETRO)}$	0,59		0,20		0,18	
Cálculos						
	Frasco 2		Frasco 3		Frasco 5	
$DPR_{rep.exp} (\%)$	0,93		0,91		1,06	
$DPR_{Repro.teórico}(\%)$	4,68		4,68		4,68	
$DPR_{rep.teórico}(\%)$	3,12		3,12		3,12	
$Razão\ de\ HorRat$	0,30		0,29		0,34	

Critérios de aceitação

Concentração em fração de massa considerada= 0,0035

Razão de HorRat <2

INMETRO (2011) DPR< 5,3%

ANVISA (2003) DPR< 5 %

s , desvio padrão; M , média; n , número de réplicas; $DPR_{(ANVISA)}$, desvio padrão relativo, calculado por sugestões da Anvisa (2003); $DPR_{(INMETRO)}$, desvio padrão relativo calculado pelas sugestões INMETRO (2011); $DPR_{rep.exp.} (\%)$, desvio padrão relativo (em condições de repetibilidade); $DPR_{Repro.teórico}$, desvio padrão relativo em condições de reprodutibilidade, teórico; $DPR_{rep.teórico}$, desvio padrão relativo teórico (em condições de repetibilidade); $HorRat_{rep}$, razão de Horwitz (para repetibilidade). Critério considerado pela Anvisa (2003), valores de $DPR_{rep.exp}(\%)$ menores que 5%; critério considerando a razão de $HorRat_{rep}$, menor ou igual a 2.

4.2.2 Exatidão

A comprovação da exatidão foi elaborada na seção 4.6, como parâmetro de avaliação e escolha entre a capacidade preditiva dos modelos de regressão. Como de praxe nos trabalhos analíticos com calibração multivariada (ADAMAPOULOS; GOULA, 2004; QU et al., 2008; FILGUEIRAS et al., 2012; DE SOUZA et al., 2013; ALVES; POPPI, 2016; GUÉVÉLOU; ALLEN Jr, 2016; GUO; NI; KOKOT, 2016; LI et al., 2016; LI; QU, 2016), a exatidão é expressa através da raiz quadrada do erro médio quadrático de previsão (RMSEP, *Root Mean Squares Error of Prediction*) conforme já apresentado neste trabalho. De forma a visualizar a capacidade de

previsão do modelo e contemplar as exigências da Resolução nº 899 (ANVISA, 2003), cada resultado do set de validação externa foi confrontada com os valores previstos pela regressão multivariada concluída (PLS com tratamento MSC e CM), resultando em mais de 90% das amostras com exatidão entre 95 e 105% (Tabela 27).

Tabela 27. Resultados (mg/dose) e percentual(%) de exatidão para os valores da validação externa. Sinalizado em vermelho as amostras que apresentaram valores inferiores a 95%.

Referência	NIR	%	Referência	NIR	%	Referência	NIR	%			
1	0,04239	0,04114	97	34	0,04274	0,04448	104	67	0,05274	0,0539	102
2	0,04395	0,04311	98	35	0,05128	0,04958	97	68	0,04786	0,04624	97
3	0,04188	0,04145	99	36	0,06281	0,06192	99	69	0,04809	0,04672	97
4	0,04346	0,04469	103	37	0,04284	0,04388	102	70	0,04798	0,04875	102
5	0,05034	0,04434	88	38	0,04444	0,04234	95	71	0,05542	0,0547	99
6	0,04084	0,04152	102	39	0,05422	0,05631	104	72	0,04542	0,04425	97
7	0,04649	0,04700	101	40	0,04298	0,04388	102	73	0,05493	0,0525	96
8	0,03907	0,04005	103	41	0,05372	0,05341	99	74	0,04711	0,04779	101
9	0,05665	0,05509	97	42	0,05567	0,0545	98	75	0,06177	0,06252	101
10	0,05786	0,05601	97	43	0,04861	0,05074	104	76	0,05542	0,05324	96
11	0,05011	0,04511	90	44	0,04591	0,04671	102	77	0,05395	0,05219	97
12	0,05177	0,05350	103	45	0,04639	0,04304	93	78	0,05733	0,06037	105
13	0,04737	0,04517	95	46	0,04956	0,04662	94	79	0,04937	0,05169	105
14	0,04546	0,04728	104	47	0,04346	0,04399	101	80	0,04446	0,04687	105
15	0,06274	0,06336	101	48	0,04786	0,04716	99	81	0,04335	0,04203	97
16	0,04786	0,04657	97	49	0,04688	0,04227	90	82	0,04591	0,0445	97
17	0,06572	0,06889	105	50	0,05421	0,05308	98	83	0,04633	0,04477	97
18	0,04403	0,04263	97	51	0,03907	0,04098	105	84	0,04956	0,04975	100
19	0,05079	0,0496	98	52	0,04177	0,0434	104	85	0,04503	0,04333	96
20	0,05688	0,05416	95	53	0,0403	0,04141	103	86	0,04523	0,04297	95
21	0,05639	0,05379	95	54	0,05395	0,05239	97	87	0,05726	0,05927	104
22	0,05233	0,04868	93	55	0,04344	0,04479	103	88	0,05616	0,05547	99
23	0,04379	0,04514	103	56	0,04398	0,04587	104	89	0,05372	0,05388	100
24	0,04744	0,04958	105	57	0,05591	0,05467	98	90	0,0453	0,04338	96
25	0,0503	0,05164	103	58	0,05763	0,05954	103	91	0,05637	0,05517	98
26	0,05226	0,05325	102	59	0,05737	0,05211	91	92	0,04249	0,04432	104
27	0,06835	0,06515	95	60	0,04988	0,05224	105	93	0,04786	0,04914	103
28	0,05442	0,05186	95	61	0,05298	0,05061	96	94	0,05444	0,05338	98
29	0,04395	0,0423	96	62	0,04786	0,04162	87	95	0,05728	0,06017	105
30	0,04933	0,04946	100	63	0,04177	0,04376	105	96	0,04667	0,04612	99
31	0,04451	0,0447	100	64	0,04493	0,04526	101	97	0,05495	0,05235	95
32	0,05421	0,05474	101	65	0,04884	0,04459	91	98	0,04911	0,04748	97
33	0,04542	0,04601	101	66	0,04621	0,04827	104	99	0,05149	0,05234	102

4.2.3 Robustez

Para a avaliação da variação da temperatura ambiente (Tabela 28), verificou-se através da análise de variância, ANOVA, entre todas as temperaturas consideradas, que o valor de *p-valor* (0,0059) observado, não foi superior a 0,05, indicando heterogeneidade entre as variâncias, ou seja, que ao menos uma é diferente das demais variâncias. Dessa forma, foi necessária a aplicação do teste *F* (SNEDECOR; COCHRAN 1989) para cada variação em relação à temperatura de trabalho, tomada como referência.

A heterocedasticidade somente foi identificada para a variação de temperatura de 27°C, que apresentou variância não homogênea em relação à amostra nas condições normais (23°C), com a estatística F_{cal} (14,94) maior que F_{tab} (3,18) para uma significância α de 0,05, a um nível de confiança de 95%. Para essa variação foi aplicada a estatística *t* com variâncias amostrais não homogêneas (Equações 28 e 29). O método não foi considerado robusto para essa temperatura, visto que a estatística $t_{calculado}$ foi maior que $t_{tabelado}$, a significância $\alpha/2$ de 0,025, para um nível de confiança de 95%, com 10 graus de liberdade ($v=10,28$), definidos pela equação 29.

Para as demais variações de temperatura, com a homoscedasticidade entre as variâncias comprovadas através dos valores de F_{cal} (Tabela 28), as médias foram comparadas através do teste *t* para variâncias homogêneas. O método foi considerado robusto para as variação de 22°, 24° e 25°C, visto que a estatística $t_{calculado}$ (Equação 26) foi menor que $t_{tabelado}$, com $\alpha/2$ de 0,025, para um nível de confiança de 95%, com 9 (n-1) graus de liberdade.

Tabela 28. Avaliação da temperatura ambiente.

Frasco 667					
Temperatura	23°C (Ref.)	25°C	22°C	27°C	24°C
<i>Datas</i>	13.10.15	16.10.15	28.10.15	26.11.2016	07.01.2016
$M(n=10)$	0,04841	0,04824	0,04812	0,04729	0,04862
s	0,00037	0,00064	0,00047	0,00143	0,00055
s^2	1,37E-07	4,07E-07	2,21E-07	2,04E-06	3,02E-07
$DPR(\%)$	0,76	1,32	0,98	3,02	1,13
t_{tab}		2,26	2,26	2,23	2,26
t_{cal}		0,71	-0,97	2,39	-1,01
F_{tab}		3,18	3,18	3,18	3,18
F_{cal}		2,98	1,62	14,94	2,21
ANOVA	F	p -valor	F crítico		
$\alpha : 0,05$	4,167	0,0059	2,579		

Critérios de aceitação: 95% confiança, $t_{tab}=2,26$ ($n-1=9$); $t_{tab}=2,23$ ($v=10,2$); $f_{tab}=3,18$; t_{cal} , estatística t calculada; F_{cal} , estatística f calculada; t_{tab} estatística t tabelada; F_{tab} estatística f tabelada, F , estatística da análise de variância; p -valor, probabilidade da estatística calculada; s , desvio padrão; DPR , desvio padrão relativo; M , média aritmética; n , número de réplicas; ($n-1$), graus de liberdade para estatística t (homocedástica); v , graus de liberdade para estatística t (heterocedástica). s^2 , variância.

Para a avaliação da variação do tempo antes da leitura (Tabela 29), pós degelo, verificou-se através da análise de variância, ANOVA, entre todos os tempos considerados, que o valor de p -valor (0,43) observado, foi superior a 0,05, indicando homogeneidade entre as variâncias. Dessa forma, as médias foram comparadas através do teste t para variâncias homogêneas (Equações 26 e 27). O método foi considerado robusto para todos os tempos avaliados (Tabela 29), visto que a estatística $t_{calculado}$ (Equação 26) foi menor que $t_{tabelado}$, com $\alpha/2$ de 0,025, para um nível de confiança de 95%, com 9($n-1$) graus de liberdade.

Tabela 29. Avaliação do tempo de leitura a partir do degelo (mesmo dia).

Frasco 695							
	10 min (Ref.)	1 min	15 min	1 horas	2 horas	4 horas	6 horas
$M(n=10)$	0,04516	0,04486	0,04507	0,04506	0,04519	0,04521	0,04520
s	0,00035	0,00040	0,00048	0,00030	0,00033	0,00044	0,00039
s^2	1,23E-07	1,60E-07	2,30E-07	9,20E-08	1,08E-07	1,90E-07	1,55E-07
$DPR(\%)$	0,78	0,89	1,06	0,67	0,73	0,97	0,87
t_{cal}		1,84	0,81	0,73	-0,19	-0,25	-0,27
ANOVA	F	p -valor	F crítico				
$\alpha : 0,05$	1,005	0,430	2,246				

Critérios de aceitação: 95% confiança, $t_{tab}=2,26$ ($n-1=9$); t_{cal} , estatística t calculada; t_{tab} estatística t tabelada; F , estatística da análise de variância; p -valor, probabilidade da estatística calculada;

DPR, desvio padrão relativo; *s*, desvio padrão; *M*, média aritmética; *n*, número de réplicas. s^2 , variância.

Para a avaliação da troca de lâmpada, verificou-se através da análise de variância, ANOVA, para cada amostra, que os valores de *p-valor* observados foram superior a 0,05, indicando homogeneidade entre as variâncias. Dessa forma, as médias foram comparadas através do teste *t* para variâncias homogêneas (Equações 26 e 27). O método foi considerado robusto quanto a troca deste acessório instrumental (Tabela 30), visto que a estatística $t_{calculado}$ (Equação 26) foi menor que $t_{tabelado}$ a uma significância $\alpha/2$ de 0,025, para um nível de confiança de 95%, com 9 (*n*-1) graus de liberdade.

Tabela 30. Avaliação após a troca de lâmpada.

	Frasco 2		Frasco 3		Frasco 5	
	02/02/16	13/04/16	02/02/16	13/04/16	02/02/16	13/04/16
	antes	depois	antes	depois	antes	depois
<i>M</i> (<i>n</i> =10)	0,05241	0,05231	0,04245	0,04238	0,04742	0,04707
<i>s</i>	0,00011	0,00017	0,00017	0,00020	0,00032	0,00049
s^2	1,19E-08	3,06E-08	2,77E-08	4,15E-08	1,01E-07	2,44E-07
<i>DPR</i> (%)	0,21	0,33	0,39	0,48	0,67	1,05
t_{cal}	1,93		0,70		1,56	
<i>p-valor</i> -(ANOVA)	0,1500		0,4113		0,0784	
<i>Fcal</i> (ANOVA)	2,262		0,708		3,483	
<i>F crítico</i> (ANOVA)	4,414		4,414		4,414	

Critérios de aceitação: 95% confiança, $t_{tab}=2,26$ (*n*-1=9); t_{cal} , estatística *t* calculada; t_{tab} estatística *t* tabelada; *F*, estatística da análise de variância; *p-valor*, probabilidade da estatística calculada; *s*, desvio padrão; *M*, média; *n*, número de réplicas; s^2 , variância.

Para a avaliação da variação da posição do frasco em relação a um referencial (Tabela 31), verificou-se através da análise de variância, ANOVA, que os valores de *p-valor* observados, foram superior a 0,05, indicando homogeneidade entre as variâncias para os frascos avaliados. Dessa forma, as médias foram comparadas através do teste *t* para variâncias homogêneas (Equações 26 e 27). O método foi considerado robusto quanto a variação da posição de leitura do frasco da amostra, visto que a estatística $t_{calculado}$ (Equação 26) foi menor que $t_{tabelado}$, com $\alpha/2$ de

0,025, para um nível de confiança de 95%, com 9 graus de liberdade. A avaliação para mais de um frasco, sugere que a diferença entre os materiais dos frascos, possivelmente não contribuem para alterações significativas, nos valores previstos pelo método, ao se variar a posição do frasco.

Tabela 31. Avaliação da variação da posição do frasco.

	Frasco 640			
	1 (Ref.)	2 (90°)	3(180°)	4(360°)
$M(n=10)$	0,04430	0,04431	0,04429	0,04431
s	0,00009	0,00012	0,00010	0,00012
s^2	8,87E-09	1,43E-08	9,88E-09	1,52E-08
$DPR(\%)$	0,21	0,27	0,22	0,28
t_{cal}		-0,24	0,28	-0,31
ANOVA	F	p-valor	F crítico	
	0,0789	0,971	2,90	
	Frasco 642			
	1 (Ref.)	2 (90°)	3(180°)	4(360°)
$M(n=10)$	0,04484	0,04485	0,04483	0,04482
s	0,00010	0,00013	0,00012	0,00033
s^2	1,09E-08	1,79E-08	1,37E-08	1,10E-07
$DPR(\%)$	0,23	0,30	0,26	0,74
t_{cal}		-0,23	0,37	0,13
ANOVA	F	p-valor	F crítico	
	0,0453	0,987	2,90	
	Frasco 646			
	1 (Ref.)	2 (90°)	3(180°)	4(360°)
$M(n=10)$	0,04254	0,04268	0,04260	0,04257
s	0,00017	0,00019	0,00013	0,00018
s^2	2,94E-08	3,43E-08	1,57E-08	3,30E-08
$DPR(\%)$	0,40	0,43	0,29	0,43
t_{cal}		-1,77	-0,89	-0,33
ANOVA	F	p-valor	F crítico	
	1,105	0,362	2,90	

Critérios de aceitação: 95% confiança, $t_{tab}=2,26(n-1=9)$; t_{cal} , estatística t calculada; t_{tab} estatística t tabelada; F , estatística da análise de variância; p -valor, probabilidade da estatística calculada; s , desvio padrão; M , média; n , número de réplicas; s^2 variância

4.2.4 Teste de erros sistemáticos no modelo – viés (*bias*)

O viés (*bias*) foi calculado apenas com base nas amostras de validação, conforme a tabela 32. Com o valores de *bias* e *SDV* apresentados, um valor de $t_{bias} = 1,9798$ foi estimado para o modelo, o que está abaixo do $t_{crítico}$ de 1,9870 com 99

graus de liberdade e no nível de confiança de 95%. Ou seja, através dos cálculos utilizados, não possível comprovar a existência de erros sistemáticos incluídos no modelo, desse modo, o *bias* pode ser considerado insignificante e desprezado.

Tabela 32. Resultados para erros sistemáticos no modelo de calibração.

<i>bias</i>	-0,0004470	<i>t bias</i>	1,9798
<i>SDV</i>	0,002246	<i>t crit (n=99)</i>	1,9870

bias, erro sistemático; *SDV*, desvio padrão para a validação; *t bias*, estatística t student calculado; *t crit (n=99)*, estatística t student tabelada, para *n* graus de liberdade.

4.2.5 Recuperação

É importante considerar, como a eficiência preditiva da calibração varia em função de alterações das concentrações, tanto dos excipientes quanto do analito. Na maioria dos casos, a dispersão dos resultados aumenta com a diminuição da concentração e a recuperação pode ser diferente quando são adicionadas altas ou baixas concentrações. Por esse motivo, a recuperação foi avaliada dentro da faixa de concentração esperada para o componente de interesse (de 0,2% a 0,35%) inferindo uma faixa de recuperação de 95 a 105% (Tabela 33).

Tabela 33. Avaliação para incremento de 0,02 mg.dose⁻¹ em nitrogênio proteico através de solução padrão de ovoalbumina.

	Concentrações (mg/dose) originais		Concentrações (mg/dose) reliofilizadas		<i>Dif</i>	Concentração esperada	<i>Rec.(%)</i>
	<i>M(n=3)</i>	<i>s</i>	<i>M(n=3)</i>	<i>s</i>			
526	0,05079	1,34E-04	0,07190	7,37E-05	0,021	0,07079	101,6
529	0,05098	1,76E-04	0,07323	6,43E-05	0,022	0,07098	103,2
533	0,04425	1,23E-04	0,06357	1,17E-04	0,019	0,06425	98,9
534	0,05360	1,35E-04	0,07504	3,46E-05	0,021	0,07360	102,0
535	0,04316	1,88E-04	0,06401	1,48E-04	0,021	0,06316	101,4
536	0,04122	1,34E-04	0,06023	1,89E-04	0,019	0,06122	98,4
537	0,04421	1,67E-04	0,06508	2,10E-04	0,021	0,06421	101,4
538	0,04277	2,16E-04	0,06189	1,56E-04	0,019	0,06277	98,6

M, média aritmética; *n*, replicatas; *s*, desvio padrão; *Dif.*, diferença entre a concentração original e a concentração final pós reliofilização; *Rec.*, recuperação.

A recuperação alcançada para valores de incremento de 0,02 mg/dose em nitrogênio proteico, com adição do padrão de ovoalbumina às amostras, todas apresentaram percentuais entre 95 a 105% (Tabela 34). Diferentemente, do

incremento de 0,01mg/dose em nitrogênio proteico através do padrão de ovoalbumina, com recuperações não tão satisfatórias, fora da faixa de aceitação (Tabela 34). Esse resultado pode inferir que incrementos próximos ao desvio padrão das leituras, não apresentam substancialmente recuperações aceitáveis.

Tabela 34. Teste para incremento de 0,01mg.dose⁻¹ em nitrogênio proteico através de solução padrão de ovoalbumina.

	Concentrações (mg/dose) originais		Concentrações (mg/dose) reliofilizadas		<i>Dif</i>	Concentração esperada	<i>Rec.(%)</i>
	<i>M(n=3)</i>	<i>s</i>	<i>M(n=3)</i>	<i>s</i>			
671	0,05277	1,88E-04	0,05827	1,14E-04	0,0075	0,06277	92,8
672	0,04021	1,26E-04	0,04561	1,65E-04	0,0054	0,05021	90,8
673	0,04645	1,46E-04	0,04414	1,47E-04	0,0023	0,05645	78,2
674	0,04213	2,00E-04	0,04539	1,35E-04	0,0033	0,05213	87,1
675	0,04432	1,32E-04	0,04139	1,60E-04	0,0029	0,05432	76,2
676	0,06011	1,97E-04	0,05739	1,78E-04	0,0027	0,07011	81,9
677	0,05521	1,47E-04	0,05739	1,81E-04	0,0022	0,06521	88,0

M, média aritmética; *n*, replicatas; *s*, desvio padrão; *Dif.*, diferença entre a concentração original e a concentração final pós reliofilização; *Rec.*, recuperação.

4.2.6 Avaliação da influência espectral dos excipientes da matriz

Conforme argumentado anteriormente, a importância da verificação da eficiência preditiva da calibração em função de alterações das concentrações, foram avaliados os incrementos de concentrações para cada excipiente da matriz vacinal. Para nenhum dos excipientes avaliados, foram obtidas as recuperações esperadas (Tabelas 35 e 36). Esse resultado está coerente às substâncias que não apresentaram concentrações de nitrogênio proteico, quando suas soluções foram analisadas pelo método Kjeldahl, mesmo aos compostos que apresentam ligações de nitrogênio em suas fórmulas estruturais. Comprovando que para a concentração analisada, não foi observada influência na resposta espectral, associada à variação dos componentes da matriz.

Tabela 35. Teste de recuperação com incremento de 0,02 mg.dose⁻¹ em nitrogênio presente na composição dos excipientes.

	Concentração (mg/dose)/M(n=3)			Concentração esperada	Rec. (%)
	Originais	Reliofilizadas	Dif.		
CANAMICINA					
630	0,05481	0,05399	0,00082	0,07481	4,1
631	0,05252	0,05235	0,00017	0,07252	0,8
632	0,05027	0,05230	-0,00203	0,07027	10,2
633	0,04726	0,04731	-0,00005	0,06726	0,3
634	0,05954	0,04940	0,01014	0,07954	50,7
635	0,04643	0,04682	-0,00039	0,06643	2,0
637	0,04825	0,05261	-0,00436	0,06825	21,8
638	0,04283	0,04358	-0,00075	0,06283	3,7
ERITROMICINA					
620	0,05311	0,05125	0,00186	0,07311	9,3
621	0,05454	0,04633	0,00821	0,07454	41,0
622	0,04977	0,04761	0,00216	0,06977	10,8
623	0,04386	0,04632	-0,00246	0,06386	12,3
624	0,04195	0,05526	-0,01331	0,06195	66,6
625	0,05234	0,05021	0,00213	0,07234	10,7
626	0,05167	0,05500	-0,00333	0,07167	16,7
629	0,04907	0,04901	0,00006	0,06907	0,3
GLUTAMATO					
594	0,05576	0,05924	-0,00348	0,07576	17,4
596	0,04322	0,04242	0,00080	0,06322	4,0
597	0,04871	0,04305	0,00566	0,06871	28,3
598	0,04690	0,04505	0,00185	0,06690	9,2
599	0,06147	0,05966	0,00181	0,08147	9,1
600	0,04052	0,04619	-0,00567	0,06052	28,4
601	0,06277	0,06401	-0,00124	0,08277	6,2
602	0,04448	0,05033	-0,00585	0,06448	29,3

M, média aritmética; n, replicatas; Dif., diferença entre a concentração original e a concentração final pós reliofilização; Rec., recuperação.

Tabela 36. Teste de recuperação com incremento de 0,02 mg.dose⁻¹ em nitrogênio presente na composição dos excipientes.

	Concentração (mg.dose ⁻¹)/M(n=3)			Concentração esperada	Rec. (%)
	Originais	Reliofilizadas	Dif.		
GELATINA					
512	0,04400	0,04847	-0,00447	0,06400	22,4
515	0,05302	0,04946	0,00356	0,07302	17,8
516	0,04652	0,04614	0,00038	0,06652	1,9
517	0,05554	0,05268	0,00286	0,07554	14,3
518	0,04666	0,04329	0,00337	0,06666	16,9
520	0,04212	0,04857	-0,00645	0,06212	32,2
522	0,04551	0,04027	0,00524	0,06551	26,2
524	0,04069	0,04195	-0,00126	0,06069	6,3
SACAROSE					
610	0,05143	0,05117	0,00026	0,07143	1,3
612	0,05925	0,06394	-0,00469	0,07925	23,5
613	0,04985	0,04914	0,00071	0,06985	3,6
614	0,05606	0,05421	0,00185	0,07606	9,2
616	0,04150	0,05699	-0,01549	0,06150	77,4
617	0,05026	0,05794	-0,00768	0,07026	38,4
618	0,04573	0,05373	-0,00800	0,06573	40,0
619	0,05102	0,04829	0,00273	0,07102	13,7
SORBITOL					
590	0,04815	0,05784	-0,00969	0,06815	48,5
591	0,04815	0,04769	0,00046	0,06815	2,3
592	0,04416	0,05259	-0,00843	0,06416	42,1
593	0,04272	0,05279	-0,01007	0,06272	50,4
606	0,04772	0,04095	0,00677	0,06772	33,8
607	0,04363	0,04594	-0,00231	0,06363	11,6
608	0,04399	0,04418	-0,00019	0,06399	0,9
609	0,05071	0,05696	-0,00625	0,07071	31,3
ÁGUA WFI					
540	0,06119	0,05952	-0,00167	0,08119	8,3
541	0,04064	0,04554	0,00490	0,06064	24,5
542	0,04995	0,04076	-0,00919	0,06995	46,0
543	0,04848	0,04352	-0,00496	0,06848	24,8
544	0,04687	0,04911	0,00224	0,06687	11,2
545	0,04598	0,05062	0,00464	0,06598	23,2
546	0,05951	0,05706	-0,00245	0,07951	12,2
547	0,04629	0,05218	0,00589	0,06629	29,5

M, média aritmética; n, replicatas; Dif., diferença entre a concentração original e a concentração final pós reliofilização; Rec., recuperação.

4.3 OUTRAS AVALIAÇÕES

4.3.1 Precisão (*inter meses*)

Embora não esteja no procedimento exigido pelos órgãos considerados neste trabalho, foi avaliado durante todos os meses de estudo, a variação da aquisição espectral das amostras controles (Tabela 37).

Tabela 37. Precisão das amostras controles durante os meses e dias de estudo.

Datas	Frasco 2 (mg.dose ⁻¹)		Frasco 3 (mg.dose ⁻¹)		Frasco 5 (mg.dose ⁻¹)	
	<i>M</i> (<i>n</i> =3)	<i>s</i> (<i>n</i> =3)	<i>M</i> (<i>n</i> =3)	<i>s</i> (<i>n</i> =3)	<i>M</i> (<i>n</i> =3)	<i>s</i> (<i>n</i> =3)
04/08/15	0,05146	5,9E-05	0,04242	4,5E-05	0,04752	7,0E-05
05/08/15	0,05155	1,9E-04	0,04215	1,6E-04	0,04841	2,7E-04
06/08/15	0,05120	1,0E-03	0,04238	1,4E-04	0,04742	2,0E-05
07/08/15	0,05242	9,5E-05	0,04311	9,5E-05	0,04760	7,0E-05
08/08/15	0,05111	1,0E-04	0,04242	7,6E-05	0,04742	8,9E-05
09/08/15	0,05126	7,2E-05	0,04197	1,7E-04	0,04684	1,5E-05
12/08/15	0,05238	2,5E-04	0,04268	1,5E-04	0,04717	1,4E-04
14/08/15	0,05221	2,3E-04	0,04335	1,8E-04	0,04757	3,0E-04
18/08/15	0,05232	1,1E-04	0,04171	1,5E-04	0,04864	1,6E-04
19/08/15	0,05314	4,6E-05	0,04241	4,2E-04	0,04752	2,6E-04
31/08/15	0,05135	1,0E-04	0,04246	1,6E-04	0,04822	1,1E-04
16/09/15	0,05017	4,6E-05	0,04225	4,4E-05	0,04757	2,4E-04
18/09/15	0,04967	7,8E-04	0,04235	5,9E-05	0,04710	6,2E-05
28/10/15	0,05237	1,4E-04	0,04223	1,9E-04	0,04785	2,8E-04
23/11/15	0,04853	1,4E-03	0,04260	2,8E-04	0,04842	1,5E-04
07/01/16	0,05160	1,3E-03	0,04173	4,5E-05	0,04728	2,5E-04
02/02/16	0,05245	5,8E-05	0,04253	1,8E-04	0,04755	1,8E-04
13/04/16	0,05229	2,5E-04	0,04222	2,1E-05	0,04686	3,8E-05
<i>M</i> (<i>n</i> =18)	0,05153		0,04239		0,04761	
<i>s</i> (<i>n</i> =18)	1,1E-03		4,1E-04		5,2E-04	
<i>DPR</i> (%)	2,21		0,96		1,09	

Cálculo do critério de aceitação por *HorRat*

	Frasco 2	Frasco 3	Frasco 5
<i>DPR</i> _{prec.exp} (%)	2,21	0,96	1,09
<i>DPR</i> _{Repro.teórico} (%)	4,68	4,68	4,68
<i>DPR</i> _{prec.teórico} (%)	3,12	3,12	3,12
Razão de <i>HorRat</i> _{prec}	0,71	0,31	0,35

Critérios de aceitação

Concentração em fração de massa considerada= 0,0035; Razão de *HorRat* :<2

s, desvio padrão; *M*, média; *DPR*, desvio padrão relativo simples; *DPR*_{prec.exp} (%), desvio padrão relativo; *DPR*_{Repro.teórico}, desvio padrão relativo em condições de reprodutibilidade, teórico; *DPR*_{prec.teórico}, desvio padrão relativo teórico (em condições de repetibilidade); *HorRat*_{prec}, razão de Horwitz (para precisão). Critério considerando a razão de *HorRat*_{prec}, menor ou igual a 2.

A razão de *HorRat* calculada para todas as amostras avaliadas, foram menores que 2, demonstrando precisão adequada, indicando confiabilidade em todas as aquisições espectrais que foram realizadas nas datas apresentadas na tabela 37.

4.3.2 Avaliação da previsão do modelo com lotes de rotina

A estatística *F* do teste foi calculada através da Equação 25, para o valor previsto (NIR) de cada lote produzido em 2016, em relação ao respectivo valor fornecido pelo método de referência (KJ), conforme tabela 38. A heterocedasticidade entre as variâncias foi observada para todos os lotes avaliados, com a estatística F_{cal} maior que $F_{tab}(19,0)$ para uma significância α de 0,05, a um nível de confiança de 95%. Dessa forma, foi aplicada a estatística *t* para variâncias amostrais não homogêneas (Equações 28 e 29). A estatística $t_{calculado}$ foi menor que $t_{tabelado}$ para todos os lotes, com $\alpha/2$ de 0,025, para um nível de confiança de 95%. Essa comparação indica estatisticamente que a capacidade preditiva para a calibração multivariada em questão, diante de diferentes lotes e diferentes analistas, conforme proposta de rotina de análise, está adequada para o que se propôs.

Tabela 38. Previsão de lotes de rotina produzidos em 2016.

Lotes	1° Analista	2° Analista	F_{cal}	F_{tab}	t_{cal}	t_{tab}	GLc
	NIR	KJ					
VFC1 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,048	0,052	87,26	19,0	-2,52	4,303	2,05
Variância	1,04E-07	9,08E-06					
VFC2 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,047	0,040	229,44	19,0	2,55	4,303	2,02
Variância	1,10E-07	2,52E-05					
VFC3 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,042	0,041	185,61	19,0	0,39	4,303	2,02
Variância	1,96E-07	3,63E-05					
VFC4 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,043	0,044	144,39	19,0	-1,18	4,303	2,03
Variância	2,26E-08	3,27E-06					
VFC5 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,042	0,043	116,81	19,0	-0,72	4,303	2,03
Variância	9,40E-08	1,10E-05					
VFC6 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,044	0,044	59,12	19,0	-0,29	4,303	2,07
Variância	1,07E-07	6,30E-06					
VFC7 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,039	0,041	71,27	19,0	-1,87	4,303	2,06
Variância	4,09E-08	2,91E-06					
VFC8 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,039	0,040	25,51	19,0	-0,84	4,303	2,16
Variância	5,69E-08	1,45E-06					
VFC9 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,044	0,048	33,84	19,0	-3,04	4,303	2,12
Variância	1,19E-07	4,03E-06					
VFC10 (mg.dose ⁻¹) <i>M</i> (<i>n</i> =3)	0,052	0,057	274,98	19,0	-1,35	4,303	2,01
Variância	1,80E-07	4,94E-05					

VFC, vacina de febre amarela 10 doses; t_{cal} , estatística t calculada; t_{tab} estatística t tabelada; F_{cal} , estatística calculada da análise de variância; F_{tab} , estatística tabelada da análise de variância; *M*, média aritmética; *n*, número de réplicas; *GLc*, graus de liberdade combinado.

4.4 ESTIMATIVA DA INCERTEZA DA CALIBRAÇÃO MULTIVARIADA

4.4.1 Estimativa da incerteza dos resultados do método de referência

O procedimento da estimativa da incerteza depende do conhecimento detalhado da técnica analítica, ou seja, da natureza da medição e seus pontos críticos. Dessa forma, foi compreendido nos estudos apresentados anteriormente, o

quanto é importante a confiabilidade dos dados de referência e como eles afetam fortemente a previsão do modelo multivariado. Assim, o primeiro passo foi a identificação das fontes de incerteza da metodologia analítica, empregada para fornecer os valores de referência, que estão compreendidos na etapa de fatoração do ácido e na etapa de titulação da amostra (Tabelas 39 e 40).

Para a fatoração do ácido, foi considerada como fonte de incerteza do tipo A, a grandeza proveniente de uma série de repetições do procedimento de titulação, fornecendo assim, um desvio padrão experimental da média (s). Àquelas associadas aos materiais de medição, por dispensarem a análise estatística de observações repetidas, foram consideradas fontes do tipo B, tais como a pureza do padrão primário, a balança e a bureta de vidro (Tabela 39). Todas as grandezas dessas componentes foram importadas dos certificados de calibração ou do fabricante. No mesmo fundamento, as componentes da incerteza foram determinadas, para a etapa da titulação das amostras (Tabela 40). A resolução dos instrumentos de medição, tais como balança e bureta, foi incluída ao cálculo da incerteza padrão combinada por inferir dúvida na leitura da medição.

Tabela 39. Fontes de incerteza da fatoração do ácido.

u_{1a}	Repetitividade da titulação = $s/\text{raiz}(n)$ s = desvio padrão amostral entre os resultados de fator n = repetições
u_{1b}	Pureza não corrigida do padrão primário (certificado)
u_{2b}	Incerteza da balança (certificado de calibração)
u_{3b}	Erro não corrigido da balança (certificado de calibração)
u_{4b}	Resolução da balança (certificado de calibração)
u_{5b}	Incerteza da bureta de vidro de 25 mL (certificado de calibração)
u_{6b}	Resolução da bureta (certificado de calibração)
u_{7b}	Erro não corrigido da bureta (certificado de calibração)
u_c	Incerteza padrão combinada
U_f	Incerteza padrão expandida (do fator da solução do ácido)

u_a , incerteza padrão tipo A, u_b , incerteza padrão tipo B.

Tabela 40. Fontes de incerteza da titulação da amostra, após destilação.

u_{1a}	Repetitividade atribuída ao laboratório = $s/\text{raiz}(n)$ s = desvio padrão amostral entre as replicatas
u_{1b}	Incerteza do fator da solução do ácido titulante
u_{2b}	Incerteza da micropipeta de 5mL (Certificado de calibração)
u_{3b}	Erro não corrigido da micropipeta de 5mL (Certificado de calibração)
u_{4b}	Resolução da micropipeta de 5mL (Certificado de calibração)
u_{5b}	Incerteza da micropipeta de 1mL (Certificado de calibração)
u_{6b}	Erro não corrigido da micropipeta de 1mL (Certificado de calibração)
u_{7b}	Resolução da micropipeta de 1mL (Certificado de calibração)
u_{8b}	Incerteza da bureta (do volume gasto na titulação da amostra) certificado de calibração
u_{9b}	Resolução da bureta (do volume gasto na titulação da amostra) Certificado de calibração
u_{10b}	Incerteza da bureta (do volume gasto na titulação do branco) Certificado de calibração
u_{11b}	Resolução da bureta (do volume gasto na titulação do branco) Certificado de calibração
u_c	Incerteza padrão combinada
U_r	Incerteza padrão expandida (do método Kjeldahl)

u_a , incerteza padrão tipo A, u_b , incerteza padrão tipo B.

Na sequência, foram calculados os coeficientes de sensibilidade, que para a etapa de fatoração do ácido incluem a massa do padrão primário e o volume do ácido gasto na fatoração; e para a etapa de titulação da amostra, incluem o volume de ácido gasto na titulação da amostra, o volume gasto de ácido na titulação do branco, o volume de dissolução da vacina, o volume de alíquota digerida e o fator do ácido titulante (Tabelas 41 e 42).

Tabela 41. Cálculo dos coeficientes de sensibilidade para a etapa de fatoração do ácido.

Variável	Valor	$\Delta(x,y)$	R_n	Δ_R	C_i	Média C_i
$V_a(mL)$	13,8	0,05	0,9920	-0,0036	-0,072	-0,072
		-0,05	0,9992	0,0036	-0,072	
$m(g)$	0,023	0,001	1,0383	0,0427	42,7273	42,7273
		-0,001	0,9528	-0,0427	42,7273	

m , massa do padrão primário (g); V_a , volume do ácido gasto na fatoração (mL); $\Delta(x,y)$, delta da variável, R_n , novo resultado calculado; Δ_R , delta do resultado (novo); C_i , coeficiente de sensibilidade.

Tabela 42. Cálculo dos coeficientes de sensibilidade para a etapa de titulação da amostra.

Variável	Valor	$\Delta(x,y)$	R_n	Δ_R	C_i	Média C_i
$V_{a,m}(mL)$	0,7	0,05	0,06544	0,004884	0,09767	0,09767
		-0,05	0,05567	-0,004884	0,09767	
$V_{br}(mL)$	0,08	0,05	0,05567	-0,004884	-0,09767	-0,09767
		-0,05	0,06544	0,004884	-0,09767	
F_a	0,99667	0,005	0,06086	0,000304	0,06076	0,06076
		-0,005	0,06025	-0,000304	0,06076	
$V_d (mL)$	5,000	0,005	0,06062	0,000061	0,01211	0,01211
		-0,005	0,06050	-0,000061	0,01211	
$V_{ali} (mL)$	1,000	0,00002	0,06056	-0,000001	-0,06056	-0,06056
		-0,00002	0,06056	0,000001	-0,06056	

V_{am} , volume de ácido gasto na titulação da amostra (mL); V_{br} , volume gasto de ácido na titulação do branco; V_d , volume de dissolução da vacina; V_{ali} , volume de alíquota digerida; F_a , fator do ácido titulante; $\Delta(x,y)$, delta da variável, R_n , novo resultado calculado; Δ_R , delta do resultado (novo); C_i , coeficientes de sensibilidade.

As tabelas 43 e 44 apresentam os resultados do cálculo da incerteza padrão combinada e da incerteza expandida para o método Kjeldahl. A distribuição de probabilidade retangular foi atribuída à resolução da leitura dos equipamentos digitais, tais como balança e bureta digital, por apresentarem uma probabilidade constante na ocorrência dos resultados dentro do intervalo de leitura. De acordo com a recomendação do Guia para a Expressão da Incerteza de Medição - GUM (INMETRO, 2008), o valor do fator de abrangência $k(1,73)$, para um intervalo com nível da confiança $PA(100)$, foi adotado para distribuições retangulares. As vidrarias e equipamentos analógicos, por produzirem valores que possivelmente espera-se a ocorrência no centro do intervalo das leituras, fornecem comportamento de uma distribuição de probabilidade triangular. Foi considerado o valor do fator de abrangência $k(2,45)$, como divisor para distribuições triangulares. Foi atribuída uma distribuição normal, para as incertezas procedentes de condições repetidas. Os dados obtidos por fontes com pouca informação da natureza das medições, como a pureza do padrão primário, foram adotadas para a distribuição de probabilidade como “padrão assumido” (com $PA=68\%$), ou seja, foi considerado um fator divisor igual a 1. Nota-se que a incerteza proveniente da pureza do padrão primário foi a de maior contribuição para a incerteza combinada final, enquanto as incertezas associadas às medições na balança e na bureta foram de menor impacto no cálculo (Tabela 43).

Tabela 43. Estimativa da incerteza da fatoração do ácido.

	$u (\pm)$	P_A (%)	Distribuição de probabilidade		C_i	$u_i \pm$ $P_A=68\%$	v	u_i^4 / v
				divisor		($\mu\text{g/mL}$)		
u_{1a}	0,001775	68	normal	1	1	0,001775	3	3,31E-12
u_{1b}	0,00022g	68	padrão	1	42,727 g ⁻¹	0,009400	infinito	----
u_{2b}	0,0001g	95,45	normal	k =2,07	42,727 g ⁻¹	0,002064	40	4,54E-13
u_{3b}	0g	68	padrão	1	42,727 g ⁻¹	---	infinito	----
u_{4b}	0,00005g	100	retangular	$\sqrt{3}=1,73$	42,727 g ⁻¹	0,001233	infinito	----
u_{5b}	0,03mL	95,45	normal	k =2	-0,072 mL ⁻¹	0,001078	infinito	----
u_{6b}	0,05mL	100	triangular	$\sqrt{6}=2,45$	-0,072 mL ⁻¹	0,001467	infinito	----
u_{7b}	-0,06mL	68	padrão	1	-0,072 mL ⁻¹	0,004313	infinito	----
u_c	0,01092	68	normal					
U_f	0,02184	95,45	normal	k=2				$v_{eff}=3774$

P_A , probabilidade de abrangência; k , fator de abrangência; v , graus de liberdade; v_{eff} , graus de liberdade efetivo; u_i , incerteza padrão; C_i , coeficientes de sensibilidade.

Tabela 44. Estimativa da incerteza da metodologia de referência (Kjeldahl).

	$u (\pm)$	P_A (%)	Distribuição de probabilidade		C_i	$u_i \pm$ ($P_A=68\%$)	v	u_i^4 / v
				Divisor				
u_{1a}	0,00059*	68	normal	1	1	0,000588	9	1E-14
u_{1b}	0,0218	95,45	normal	k =2	0,06076*	0,000662	***	---
u_{2b}	0,003mL	95,45	normal	k =2	0,01211**	1,817E-05	***	---
u_{3b}	0,018mL	68	padrão	1	0,01211**	0,000218	***	---
u_{4b}	0,005mL	100	triangular	$\sqrt{6}=2,45$	0,01211**	0,00002	***	---
u_{5b}	0,00062mL	95,45	normal	k =2,32	-0,06056**	1,618E-05	***	---
u_{6b}	0,00378mL	68	padrão	1	-0,06056**	0,000229	***	---
u_{7b}	0,001mL	100	triangular	$\sqrt{6}=2,45$	-0,06056**	0,00002	***	---
u_{8b}	0,008mL	68	normal	k =2	0,09767**	0,000391	***	---
u_{9b}	0,005mL	100	retangular	$\sqrt{3}=1,73$	0,09767**	0,000282	***	---
u_{10b}	0,008mL	68	normal	k =2	-0,09767**	0,000391	***	---
u_{11b}	0,005mL	100	retangular	$\sqrt{3}=1,73$	-0,09767**	0,000282	***	---
u_c	0,001161*	68	normal					
U_r	0,002345*	95,45	normal	k=2,018				$v_{eff}=137,62$

P_A , nível de confiança ou probabilidade de abrangência; k , fator de abrangência; v , graus de liberdade; u_i , incerteza padrão; C_i , coeficientes de sensibilidade, v_{eff} , graus de liberdade efetivo; (*) em mg/dose; (**) em mg/mL.dose; (***) infinito.

4.4.2 Estimativa da incerteza dos resultados da calibração multivariada

As principais fontes de incerteza associadas à calibração multivariada neste estudo, foram resumidas na tabela 45. De posse da incerteza referente à metodologia analítica do Kjeldahl (0,002345mg/dose), nota-se o resultado da incerteza da calibração multivariada (0,002867mg/dose) está acima dos valores padrões (Tabela 46). A contribuição significativa de 54% da incerteza padrão do método clássico Kjeldahl sobre a soma das incertezas padrão (u_i), fortalece o incremento dos desvios inerentes ao método adotado, para fornecer os valores de referência ao modelo multivariado. Essa avaliação reflete a forte influência da confiabilidade dos valores de referência.

Tabela 4544. Fontes de incerteza do método multivariado.

u_{1a}	Repetitividade das leituras = $s/\text{raiz}(n)$ s = desvio padrão amostral entre as replicatas das leituras
u_{2a}	Posição do frasco = $s/\text{raiz}(n)$
u_{1b}	Incerteza nos valores de referência utilizados na calibração
u_{2b}	Erro da calibração multivariada (RMSEC)
u_c	Incerteza padrão combinada
$U_{\text{calibração}}$	Incerteza padrão expandida

u_a , incerteza padrão tipo A, u_b , incerteza padrão tipo B.

Tabela 46. Estimativa da incerteza para a calibração multivariada.

	$u (\pm)$	P_A (%)	Distribuição de probabilidade	Divisor	C_i	$u_i \pm P_A=68\%$ ($\mu\text{g/mL}$)	v	C_o (%)	u_i^4 / v
u_{1a}	0,0001069	68	normal	1	1	1E-04	9	4,97	1,45E-17
u_{2a}	5,142E-06	68	padrão	1	1	5E-06	3	0,239	2,33E-22
u_{1b}	0,002345*	95,45	normal	k = 2,02	1	0,001	**	53,9	
u_{2b}	0,00088	68	padrão	1	1	9E-04	98	40,9	6,12E-15
u_c	0,001	68	normal	1					
$U_{\text{calibração}}$	0,002867 *	95,45	normal	k = 1,96					$v_{\text{eff}} = 742$

P_A , nível de confiança ou probabilidade de abrangência; k , fator de abrangência; v , graus de liberdade; u_i , incerteza padrão; C_i , coeficientes de sensibilidade, v_{eff} , graus de liberdade efetivo, C_o , contribuição em percentual no resultado da incerteza padrão em relação a soma total das incertezas padrão; (*) em mg/dose; (**) infinito.

4.5 COMPARAÇÃO ENTRE OS MÉTODOS

A tabela 47, ordena de forma resumida, as necessidades e resultados para os dois métodos analíticos discutidos neste projeto. O método clássico do kjeldahl é um método bastante conhecido e utilizado, de um modo geral, mas apresenta algumas dificuldades de procedimento analítico, podendo gerar resultados equivocados. Essas dificuldades podem estar na digestão dos diferentes materiais, em que o tempo de mineralização deve ser bem monitorado em decorrência do uso do catalisador (Se + Hg), que provoca em tempos longos a formação de N livre (FONTANA,1953).

Outra etapa extremamente crítica no procedimento, talvez a principal fonte de variabilidade nos resultados, está na precipitação das proteínas e na separação do sobrenadante. A não observância à essa separação do sobrenadante antes de transferir o precipitado para a digestão, parte dos excipientes que contém nitrogênio em suas fórmulas estruturais, podem ser carregados juntamente com o precipitado, para a determinação de nitrogênio. Isso pode ocasionar grande variabilidade de algumas réplicas, necessitando de um número adequado acima de três, para retirar boas conclusões estatísticas dos resultados. A falta de atenção e o não controle, sobre essas etapas críticas do método, ocasiona a necessidade de repetições de análise.

O uso demasiado de soluções ácidas e básicas, proporciona um grande acúmulo de resíduo químico, e a necessidade de água para o resfriamento da destilação e água para inúmeras lavagens do sistema.

O método por infravermelho próximo (NIR) contorna os problemas ambientais e de procedimento analítico relatado, trazendo agilidade na liberação dos resultados e é uma técnica operacionalmente limpa. Entretanto quando associada à calibração multivariada, exige um aprofundamento multidisciplinar para o seu desenvolvimento e validação. Ao absorver matematicamente a incerteza dos valores padrões, o cuidado com a escolha da técnica de referência é imprescindível.

Tabela 47. Resumo comparativo entre os métodos.

	Método NIR	Método Kjeldahl
Amostra	Não é destruída	Destruída
Reagentes	Nenhum	Àcido bórico P.A Àcido sulfúrico P.A Àcido tricloroacético PA Àcido clorídrico Catalisador (mistura selênica) Verde de bromocresol Hidróxido de sódio P.A Vermelho de metila
Resíduo	Nenhum	Descarte adequado com custo
Equipamentos e Materiais	Espectrômetro NIR	Destilador Digestor Tubos de digestão Centrifuga Tubos falcons de centrifuga Micropipetas/buretas
Utilidades	Nenhuma	Água corrente
Tempo de análise	5 minutos	8 horas
$M(n=10)$	0,04358	0,04410
s	3,27E-04	2,51E-03
s^2	1,07E-07	6,30E-06
$DPR(\%)$	0,75	5,69

s , desvio padrão; M , média; n , número de réplicas; s^2 , variância; $DPR(\%)$, desvio padrão relativo.

5 CONCLUSÃO

A espectroscopia NIR apresentou vantagens sobre o método Kjeldahl usado na rotina, pois permitiu a eliminação do uso de reagentes potencialmente perigosos, como ácido sulfúrico concentrado e hidróxido de sódio a 40%. A análise permitiu uma redução do tempo total de análise para cerca de 3 minutos e eliminou a geração de resíduos. As aplicações foram desenvolvidas em condições reais de análise, sem alterar os protocolos e a rotina convencional do laboratório.

A análise possibilita o reuso das amostras para estudos futuros, como por exemplo estudo de estabilidade da vacina.

O desenho experimental demonstrou que a etapa de escolha dos valores de referência é a fase mais importante de uma modelagem e a confiança que os mesmos, devem proporcionar à calibração multivariada é imprescindível. A aplicação da análise discriminante entre os valores de referência, através do cálculo matemático de Mahalanobis, reavaliou a veracidade dos valores de referência versus o comportamento espectral desse conjunto, tornando a calibração mais robusta em suas previsões.

O método foi validado, sendo os resultados obtidos considerados precisos, exatos e robustos para suas previsões, a um nível aceitável de incerteza para o método. A metodologia empregada harmonizou a validação tradicional univariada, com os requisitos dos métodos multivariados.

Finalmente, este trabalho apresentou a implementação e validação de modelos multivariados, através de uma robusta técnica de espectroscopia de infravermelho próximo, para determinação de nitrogênio proteico. Mostrando-se uma alternativa viável ao método tradicional, e uma possível ferramenta orientativa, para estudos posteriores na área de controle de qualidade, voltadas às exigências da vigilância sanitária.

6 PROPOSTAS FUTURAS

6.1 Monitoramento da calibração multivariada

Para o controle de qualidade, a grande preocupação está em torno da durabilidade do bom desempenho de um método multivariado. Questão não muito abordada na literatura. Dessa forma a avaliação da vida útil dessa calibração é uma das propostas de trabalho futuro desta dissertação, para monitoramento da continuidade do poder de previsão do modelo PLS. Diante de amostras de vacina VFC da rotina do laboratório, através de cartas controles, com base nos erros de previsão e avaliações de exatidão, acompanhando-os por um prazo mais longo. Caso o limite definido seja ultrapassado, o modelo deve ser revisado.

6.2 Linearidade, sensibilidade e limites para calibração multivariada

Em calibrações univariadas, os cálculos das figuras de mérito são mais bem estabelecidos e mais simples de serem abordados, enquanto que para calibrações multivariadas, as pesquisas ainda estão em desenvolvimento e não totalmente definidas e não tão bem fundamentadas. De maneira bastante similar à calibração univariada, as figuras de mérito como exatidão, precisão, robustez e *bias* são estimadas em calibração multivariada. O que não ocorre para parâmetros como linearidade, sensibilidade, razão sinal/ruído, ajuste, seletividade, limites e intervalos de confiança, que seus cálculos são mais complexos, pois somente podem ser estimados mediante o cálculo do sinal analítico líquido (NAS). O procedimento para a estimativa do NAS envolve um conjunto de equações algébricas e manipulações de matrizes, não tão claras nas publicações em que foram abordados até o momento.

Devendo assim, a determinação da sensibilidade e seletividade, mediante o cálculo do NAS, para contemplar totalmente as exigências da categoria da análise de nitrogênio proteico, no controle de qualidade da vacina de febre amarela de 10 doses.

REFERÊNCIAS

ADAMOPOULOS , K. G. GOULA, A. M. Application of near-infrared reflectance spectroscopy in the determination of major components in taramosalata. **Journal of Food Engineering** , p.199–207,2004.

ALMEIDA. F. M. N. **Espectroscopia de Infravermelho Próximo com Transformada de Fourier (FT-NIR) na Caracterização de Farinhas para Alimentação Pueril**. 2009.84 f. Dissertação(Mestrado em Engenharia Biológica)- Universidade técnica de Lisboa, Instituto superior técnico,2009. Disponível em:<[https://fenix.tecnico.ulisboa.pt/downloadFile/395139422094/Tese%20FINA L%20Filipa%20Almeida.pdf](https://fenix.tecnico.ulisboa.pt/downloadFile/395139422094/Tese%20FINA%20Filipa%20Almeida.pdf)>. Acesso em 20 fev.2016.

ALVES , J. C.L.; POPPI, R. J. Quantification of conventional and advanced biofuels contents in diesel fuel blends using near-infrared spectroscopy and multivariate calibration. **Fuel**, n.165, p.379–388, 2016.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS / INSTITUTO NACIONAL DE METROLOGIA- ABNT/INMETRO. **Guia para a Expressão da Incerteza de Medição, GUM**. 3ª ed. brasileira em língua portuguesa, Rio de Janeiro, p.1211, 2003.

ASSOCIATION OF OFFICIAL ANALYTICAL CHEMISTS- AOAC. **Guidelines for collaborative study procedures to validate characteristics of a method of analysis**. Official methods of analysis. 2002. Disponível em:<http://www.aoac.org/imis15_prod/AOAC_Docs/StandardsDevelopment/SLV_Guidelines_Dietary_Supplements.pdf>. Acesso em: 14 mar.2015.

BADDINI, A. L. Q. et al. Determination of total protein in hyperimmune serum samples by near-infrared spectrometry and multivariate calibration. **Anal. Biochem.** v. 397, p.175–180, 2010.

BAZILIO, F. S. **Determinação da migração específica dos plastificantes ftalato de di-(2-etilhexila) e adipato de di-(2-etil-hexila) de filmes flexíveis de PVC para alimentos gordurosos: validação de método e controle sanitário do filme flexível de PVC**.2014. 141 f. Dissertação (Mestrado em Vigilância Sanitária) - Instituto Nacional de Controle de Qualidade em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro, 2014.

BENCHIMOL, J. L(coord). **Febre amarela: a doença e a vacina, uma história inacabada**. Rio de Janeiro: Editora Fiocruz, 2001.

BERTOLLI FILHO, C. For a new history of yellow fever and the yellow fever vaccine in Brazil. **História, ciências, saúde-Manguinhos**, v. 9, seção1, p.213-20, 2002.

BOEIRA, R.C.MAXIMILIANO, V.C. B. **Análise de nitrogênio em lodos de esgoto úmidos. ICTR 2004, Congresso Brasileiro de Ciência e Tecnologia em Resíduos e Desenvolvimento Sustentável**. Costão do Santinho, Florianópolis, S. C. 2004. Disponível: <<https://www.ipen.br/biblioteca/cd/ictr/2004/ARQUIVOS%20PDF/09/09-037.pdf>>. Acesso em 20 mar. 2015.

BOTELHO, B.G. **Desenvolvimento e validação de métodos para análise direta de alimentos usando ferramentas quimiométricas, espectroscopia no infravermelho e imagens digitais**. 2014. 136 f. Tese (Doutorado em Química)- Universidade Federal de Minas Gerais, Departamento de Química, Belo Horizonte. 2014. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/SFSA-9TMH34/tese_bruno_g_botelho_2014.pdf?sequence=1>. Acesso em 20 fev. 2016.

BRAGA, J. W. B.; POPPI, R. J.. Validação de modelos de calibração multivariada: uma aplicação na determinação de pureza polimórfica de carbamazepina por espectroscopia no infravermelho próximo. **Química Nova**, v. 27, n. 6, p.1004-1011, 2004.

BRAMBILLA, K.J.C. et al. **Análise do teor de nitrogênio em tecido vegetal utilizando espectroscopia no infravermelho próximo. V Evento de Iniciação Científica da Embrapa Florestas**. Embrapa. Colombo, PR, 2006.
BRASIL. Lei n.º 8080, de 19 de setembro de 1990. Dispõe sobre as condições para a promoção, proteção e recuperação da saúde, a organização e o funcionamento dos serviços correspondentes e dá outras providências. **Diário Oficial da República Federativa do Brasil**. Brasília, Seção 1, n.182, p.1, 20 set. 1990.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. Consulta Pública n.º 72, de 04 de novembro de 2009. Proposta de resolução que dispõe sobre os procedimentos e condições de realização de estudos de estabilidade para o registro ou alterações pós-registro de produtos biológicos. **Diário Oficial [da República Federativa do Brasil]**. Brasília, Seção 1, n.214, p.78, 10 nov. 2009.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. Resolução n.º 17 de 16 de abril de 2010. Dispõe sobre as Boas Práticas de Fabricação de Medicamentos. **Diário Oficial da República Federativa do Brasil**. Brasília, DF, Seção 1, n.73, p.94, 19 abr. 2010. Disponível em: <http://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2010/res0017_16_04_2010.html>. Acesso em: 18 mar. 2015.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA).

Resolução RDC n. 49, de 23 de novembro de 2010. Aprova a Farmacopéia Brasileira, 5 ed, v.2, p.1363 -1364, e dá outras providências. **Diário Oficial da República Federativa do Brasil**. Brasília, DF, 24 nov. 2010. Disponível em:<http://www.anvisa.gov.br/hotsite/cd_farmacopeia/pdf/volume2.pdf>. Acesso em :30 mar. 2015.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. RESOLUÇÃO-RE Nº 899, de 29 de maio de 2003. Guia para validação de métodos analíticos e bioanalíticos. **Diário Oficial da República Federativa do Brasil**. Brasília, DF, 02 jun. 2003. Disponível em: <http://portal.anvisa.gov.br/wps/wcm/connect/4983b0004745975da005f43fbc4c6735/RE_899_2003_Determina+a+publica%C3%A7%C3%A3o+do+Guia+para+valida%C3%A7%C3%A3o+de+m%C3%A9todos+anal%C3%ADticos+e+bioanal%C3%ADticos.pdf?MOD=AJPERES>. Acesso em:27 mar.2015.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. Consulta Pública nº 129, de 12 de fevereiro de 2016. Guia para validação de métodos analíticos e bioanalíticos. **Diário Oficial da República Federativa do Brasil**. Brasília, DF, 15 Fev. 2016. Disponível em: <http://portal.anvisa.gov.br/wps/content/anvisa+portal/anvisa/regulacao+sanitaria/assuntos+de+interesse/consultas+publicas/assuntos+de+interesse/consultas+publicas+e+m+andamento/20160212129>. Acesso em 25 fev, 2016.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde (SVS). **Manual de Vigilância Epidemiológica da Febre Amarela revisado**. Brasília, DF, 2.ed., 2004.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde (SVS). Febre Amarela: Aspectos epidemiológicos da febre amarela silvestre e a vigilância intensificada durante período de monitoramento 2012-2013. **Boletim Epidemiológico**, Brasil, v.45, n.7, 2014. Disponível em:<<http://portalsaude.saude.gov.br/index.php/o-ministerio/principal/leia-mais-o-ministerio/197-secretaria-svs/11955-boletins-epidemiologicos-arquivos>>. Acesso em: 17mar.2015.

BRITO, N. M. et al. Validação de métodos analíticos: estratégia e discussão. **Pesticidas, R. Ecotoxicol e Meio ambiente**, Curitiba, v.13, p.129-146, jan./dez. 2003. Disponível em:<<http://ojs.c3sl.ufpr.br/ojs/index.php/pesticidas/article/download/3173/2546>>. Acesso em: 16 mar.2015.

BRITTAIN, H. G. Validação de métodos analíticos não cromatográficos. **Pharmaceutical technology**, p.4-9, jun.1998.
BRYAN, C. S.; MOSS, S. W.; KAHN, R.J. Yellow fever in the Americas. **Infectious Disease Clinics of North America**, v. 18, n.2, p. 275–292, Jun. 2004. Disponível

em:<http://ac.els-cdn.com.ez68.periodicos.capes.gov.br/S0891552004000224/1-s2.0-S0891552004000224-main.pdf?_tid=310e20fe-c366-11e4-b949-00000aacb360&acdnt=1425580620_39bf1c207e9e6942e54ade77df1dd15a>. Acesso em: 05 mar. 2015.

CARVALHO, C; MACHADO, R. B; TIMM, L. B. **O direito sanitário brasileiro**. 1. ed. Quartier latin, São Paulo, 2004.

COELHO, D. S.; CORTEZ, J. W.; OLSZEWSKI, N. Variabilidade espacial da resistência mecânica à penetração em vertissolo cultivado com manga no perímetro irrigado de Mandacaru, Juazeiro, Bahia, Brasil, **Rev. Bras. Ciênc. Solo**, Viçosa, v.36, n.3, mai-jun., 2012. Disponível em:<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832012000300007>. Acesso em: 20 mai. 2016.

COX, M.; VASCONCELOS, V.R.; LEITE, E.R. Utilização da técnica da Espectroscopia e Reflectância do InfraVermelho Próximo (NIRS) na determinação do valor nutricional da dieta de ovinos em caatinga. *Revista Científica de Produção Animal*, v.2, n.1, p.31-40, 2000.

CQAQUIMICA. Determinação de nitrogênio e proteína pelos métodos Kjeldahl e DUMAS. **Revista Analytica**, n. 73, out-nov. 2014. Disponível em:<http://www.revistaanalytica.com.br/revista_digital/73/files/assets/downloads/page0052.pdf>. Acesso em 20 mar.2015.

DE SOUZA, A.M. et al. Experimento didático de quimiometria para calibração multivariada na determinação de paracetamol em comprimidos comerciais utilizando espectroscopia no infravermelho próximo: um tutorial, parte ii. **Química Nova**, v. 36, n. 7, p.1057-1065, 2013.

DE SOUZA, M.L et al. Aplicação de espectroscopias raman e infravermelho na identificação e quantificação de plastificantes em filmes comerciais de pvc esticável. **Química nova**, vol. 32, n. 6, p.1452-1456, 2009.

EURACHEM/CITAC WORKING GROUP. **Quantifying Uncertainty in Analytical Measurement**. 3. ed, London, 2012. Disponível em:<www.measurementuncertainty.org>. Acesso em :01 abr.2015.

EURACHEM/CITAC WORKING GROUP. **The fitness for purpose of analytical methods: A laboratory guide to method validation and related topics**. 2a ed., London, 2014. Disponível em:<www.measurementuncertainty.org>. Acesso em :01 abr.2015

FERRAO, et al. Determinação simultânea dos teores de cinza e proteína em farinha de trigo empregando nirr-pls e drift-pls1. **Ciênc. Tecnol. Aliment.**, Campinas, v. 24, n. 3, p. 333-340, jul.-set. 2004.

FERREIRA M.M.C et al. Quimiometria i: Calibração multivariada, um tutorial. **Química nova**, v. 22 iss:5 p. 724 -731. 1999.

FERREIRA NETO, F. **Determinação do teor de diclofenaco de sódio em comprimidos por espectroscopia no infravermelho próximo-NIR com calibração multivariada PLS**. 2012. 62f. Dissertação (Mestrado em Química) - Instituto de Química-Universidade Federal do Rio Grande do Norte, Natal/RN, 2012.

FERREIRA, R. L. **Detecção de micoplasmas por reação em cadeia da polimerase (PCR) em produtos intermediários da vacina contra a febre amarela produzida em Bio-Manguinhos/Fiocruz**. 2007. 95 f. Dissertação (Mestrado Profissional em Tecnologia de Imunobiológicos) – Instituto de Tecnologia em Imunobiológicos, em parceria com o Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, 2007.

FIDENCIO, P.H. Análise de solos por espectroscopia no infravermelho próximo e aplicação de métodos quimiométricos. 2001. 138f. Tese (Doutorado em Ciências-Química Analítica). Universidade Estadual Campinas .Instituto de química. Campinas/SP. 2001.

FILGUEIRAS, P. R. Avaliação de modelos de calibração PLS e SVM na determinação do carbono orgânico do solo por espectroscopia NIR. **Sociedade Brasileira de Química (SBQ) 35ª Reunião Anual da Sociedade Brasileira de Química**, 2012

FINDLAY, G. M. John Williams and the early history of yellow fever. **British medical journal**, v.2, p.474-476, 1948.

FONTANA, P. Considerações sobre a dosagem do nitrogênio pelo método de Kjeldahl. Laboratório de Bioquímica do Instituto Oswaldo Cruz. **Memórias do Instituto de Oswaldo Cruz**, n.51, p.278-288, 1953.

FOUNTAIN, W. et al. Near-infrared spectroscopy for the determination of testosterone in thin-film composites. **Journal of Pharmaceutical and Biomedical Analysis**, v.33, n.2, p.181-189, 2003.

GALVANI, F. ; GAERTNER, E. **Adequação da Metodologia Kjeldahl para**

determinação de Nitrogênio Total e Proteína Bruta. Circular técnica, n.63, Embrapa, Corumbá, MS. p. 1517-1965, maio, 2006.

GARDNER, C. L.; RYMAN, K.D. Yellow Fever: A Reemerging Threat. **Clinics in Laboratory Medicine**, v.30, (1), p. 237–260, mar. 2010.

GARSKE, T. et al. Yellow fever burden estimation: Summary. **WHO**, 2013. Disponível em: <http://www.who.int/csr/disease/yellowfev/YellowFeverBurdenEstimation_Summary2013.pdf>. Acesso em: 17 mar.2015.

GOMES, S. L. et al. **Vacinação em massa contra febre amarela na África. Observatório Epidemiológico, 6ª Semana Epidemiológica** [Publicação Científica do Curso de Bacharelado em Enfermagem do CEUT], 17.ed, 2010. Disponível em: <<http://www.ceut.com.br/observatorio/edicao%2017.pdf>>. Acesso em: 17 mar. 2015.

GONÇALVES, J. L. et al. **Predição da proteína bruta da dieta selecionada por ovinos em pastagem de capim-tanzânia por meio da espectroscopia fecal (NIRS fecal).** VIII congresso nordestino de produção animal, Embrapa, Fortaleza, CE, 2013.

GONÇALVES, E.B.; ALVES, A. P. G.; MARTINS, P.A. Questões críticas em validação de métodos analíticos. **Embrapa, Informação Tecnológica**, Brasília, DF, 2011. Disponível em: <<http://www.alice.cnptia.embrapa.br/alice/bitstream/doc/901074/1/Texto40080911.pdf>>. Acesso em: 14 mar. 2015

GREDILLA A. et al .Non-destructive Spectroscopy combined with chemometrics as a tool for Green Chemical Analysis of environmental samples: A review. **Trends in Analytical Chemistry**, n. 76, p. 30–39, 2016.

GUÉVÉLOU, E; ALLEN S. K. Use of Near Infrared Reflectance Spectroscopy (NIRS) for the rapid compositional analysis of di-, tri-, and tetraploid eastern oysters (Crassostrea virginica). **Aquaculture**, n. 459, p.203-209, 2016.

GUO, Y.; NI Y.; KOKOT, S. Evaluation of chemical components and properties of the jujube fruit using near infrared spectroscopy and chemometric. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, n 153, p. 79-86, 2016.

HATANAKA R. R. **Desenvolvimento, validação, avaliação da incerteza de medição e análise de custos de método baseado em espectroscopia no infravermelho e análise multivariada para previsão de parâmetros de qualidade de amostras de E85.** 2015. 150f. Tese (Doutorado em Química)-Universidade

Estadual Paulista, Instituto de Química, Araraquara, 2015.

HOMMA, A. The Brazilian vaccine manufacturer's perspective and its current status. **Biologicals**, v.37 iss:3, p.1045-1056, 2009.

HOMMA, A. Vacina Febre Amarela: sucessos e desafios. **Curso de vacinas em saúde pública**, Ensp/Fiocruz, 15-19 jul. 2013.

HORWITZ, W.; ALBERT, R. The Horwitz Ratio (HorRat): A Useful Index of Method Performance with Respect to Precision. **Journal of AOAC International**, v. 89, n. 4, p. 1095-1109, 2006.

INMETRO. JCGM 100:2008, Grupo de trabalho para tradução do GUM 2008. BARATTO, A.C. (Coord.) et al. **Avaliação de dados de medição — Guia para a expressão de incerteza de medição**. Primeira edição especial, set. 2008. Disponível em: <www.bipm.org>. Acesso em 30.03.2015.

INSTITUTO DE TECNOLOGIA EM IMUNOBIOLOGICOS. FIOCRUZ - BIOMANGUINHOS. Critérios para aprovação da vacina de febre amarela(atenuada) 05,10 e 50 doses. **Documento Interno: Instrução de trabalho n.0359**. Rev. 01, Rio de Janeiro, 2013b.

INSTITUTO DE TECNOLOGIA EM IMUNOBIOLOGICOS. FIOCRUZ - BIOMANGUINHOS. **Febre amarela: sintomas, prevenção e transmissão**. Rio de Janeiro, 2014a. Disponível em: <<http://www.bio.fiocruz.br/index.php/febre-amarela-sintomas-transmissao-e-prevencao>>. Acesso em: 20 mar. 2015.

INSTITUTO DE TECNOLOGIA EM IMUNOBIOLOGICOS. FIOCRUZ - BIOMANGUINHOS. Procedimento para revisão por inspeção visual manual (IVM) de produtos liofilizados. **Documento Interno: Instrução de trabalho n. 7824**. Rev. 01, Rio de Janeiro, 2014b.

INSTITUTO DE TECNOLOGIA EM IMUNOBIOLOGICOS. FIOCRUZ - BIOMANGUINHOS. Determinação do nitrogênio proteico (kjeldahl). **Documento Interno: Instrução de trabalho n. 0060**. Rev. 04, Rio de Janeiro, 2013a.

INSTITUTO NACIONAL DE CONTROLE DE QUALIDADE. FIOCRUZ - INCQS. Validação de metodologias com determinação através de curva analítica linear. **POP nº 65.3120.126**. Rev. 02. Rio de Janeiro, 2012.

INSTITUTO NACIONAL DE METROLOGIA - INMETRO. Orientações sobre Validação de Métodos de Ensaio Químicos. **DOQCGCRE-008**. Normatização e

Qualidade Industrial. Rev. 04, jul. 2011. Disponível em: <http://www.inmetro.gov.br/sidoq/arquivos/cgcre/doq/doq_cgcre-8_04.pdf>. Acesso em: 16 mar. 2015.

INSTITUTO NACIONAL DE METROLOGIA - INMETRO. **Vocabulário Internacional de Metrologia: Conceitos Fundamentais e Gerais e Termos Associados (VIM 2012)**. 1.ed, Rio de Janeiro, 2012.

INSTITUTO PORTUGUÊS DE ACREDITAÇÃO - IPAC. **Guia para a Quantificação de Incerteza em Ensaios Químicos: OGC007**, Caparica, 2007. Disponível em: <<http://www.ipac.pt/docs/publicdocs/regras/ogc007.pdf>> Acesso em: 02 jun. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q1A (R2): Stability Testing of New Drug Substances and Products**. v. 4, 2003a. Disponível em:<http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1A_R2/Step4/Q1A_R2_Guideline.pdf>. Acesso em :14 mar. 2015

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q1B: Stability Testing: Photostability Testing of New Drug Substances and Products**. v.4, 1996a. Disponível em:<<http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073373.pdf>>. Acesso em :14 mar. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q1C: Stability Testing for New Dosage Forms**. v. 4, 1996b. Disponível em:<http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1C/Step4/Q1C_Guideline.pdf>. Acesso em :14 mar. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q5C: Quality of Biotechnological Products: Stability Testing of Biotechnological/Biological Products**. v.4, 1995. Disponível em:<http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q5C/Step4/Q5C_Guideline.pdf>. Acesso em: 14 mar. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Validation of Analytical Procedures: Text and Methodology**. Geneva,

2005. Disponível

em:<http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf>. Acesso em: 14 mar. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q1D: Bracketing and Matrixing Desings for Stability Testing of New Drugs Substances and Products**. v.4, 2003d. Disponível em:<<http://www.fda.gov/downloads/Drugs/.../Guidances/ucm073379.pdf>>. Acesso em: 14 mar. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q1E: Evaluation of Stability Data**. v. 4, 2003b. Disponível em:<http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1E/Step4/Q1E_Guideline.pdf>. Acesso em: 14 mar. 2015.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE - ICH. **Q1F: Stability Data Package for Registration Applications in Climatic Zones III and IV**. v. 4; 2003c.

JENTES, E.S.et al. The revised global yellow fever risk map and recommendations for vaccination, 2010: Consensus of the informal who working group on geographic risk for yellow fever. **Lancet Infect**, v. 11, p. 622-632, 2011.

LACERDA, A. L.; MELLO, M. T. V. B. de: Produzindo um imunizante: imagens da produção da vacina contra a febre amarela. **História, Ciências, Saúde**. Manguinhos, v. 10, s. 2, p. 537-71, 2003.

LANÇAS, F. M. **Validação de métodos cromatográficos de análise**. São Carlos, SP: RIMA, 2004.

LI, W.; QU, H. Determination of total organic carbon and soluble solids contents in Tanreqing injection intermediates with NIR spectroscopy and chemometrics .**Chemometrics and Intelligent Laboratory Systems** ,2016.

LI, C. et al. Geographical traceability of Marsdenia tenacissima by Fourier transform infrared spectroscopy and chemometrics . **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, n.152, p. 391-396, 2016.

Li, W.; Qu, H. Determination of total organic carbon and soluble solids contents in Tanreqing injection intermediates with NIR spectroscopy and chemometrics . **Chemometrics and Intelligent Laboratory Systems**. 2016.

LIMA, K. M. G, et al. Determinação não destrutiva do nitrogênio total em plantas por espectroscopia de reflectância difusa no infravermelho próximo. **Química Nova**, v. 31, n. 3, p. 700-703, 2008.

LIMA, KÁSSIO M. G. et al. Non-destructive determination of total nitrogen in plants by diffuse reflectance near infrared spectroscopy. **Química Nova**, v.31, n.3, p. 700-703, 2008.

LIU, X., et al. Statistical-Based Monitoring of Multivariate Non-Gaussian Systems. **AIChE Journal**, v. 54, n. 9, p. 2379-2391, 2008. Disponível em: <<https://www.cheric.org/research/tech/periodicals/view.php?seq=754864>>. Acesso em :20 mai. 2016

LUYPAERT, J.; MASSART, D.L.; VANDER HEYDEN, D.L. Near-infrared spectroscopy applications in pharmaceutical analysis. **Talanta** 72, p. 865 883, 2007.

MAAG, G. B; ALBANO, F.M; MESQUITA, R. M. P. Proposta para a estimativa da incerteza de medição e garantia da qualidade do teste de germinação. **Informativo ABRATES**, v.23, n.3, 2013.

MALUF, D.F. et al. Determinação de sibutramina em formas farmacêuticas através de espectroscopia no infravermelho com refletância difusa e métodos de calibração multivariada. **Química nova**, v. 33, n. 3, p.649-652, 2010.

MONATH, T. P. et al. Yellow fever Vaccine. In: STANLEY, A.; PLOTKIN, W.A.; ORENSTEIN; PAULA, A.(org). **Vaccines**, Seção 2, cap.38, 6. ed, 2013. Disponível em: <https://expertconsult.inkling.com/store/book/vaccines-plotkin-orensteinoffit6th/?ia_search_q=vaccines&ia_view_type=all&ia_result_type=book>. Acesso em: 06 mar. 2015.

MONATH, T. P.; VASCONCELOS, P. F.C. Yellow fever .**Journal of Clinical Virology**, n. 64 p.160–173,2015.

MONTEIRO, A. D. **Estudo de aspectos estatísticos da espectroscopia na região do infravermelho próximo**. 2011. 161 f. Dissertação (Mestrado em Engenharia Química, COPPE) – Universidade Federal do Rio de Janeiro, RJ, 2011.

MORGANO, M. A. et al. Determinação de proteína em café cru por espectroscopia nir e regressão pls. **Ciência e Tecnologia de Alimentos**, Campinas, n 25,p. 25-31, jan.-mar. 2005.

NARDO, A. E. et al. Comparação de dois métodos analíticos para determinação de N total em grãos de soja. **VI Jornada Acadêmica da Embrapa Soja**, Embrapa Soja. Documentos, 328, p.56 -59. Jul. 2011. Disponível em:<https://www.embrapa.br/documents/1355202/1529289/Jornada_Acad%C3%AAmica_2011.pdf/3dd2b3f4-42fd-4bb5-a56e-b58f2f442a59>. Acesso em 20 mar.2015.

NASCIMENTO, A. P.R. **Quantificação dos taninos dos vinhos por transformada de fourier dos espectros no infravermelho médio (FTIR)**. 2011. 70f. Dissertação (Mestrado em Viticultura e Enologia) - Faculdade de Ciências, Universidade do Porto,Lisboa, 2011.

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

NETTO, E. J. R. et al. Controle da qualidade de vacinas contra febre amarela utilizadas no Programa Nacional de Imunizações do Brasil. **Rev. Inst. Adolfo Lutz (Impr.)**, v.70, n.4, p. 606-612, 2011.

OLIVEIRA, F.C.C. **Modelo de calibração multivariada associada a espectroscopia vibracional para análise de mistura de diesel-óleos vegetais**. 2006 Dissertação (Mestrado em Química) - Instituto de Química da universidade de Brasília,DF. 2006.

OLIVEIRA, R. R. **Aplicação da espectroscopia de infravermelho próximo para a determinação do caseinomacropéptido em leite UAT** .2010. 63 f. Dissertação (Mestrado em veterinária) – Universidade Federal de Goiás, Escola de Veterinária, 2010.

OLIVIERI, A. C. et al. Uncertainty estimation and figures of merit for multivariate calibration. **Pure and Applied Chemistry**, v. 78, n. 3, p. 633–661, 2006.Disponível em:<http://www.degruyter.com/view/j/pac.2006.78.issue-3/pac200678030633/pac200678030633.xml>>. Acesso em 20 mai.2016.

PASQUINI, C. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, São Paulo, v.14, n.2, p.198-219, mar./abr. 2003.

PATTERSON, K. D. Yellow fever epidemics and mortality in the united states. **Sot. Sri. Med.** v. 34, n. 8, p. 855-865, 1992.

PEREIRA, L. S.A. **Transferência de calibração de misturas de pós para comprimidos intactos: um novo uso em análise farmacêutica para uma ferramenta quimiométrica conhecida.** 2013. 51 f. Dissertação (Mestrado em Química Analítica) - Universidade Federal de Minas Gerais, Departamento de Química, Belo Horizonte, 2013. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/SFSA-9C6SE6/disserta__o_20_08_13__leandro_.pdf?sequence=1>. Acesso em 20 fev.2016.

PORTOLÉS, J. M. **Tratamiento numérico de los datos en el análisis cuantitativo por espectrometría vibracional.** 2007. 264f. Tese (Doutorado em Química) - Departamento de Química Analítica, Faculdade de Química, Universidade de Valencia, Valencia, 2007.

POSSAS, J. L. S. **Pré-validação do ensaio imunoenzimático de quantificação do teor de ovoalbumina na vacina contra febre amarela.** 2012. 48 f. Trabalho de Conclusão de Curso (Especialização em Vigilância Sanitária) – Instituto Nacional de Controle de Qualidade em Saúde, Fundação Oswaldo Cruz. Rio de Janeiro, 2012.

QU, N. et. al. Nondestructive determination of compound amoxicillin powder by NIR spectroscopy with the aid of chemometrics. **Spectrochimica Acta**, v. 70 p.1146–1151, 2008.

RAMOS FILHO, C. F. **Com a febre amarela não se brinca.** Jornal do Brasil. 2016. Disponível em:<http://www.jb.com.br/sociedade_aberta/noticias/2016/02/10/com-a-febre-amarela-nao-se-brinca/>. Acesso em: 1 jun. 2016

REDE METROLÓGICA. Incerteza de medição: guia prático do avaliador de laboratórios. **Procedimento do sistema de gestão da qualidade.** Porto Alegre, RS, Brasil, Rev.05, Abr.2013.

REICH, G. T. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. **Advanced Drug Delivery Reviews**, 57 p. 1109–1143, 2005.

REISDÖRFER, F. C. **Estudo visando à extensão do prazo de validade da vacina febre amarela (atenuada) 05 e 10 doses.** 2011. 112f. Dissertação (Mestrado em Tecnologia de Imunobiológicos). Instituto de Tecnologia em Imunobiológicos – Bio-Manguinhos, Manguinhos, RJ, 2011.

ROGGO, Y. et al. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies . *Journal of Pharmaceutical and Biomedical Analysis*, n.44, p.683–700, 2007.

SANTOS A. F., et al. Monitoring and Control of Polymerization Reactors using NIR Spectroscopy. **Polymer-Plastics Technology and Engineering**, v. 44, p. 1 – 61, 2005.

SANTOS, A. P. **Espectroscopia de infravermelho próximo em análises de solos e plantas**. 2011. 63 f. Dissertação (Mestrado em Agronomia/Solos. Universidade Federal de Uberlândia, 2011.

SANTOS, E. O. et al. Determination of degree of polymerization of insulating paper using near infrared spectroscopy and multivariate calibration. **Vibrational Spectroscopy**, n.52, p.154-157, 2010.

SANTOS, G. A; PEREIRA, A. B; KORNDÖRFER, G. H. Uso do Sistema de Análises por Infravermelho Próximo (NIR) para análises de matéria orgânica e fração argila em solos e teores foliares de silício e nitrogênio em cana-de-açúcar. **Bioscience Journal**, Uberlândia, v. 26, n. 1, p. 100-108, Jan./Fev. 2010.

SANTOS, M. F. **Desenvolvimento e Validação de Métodos de Espectroscopia no Infravermelho Próximo e Médio para Caracterização de Lamas de ETAR para Uso Agrícola**. 2007. 87f. Dissertação (Mestrado em Engenharia Biológica) - Instituto Superior Técnico, Universidade de Lisboa, Lisboa,2007.

SARTORIO, S.D. **Aplicações de técnica de análise multivariada em experimentos agropecuários usando o software R** . 2008. 130 f. Dissertação (Mestrado em Agronomia)- **Universidade de São Paulo**, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2008.

SAUER, E. **Desenvolvimento de metodologias multivariadas para análise de queijos por espectroscopia drift**. 2007. 157f. Tese (Doutorado em Química). Setor de Ciências Exatas, Universidade Federal do Paraná, 2007.

SCHÖNBRODT, T. et al. NIR spectroscopy—a non-destructive analytical tool for protein quantification within lipid implants. **Journal of Controlled Release** 114, p.261–267, 2006.

SENA, M. M. et al. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, 23(4), p.547,1999.

SHEFF, B. Microbe of the month: Yellow fever. **Nursing**, v.35, n.7, p.75, 2005.

SHENK, J.S.; WORKMAN, J.J.; WESTERHAUS, M.O. Application of NIR spectroscopy to agricultural products. In: BURNS, D.A.; CIURCZAK, E.W. **Handbook of near-infrared analysis**. London: New York, ed.3 cap.17, , p. 356-360,2008

SILVA, V.H.et al. Determination of naphtha composition by near infrared spectroscopy and multivariate regression to control steam cracker processes. **Fuel Processing Technology**, n.131, p. 230–237, 2015.

SILVA, B. G.; ALVES, J. R. Reação adversa à vacinação. Fleury Medicina e Saúde. 22 de nov. 2011.Disponível em:<<http://www.fleury.com.br/medicos/educacao-medica/artigos/Pages/reacao-adversa-a-vacinacao.aspx>>. Acesso em: 20 jan.2015.

SILVERSTEIN, R. M.; WEBSTER, F. X. **Identificação espectrométrica de Compostos orgânicos**. 6. ed. Rio de Janeiro: LTC-Livros Técnicos e Científicos, 2006.

SKOOG, HOLLER, NIEMAN. **Princípios de Análises Instrumental**. Analítica, tradução da 8ª edição norte-americana, editora Thomson, p. 163 – 204,2006.

SOUZA, J. S. FERRÃO, M. F. Aplicações da espectroscopia no infravermelho no controle de qualidade de medicamentos contendo diclofenaco de potássio. Parte I : Dosagem por regressão multivariada. **Brazilian Journal of Pharmaceutical Sciences**, v. 42, n. 3, jul.set., 2006.

SOUZA, S. V. C. **Procedimento para validação intralaboratorial de métodos de ensaio: delineamento e aplicabilidade em análises de alimentos**. 2007. 296 f.Tese (Doutorado em Ciência de Alimentos) - Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

STAPLES, J. E; MONATH, T. P. Yellow fever: 100 years of discovery. **Jama**, 300(8), p.960-962, 27 Aug.2008. Disponível em: <<http://jama.jamanetwork.com/article.aspx?articleid=182442>>. Acesso em: 04 mar.2015.

STUART, B. **Infrared spectroscopy: fundamentals and applications**. Wiley, ed.1st , p. 244, 2004.

SWARTZ, M. E; KRULL, I.S. Validação de métodos cromatográficos. **Pharmaceutical technology**, p.12-20, jun.1998.

TEÓFILO, R.F. **Métodos Quimiométricos: Uma Visão Geral - Conceitos básicos de quimiometria**, Universidade Federal de Viçosa, Viçosa, Vol. 1, 2013.

THOMPSON, M.; ELLISON, S.L.R.; WOOD, R. Harmonized guidelines for a singlelaboratory validation of methods of analysis. **Pure Applied Chemistry**, v.74, p. 835-855, 2002.

UNIÃO INTERNACIONAL DE QUÍMICA PURA E APLICADA - IUPAC.
Compendium of Chemical Terminology - Gold Book, Versão 2.3.3, Oxford, 2014.
Disponível em: < <http://goldbook.iupac.org/PDF/goldbook.pdf>>. Acesso em: 16 jun. 2014.

VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. Estado da arte de figuras de mérito em calibração multivariada. **Química Nova**, v. 32, n. 5, p.1278-1287.2009

VALENTINI S. R.; SOMMER W. A.; MATIOLI, G. Validação de métodos analíticos. **Arq Mudi**. 11(2), p.26-31, 2007.

VASCONCELOS, P. F. C. Febre amarela, Revisão. **Revista da Sociedade Brasileira de Medicina Tropical**. v.36(2), p.275-293, 2003.

VASCONCELOS, P. F. C. Febre amarela: reflexões sobre a doença, as perspectivas para o século XXI e o risco da reurbanização. **Revista Brasileira de Epidemiologia**, v. 5, n 2, p.244-258, 2002.

VAZ, B. G.; ROCHA, W F DE C.; NOGUEIRA, R. Validation of model of multivariate calibration: an application to the determination of biodiesel blend levels in diesel by near-infrared spectroscopy. **Journal of Chemometrics**, v. 26, n. 8-9, p. 456-461, 2012.

VIANA,L.C. Uso da espectroscopia no infravermelho próximo para a predição de propriedades tecnológicas da madeira de eucalyptus.2008.111f. Dissertação (Mestrado em Engenharia Florestal) Universidade federal de lavras, Lavras/MG, Brasil, 2008.

WORLD HEALTH ORGANIZATION (WHO). Expert Committee on Biological Standardization Sixty-first report. **Recommendations to assure the quality, safety and efficacy of live attenuated yellow fever vaccines**. Annex 5. WHO Technical Report Series n. 978, 2013. Disponível em:<http://www.who.int/biologicals/areas/vaccines/TRS_978_Annex_5.pdf?ua=1>. Acesso em : 10 de jul. 2015.

WORLD HEALTH ORGANIZATION (WHO). **Has the pattern of yellow fever in Africa changed.** Disponível em: <<http://www.who.int/emergencies/yellow-fever/mediacentre/qa/en/>>. Acesso em : 1 jun. 2016.

WORLD HEALTH ORGANIZATION (WHO). **Health Topics. Yellow Fever.** Fact sheet n° 100 on yellow fever .Geneva; 2014. Disponível em: <<http://www.who.int/mediacentre/factsheets/fs100/en/index.html>>. Acesso em : 30 mar. 2015.

WORLD HEALTH ORGANIZATION (WHO). **Yellow fever situation report.** Disponível em: <<http://www.who.int/emergencies/yellow-fever/situation-reports/2-june-2016/en/>>. Acesso em: 6 jun. 2016.

XIAO-XU, M., et al. Quantification of albumin in urine using preconcentration and near-infrared diffuse reflectance spectroscopy. **Chinese Chemical Letters**, 2016.

APÊNDICE A- PROCEDIMENTO OPERACIONAL PADRÃO PARA A DETERMINAÇÃO DE NITROGÊNIO PROTEICO EM VACINAS DE FEBRE AMARELA 10 DOSES POR NIR COM CALIBRAÇÃO MULTIVARIADA (POP).

1-ALTERAÇÕES

Revisão	Alterações
00	Emissão inicial.

2- OBJETIVO

Fixar a metodologia de previsão para análise de nitrogênio proteico em vacina de febre amarela 10 doses em frascos âmbar.

3- CAMPO DE APLICAÇÃO

Aplica-se à rotina de controle de qualidade utilizada para análise de nitrogênio proteico em vacina de febre amarela 10 doses, pela (SEPFI).

4- REFERÊNCIAS

Manual de instruções do equipamento Antaris II- FT-NIR Analyser.

Dissertação.Desenvolvimento e Validação de método não destrutivo de nitrogênio proteico em vacina de febre amarela por espectroscopia de infravermelho próximo com calibração multivariada / Vanessa Emídio Dabkiewicz.-Rio de Janeiro: INCQS/FIOCRUZ, 2016.

5- DEFINIÇÕES

Não aplicável.

6- SIGLAS

VFC- Vacina de febre amarela 10 doses

SEPFI- Seção de produto final

DEQUA-Departamento de controle de qualidade

LAFIQ-Laboratório físico químico

PLS- *Partial Least Squares* (Regressão por mínimos quadrados parciais)

MSC- *Multiplicative Scatter Correction* (Correção Multiplicativa de espalhamento de Sinal)

CM- *Mean centering* (Método centrar na média)

7- RESPONSABILIDADES

7.1- Do SEPFI

Receber do almoxarifado os frascos do lote de vacina VFC a serem analisados.

Registrar, em caderno específico, a entrada das amostras na seção.

Executar a análise.

Anotar, em caderno próprio, todas as informações necessárias, referentes ao produto e ao teste realizado.

7.2- Da chefia do SEPFI

Avaliar os resultados, encaminhar laudo, aprovando ou reprovando o lote analisado, conforme a especificação.

8- PRÉ-REQUISITOS

8.1- Características

Os frascos são fornecidos pela amostragem, acondicionados em câmara fria 2 a 8°C, até o momento da análise. Os líofilos das vacinas que representarão o lote, deverão estar íntegros e com pastilha uniforme por todo o frasco âmbar.

8.2- Especificações

O teor de Nitrogênio Proteico nas Vacinas 10 doses deve ser $\leq 0,25 \text{ mg.dose}^{-1}$.

Para a calibração multivariada o teor deve estar na faixa de trabalho de calibração de 0,04 a 0,06 mg.dose^{-1} .

8.3- Reagentes

Não aplicável

8.4- Materiais

Não aplicável

8.5- Equipamento e utilidades

- Espectrômetro NIR- TAG: ESP052
- Marca: Thermo Scientific
- Modelo: Antaris II FT-NIR Analyser
- Patrimônio do Equipamento: FO26826
- Utilidade:

8.6- Softwares

- Result Operation.
- TQ Analyst.

8.7- Condições gerais do equipamento

Condições de qualificação em validade.

Condições dos padrões internos em validade.

Coleta de espectros com background

Resolução: 4 cm^{-1}

Nº de scans: 32

8.8- Parâmetros específicos da calibração

Verificar no programa TQ Analyst as seguintes descrições:

Tipo de análise: PLS

Tipo de tratamento: MSC

Unidade: mg.dose^{-1}

Faixa de aplicação: 0,03000-0,07000

Região de trabalho: $4000\text{-}5184\text{cm}^{-1}$

Normalização: CM

Fatores:1

9- PROCEDIMENTOS

9.1- Inserindo o Set de calibração

- 1- Abrir o programa TQ Analyst.
- 2- Adicionar o conjunto de espectros selecionados à calibração através de estudo de modelagem.
- 3- Adicionar os valores de referência respectivos aos espectros do set de calibração.
- 4-Verificar os parâmetros da calibração

9.2- Condicionamento: checando o instrumento antes da leitura.

- 1- Abrir o programa Result Operation.
- 2- Na barra de ferramentas, em “maintenance”, clicar em “Instrument check”. Uma janela se abrirá e em “sample type”, escolher “transmission sample” e depois em “OK” (Você pode escolher em arquivo report”, caso queira o registro do status do equipamento, no momento da leitura).
- 3- Ainda em “maintenance”, clicar em “instrument status”. Selecionar todos os ícones do “Report Options” e, em seguida “OK”.
- 4- Caso o relatório final dessa verificação indique a necessidade de alinhamento, retornar à barra de ferramentas e em “maintenance”, escolher “Align Instrument” e, em seguida, clicar em “iniciar”.

9.3. Coleta de espectros, através do programa Result Operation.

1-Clique no ícone “Select workflow” e selecione o workflow desejado (neste caso, nomeado como “Determinação de N-proteico em VFC”). Clique em Ok e em seguida no ícone GO para iniciar a análise. Na caixa de diálogo que será aberta digite o número de amostras que serão analisadas (Obs.: Para cada amostra o método executará automaticamente três leituras e, para cada leitura, o programa perguntará o nome do arquivo. Ao final das três leituras, perguntará o nome do arquivo que será a média das três leituras, que deverá ser nomeado como **média lote n^oxxxxxxxxxx**).

NOTAS: Caso na lista do ícone “Select workflow”, não esteja o workflow desejado, em “Administration”, selecione “Manager Workflow”, uma janela se abrirá, adicione o

workflow desejado (neste caso, nomeado como “Determinação de N-proteico em VFC”, clique em “Add” em seguida. Clique em OK.

Os arquivos dos espectros coletados estarão armazenado automaticamente no seguinte caminho: O(C:)/RESULTData/Archive/ Nitrogenio proteico VFC.

9.4. Para quantificar uma amostra com espectros já obtidos.

1. **Abra o programa TQ Analyst.** Em “Files” na barra de ferramentas, selecione o método que deseja executar clicando janela de método (Neste caso o método nomeado “ Metodo PLS otimizado para N-proteico em VFC) ou escolher o nome do método no menu.

2. Clique no ícone “Quantificar” na barra de ferramentas do Analista TQ. A caixa de diálogo “Abrir Spectrum” será exibida.

3. Selecione um arquivo de dados espectrais e escolha OK.

4. Clique nas setas na barra de rolagem ou mover a caixa de rolagem para trazer cada parte da tabela de resultados.

5. Assim que os resultados forem salvos, escolha OK para fechar .

NOTAS:

TQ Analyst começa a analisar o espectro selecionado usando o método activo. Dependendo de como os parâmetros do método são definidos, você pode ser solicitado a inserir as informações necessárias, tais como um valor de caminho ótico. Siga as instruções que aparecem na tela.

Quando a análise for concluída, uma caixa de diálogo aparece mostrando os resultados. Se o método activo é para análise quantitativa (veja o parâmetro Analysis Type), os resultados indicam as concentrações dos componentes especificados em cada amostra.

Se o método é projetado para classificar o espectro da amostra desconhecida, com base em um conjunto de espectros conhecidos, os nomes e os valores dos espectros relatados serão fornecidos. A altura do pico ou área do pico podem ser relatadas se o método está configurado para fazer medições simples dos espectros que são quantificados. Dependendo de como os parâmetros na guia “Relatório” estiverem definidos, os resultados podem conter informações adicionais, tais como os títulos e nomes do método de arquivo e os espectros que são quantificados.

Se um problema ocorreu durante a análise, os dados podem ser impresso com os resultados da análise. Os parâmetros que produzem os avisos também estão incluídos na guia Relatório. Pressione o ícone “Explicar” quando a guia “Relatório” está selecionado para obter instruções sobre como interpretar dados.

É possível imprimir o conteúdo da caixa de diálogo “Quantificar”, clicando no ícone “Imprimir” na caixa de diálogo.

9.5. Para quantificar várias amostras com espectros já obtidos.

1- **Abra o programa TQ Analyst.** Em “file” na barra de ferramenta, selecione o método que deseja executar clicando na janela de método ou escolher o nome do método no menu.

2. Escolha “Resumo múltipla” (multiple summary) a partir do menu “Diagnostics”. A caixa de diálogo “Abrir” será exibida.

3. Para selecionar vários arquivos espectrais a serem quantificado, defina “Arquivos” do tipo como “.SPA/SPG.” Em seguida, selecione os arquivos de dados espectrais que pretende-se quantificar e escolha OK. Você pode precisar alterar diretórios ou unidades para localizá-los. Para selecionar mais de um arquivo, selecione o primeiro arquivo que você deseja incluir, em seguida, mantenha pressionada a tecla “Control” OU “Shift” enquanto clica nos arquivos adicionais.

4. Escolha OK para iniciar a análise. Salve o arquivo com os resultados, que foi apresentado.

NOTAS:

Se você quiser selecionar um grupo espectral, definir a caixa Arquivos do tipo lista de ".SPG." Em seguida, selecione o arquivo de grupo espectral que pretende quantificar e escolha OK.

Se você quiser quantificar uma série de arquivos espectrais que estão listados em um arquivo de entrada, altere a caixa arquivos do tipo lista para "TXT" (para um bloco de notas ou gravação de arquivos) ou "XLS" (para um Excel Arquivo). Em seguida, selecione o arquivo de entrada que contém a lista de nomes de arquivo.

O arquivo de entrada deve ser formatado corretamente para uso com Multiply Quantificar e Resumo múltipla.

Quando a análise estiver concluída, os resultados são exibidos em uma janela do aplicativo. Se TQ Analyst localiza o aplicativo de planilhas Microsoft Excel em seu computador, os dados são exibidos em uma janela do aplicativo Excel. Se o Excel não for encontrado, os dados aparecem em uma janela para um aplicativo de processamento de texto, como o bloco de notas ou escrever. Os resultados são formatados usando as configurações para os parâmetros do relatório (veja a guia Report) no método atual.

Se você fechar a janela do resultado, os dados serão salvos em um arquivo.

Se você usou múltipla Quantify para analisar as amostras, os dados são salvos em QUANTIFY.TXT (ou QUANTIFY.XLS). Se você usou Resumo múltipla para analisar as amostras, os dados são salvos em summary.txt (ou SUMMARY.XLS).

Atenção: O conteúdo do arquivo QUANTIFY.xxx (ou SUMMARY.xxx) serão substituídas na próxima vez que você usar vários "quantificar" ou "Resumo múltipla". Se você quiser manter seus resultados de análise, use o comando "Salvar como" para salvá-los com um novo nome de arquivo.

10- RELAÇÃO DE CÓPIAS DISTRIBUÍDAS

DEQUA, LAFIQ.

10.1- UO incluídas nesta revisão

Não aplicável.

10.2- UO excluídas nesta revisão

Não aplicável.

11- ANEXOS

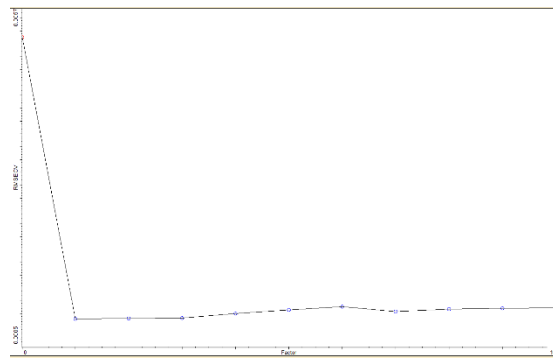
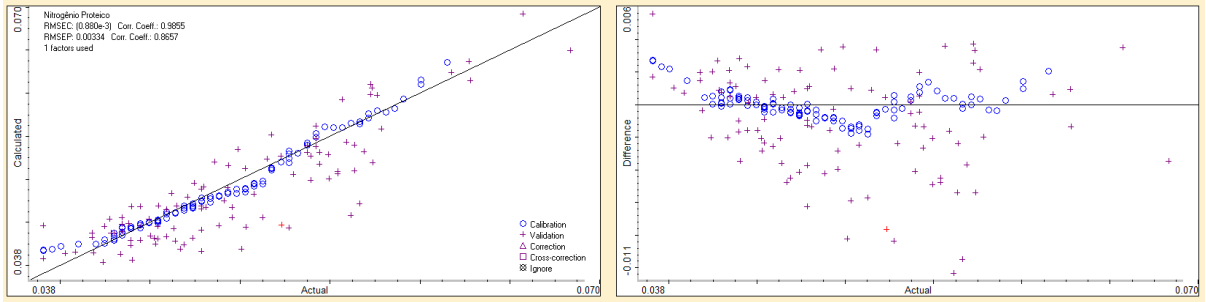
Não aplicável.

12- FOLHA DE APROVAÇÃO

NOME	ASSINATURA
ELABORADO	<hr/>
VERIFICADO	<hr/>
VERIFICADO	<hr/>
APROVADO	<hr/>

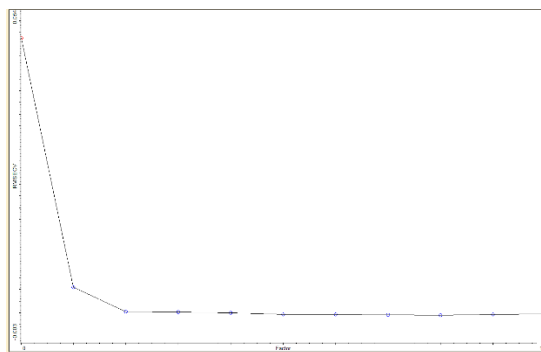
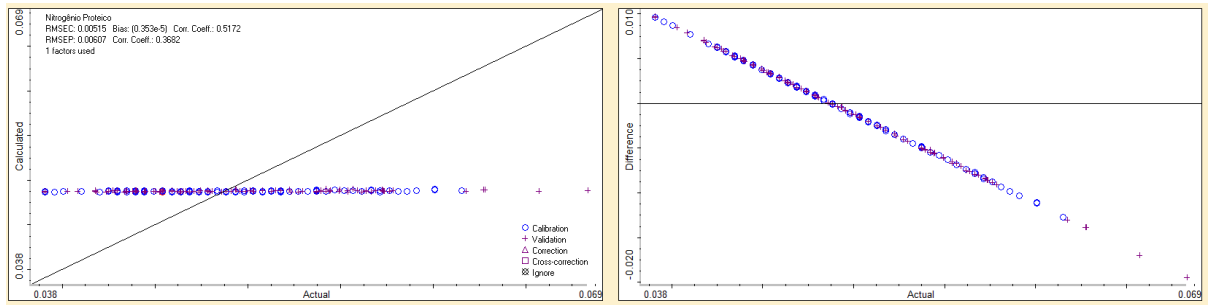
ANEXO A- ESPECTROS NIR.

PLS/MSC/CM 1FATOR



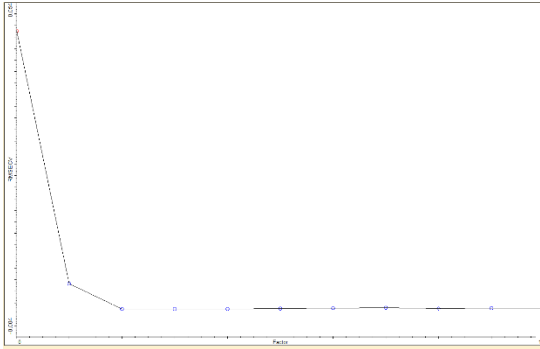
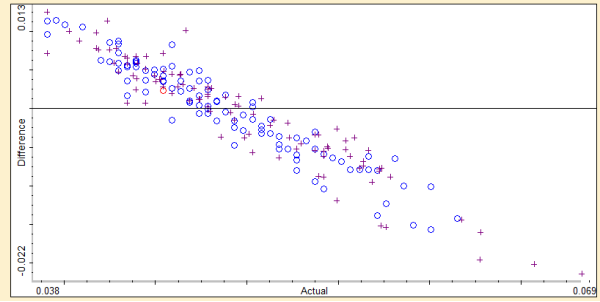
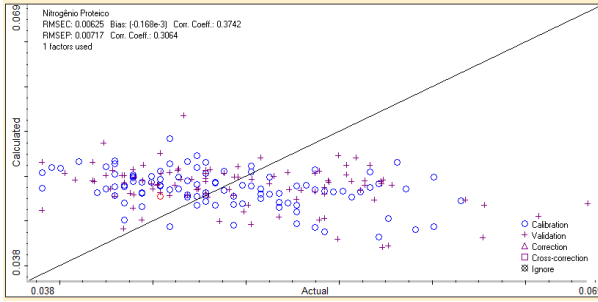
Factor	PRESS	RMSECV
0	0.00279	0.00531
1	0.00008	0.00092
2	0.00009	0.00093
3	0.00009	0.00093
4	0.00010	0.00100
5	0.00011	0.00106
6	0.00012	0.00111
7	0.00011	0.00103
8	0.00011	0.00107
9	0.00012	0.00109
10	0.00012	0.00109

PLS/MSC



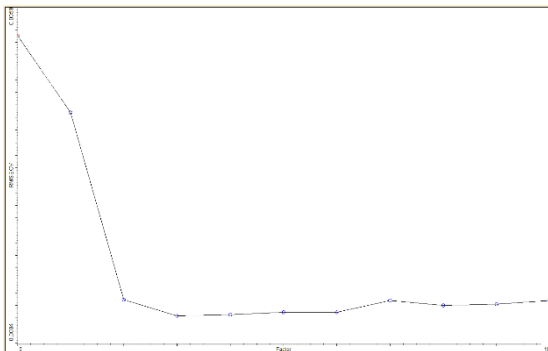
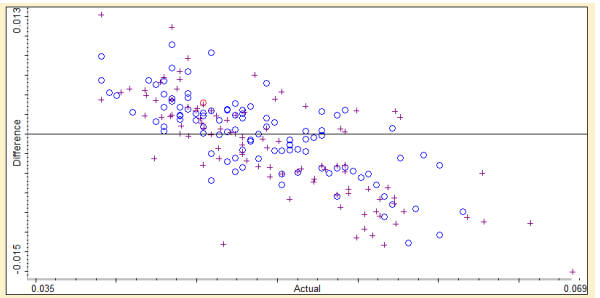
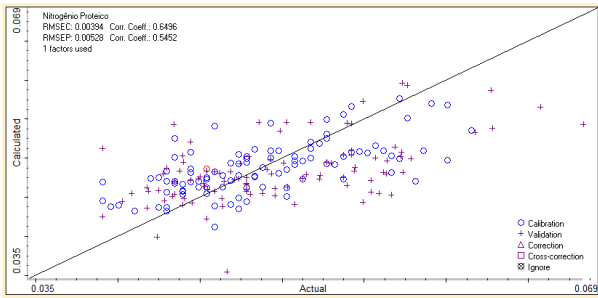
Factor	PRESS	RMSECV
0	0.23802	0.04903
1	0.00275	0.00527
2	0.00008	0.00092
3	0.00009	0.00093
4	0.00009	0.00093
5	0.00010	0.00102
6	0.00011	0.00103
7	0.00012	0.00110
8	0.00010	0.00103
9	0.00012	0.00108
10	0.00012	0.00108

PLS



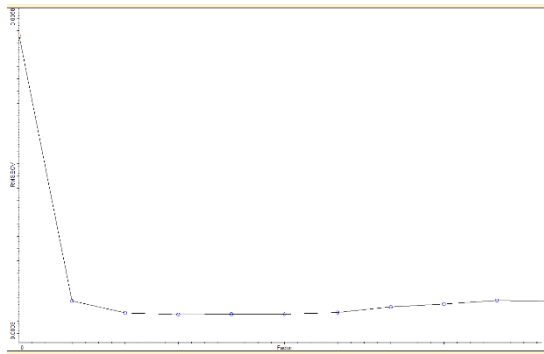
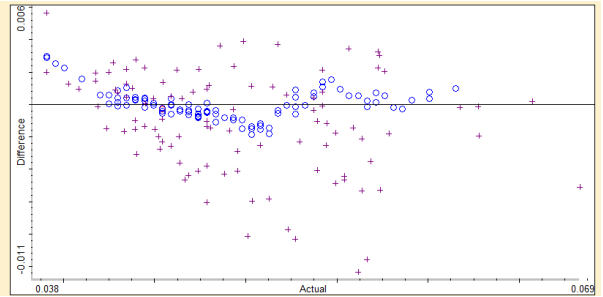
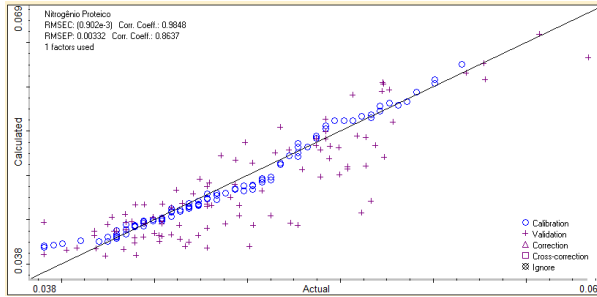
Factor	PRESS	RMSECV
0	0.23802	0.04903
1	0.00404	0.00639
2	0.00046	0.00217
3	0.00046	0.00216
4	0.00038	0.00196
5	0.00029	0.00171
6	0.00029	0.00170
7	0.00028	0.00167
8	0.00025	0.00158
9	0.00029	0.00171
10	0.00032	0.00180

PLS/CM



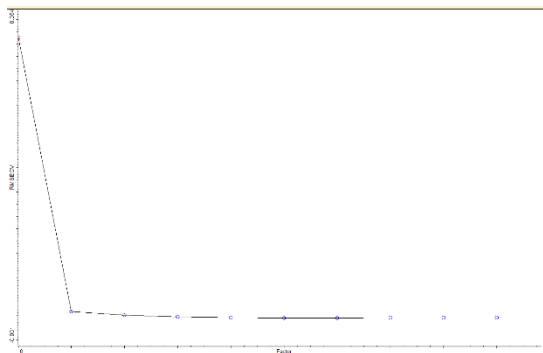
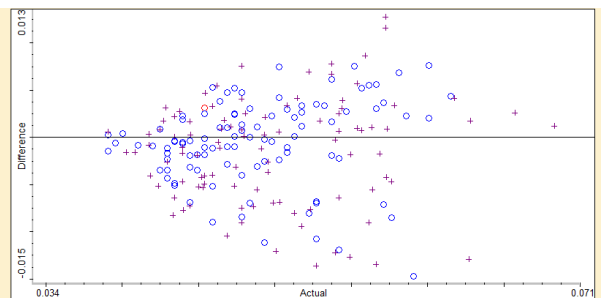
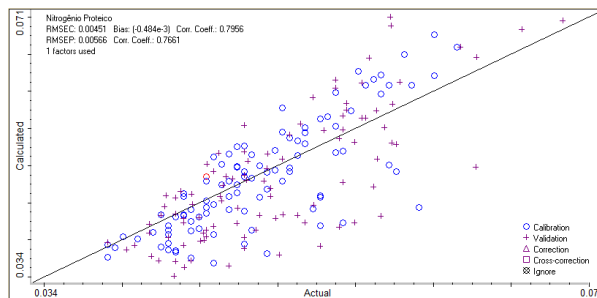
Factor	PRESS	RMSECV
0	0.00279	0.00531
1	0.00165	0.00408
2	0.00012	0.00109
3	0.00007	0.00083
4	0.00007	0.00085
5	0.00008	0.00089
6	0.00008	0.00089
7	0.00012	0.00108
8	0.00010	0.00100
9	0.00010	0.00102
10	0.00012	0.00108

PLS/SNV/CM



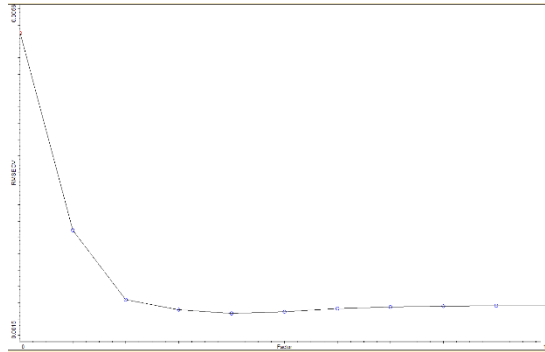
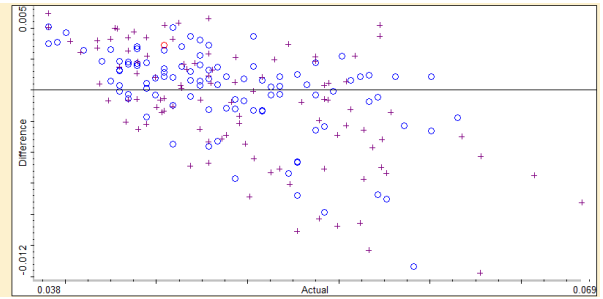
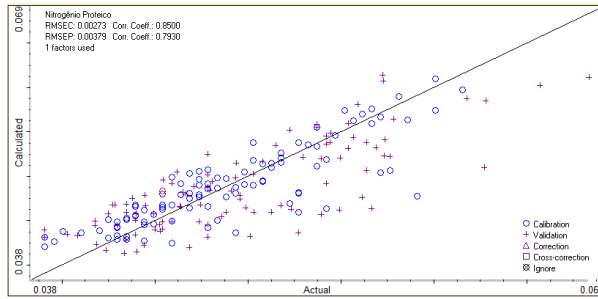
Factor	PRESS	RMSECV
0	0.00279	0.00531
1	0.00009	0.00094
2	0.00005	0.00074
3	0.00005	0.00071
4	0.00005	0.00072
5	0.00005	0.00072
6	0.00006	0.00075
7	0.00007	0.00084
8	0.00008	0.00089
9	0.00009	0.00095
10	0.00009	0.00093

PLS/SNV



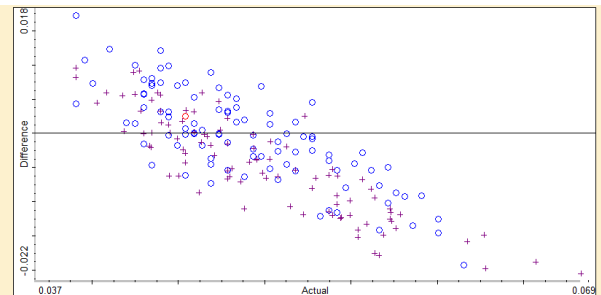
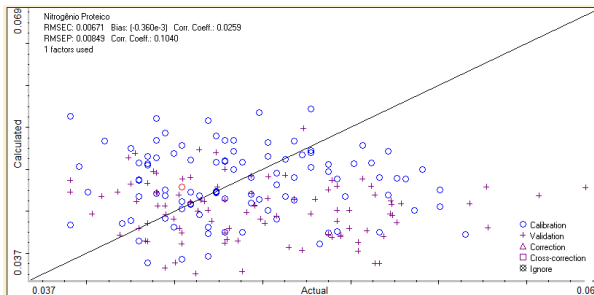
Factor	PRESS	RMSECV
0	0.23802	0.04903
1	0.00206	0.00456
2	0.00156	0.00397
3	0.00138	0.00374
4	0.00129	0.00361
5	0.00123	0.00353
6	0.00124	0.00354
7	0.00126	0.00357
8	0.00128	0.00359
9	0.00128	0.00360
10	0.00128	0.00360

PLS/1DER/SG/CM

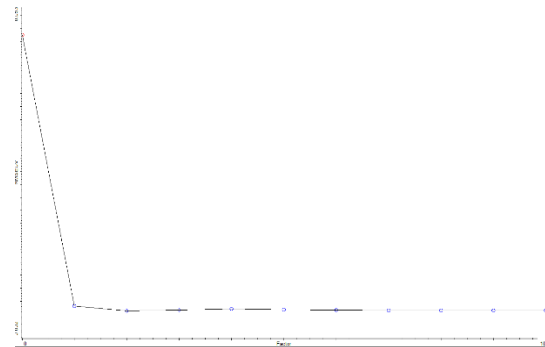


Factor	PRESS	RMSECV
0	0.00279	0.00531
1	0.00083	0.00289
2	0.00041	0.00204
3	0.00036	0.00192
4	0.00035	0.00187
5	0.00035	0.00189
6	0.00037	0.00193
7	0.00038	0.00195
8	0.00038	0.00196
9	0.00038	0.00196
10	0.00038	0.00196

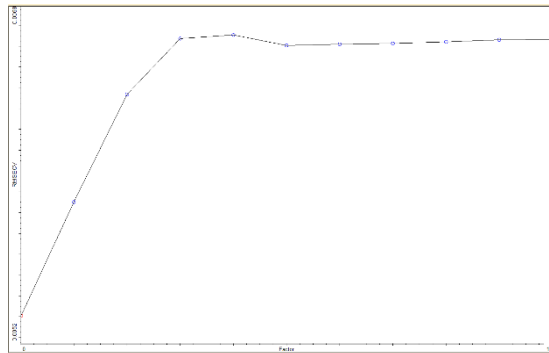
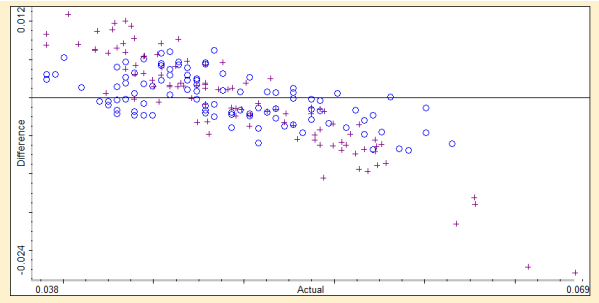
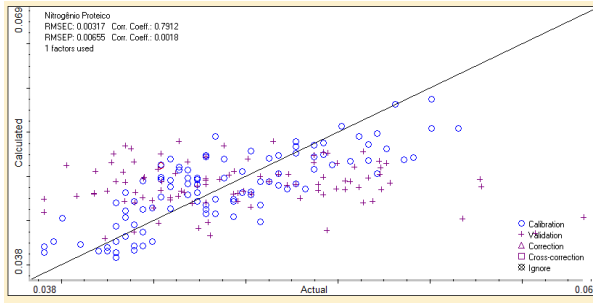
PLS/1DER/SG 5P 2 ORD



Factor	PRESS	RMSECV
0	0.23802	0.04903
1	0.00520	0.00725
2	0.00419	0.00651
3	0.00438	0.00665
4	0.00458	0.00680
5	0.00448	0.00673
6	0.00436	0.00664
7	0.00432	0.00660
8	0.00428	0.00657
9	0.00426	0.00656
10	0.00424	0.00655



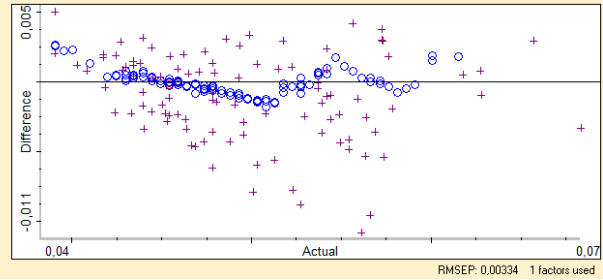
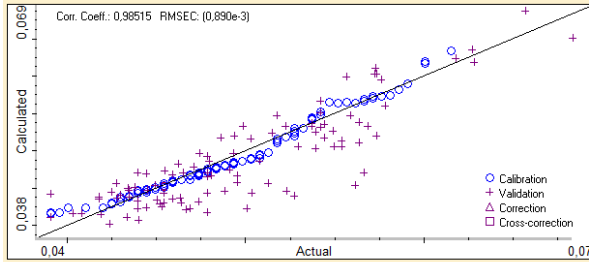
PLS/2 DER/SG



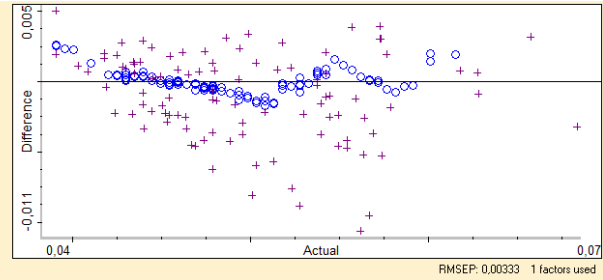
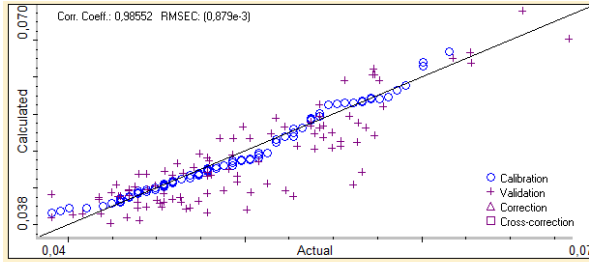
Factor	PRESS	RMSECV
0	0.00279	0.00531
1	0.00339	0.00585
2	0.00401	0.00637
3	0.00436	0.00664
4	0.00438	0.00665
5	0.00432	0.00660
6	0.00433	0.00661
7	0.00433	0.00661
8	0.00434	0.00662
9	0.00435	0.00663
10	0.00435	0.00663

ANEXO B- ESTUDO DAS REGIÕES

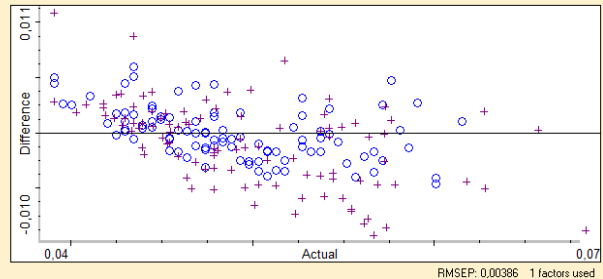
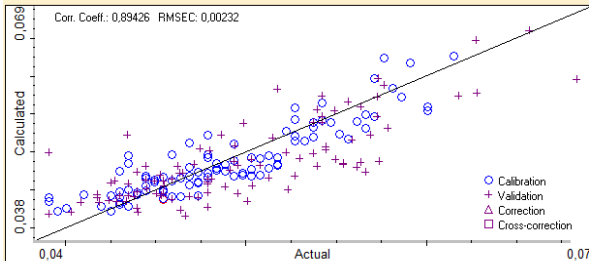
PLS/MSC/CM, TODO ESPECTRO



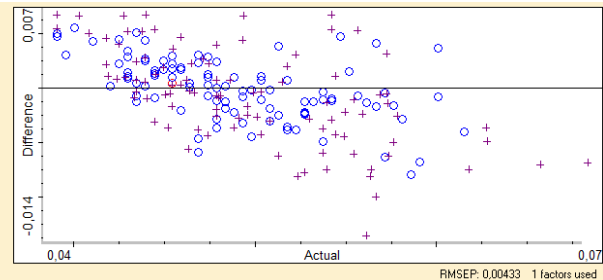
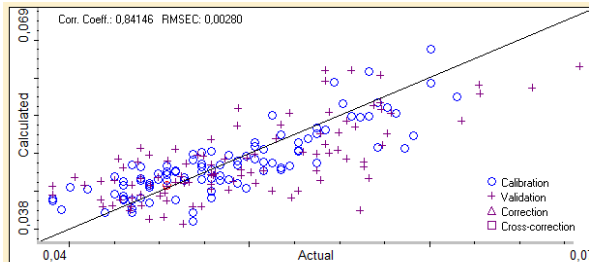
REGIAO A



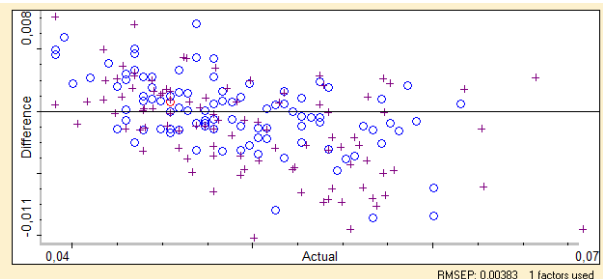
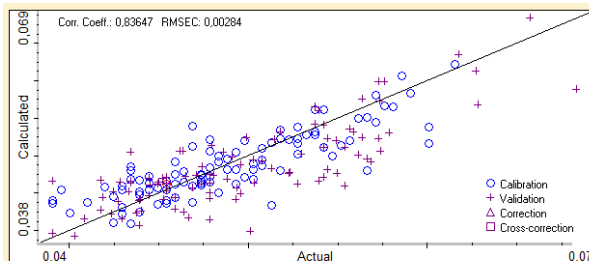
REGIAO B



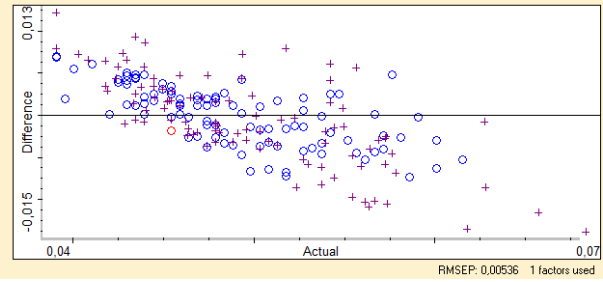
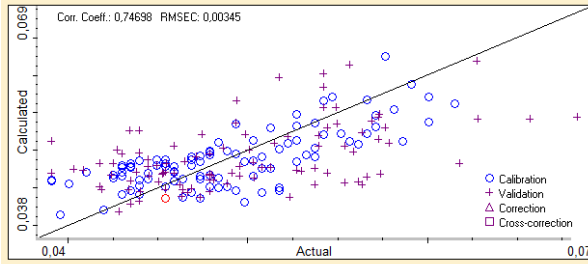
REGIAO C



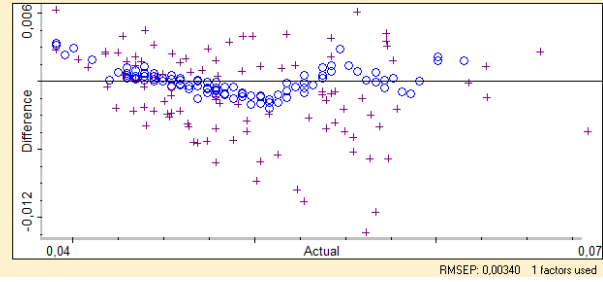
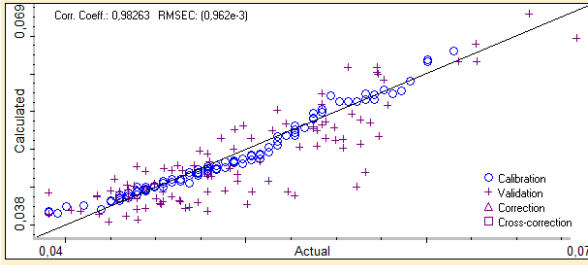
REGIAO D



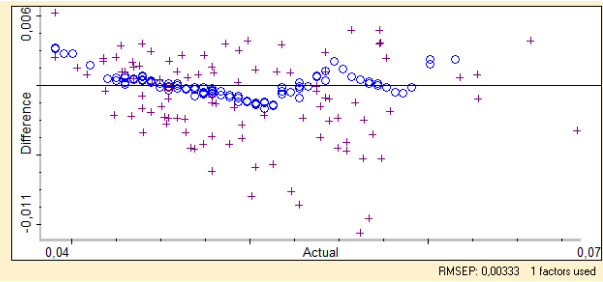
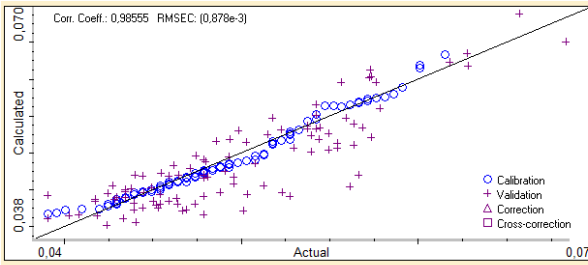
REGIAO E



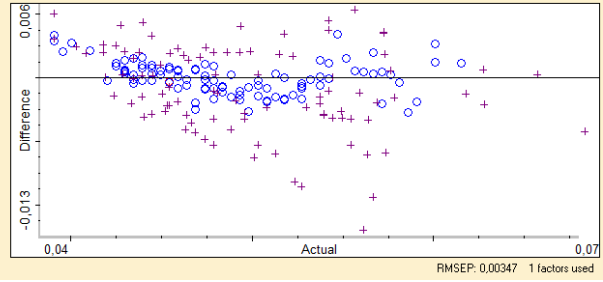
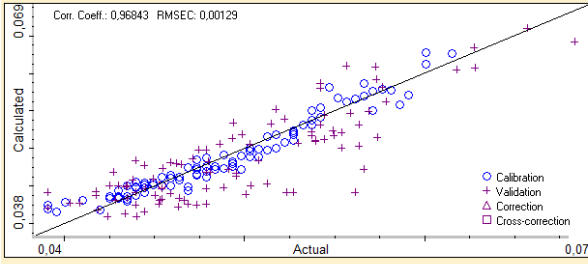
REGIAO F



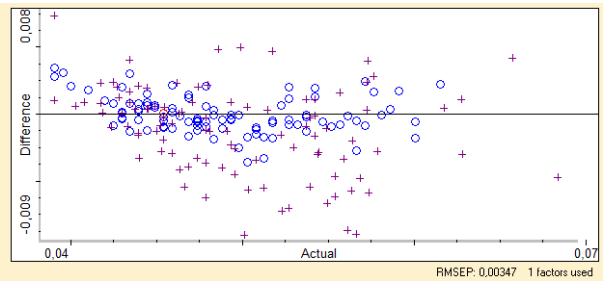
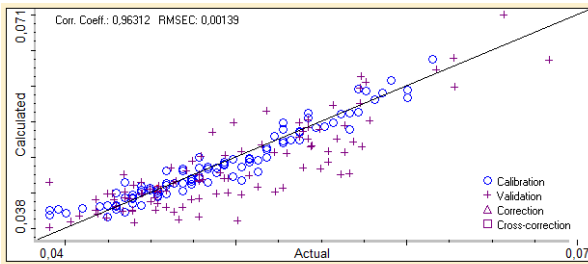
PLS/MSC/CM REGIAO AA



REGIAO BB

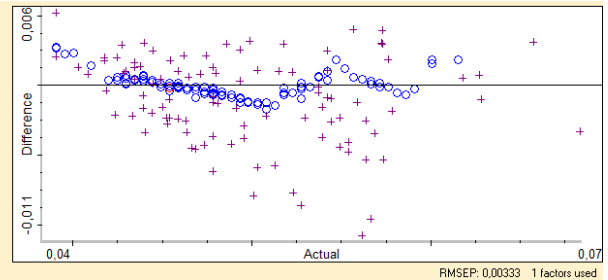
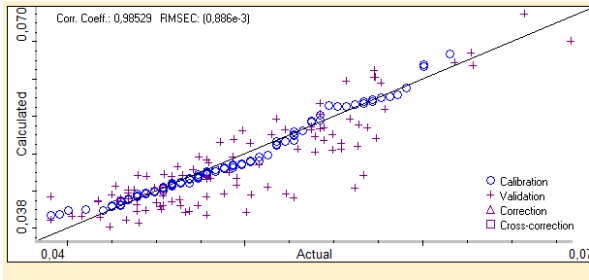


REGIAO CC

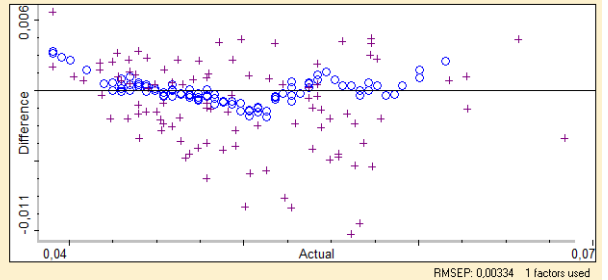
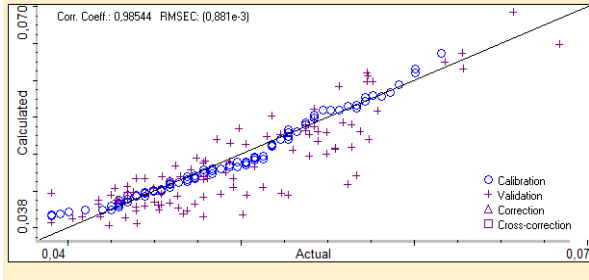


COMBINAÇÕES

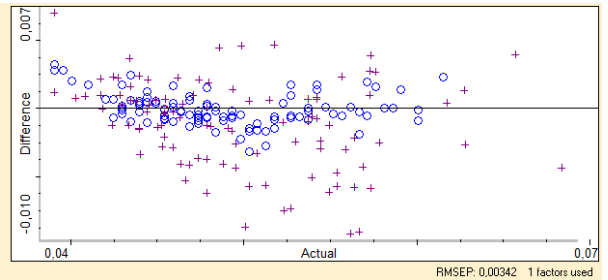
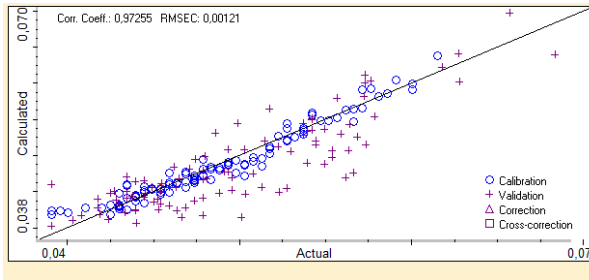
AA+BB



AA+CC

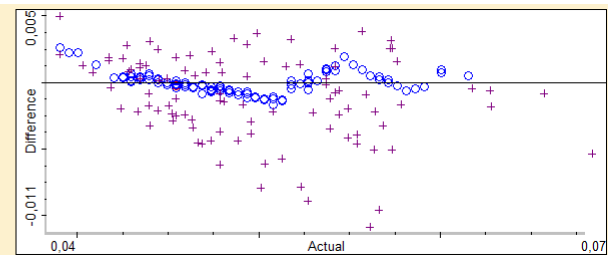
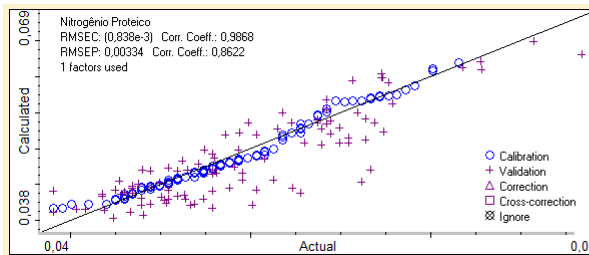


BB+CC

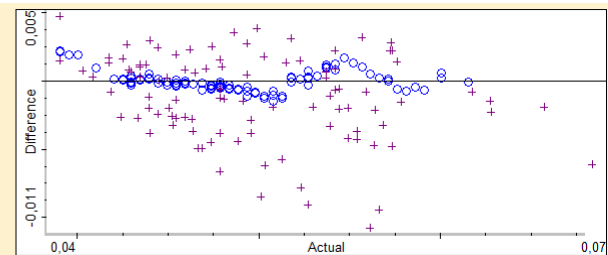
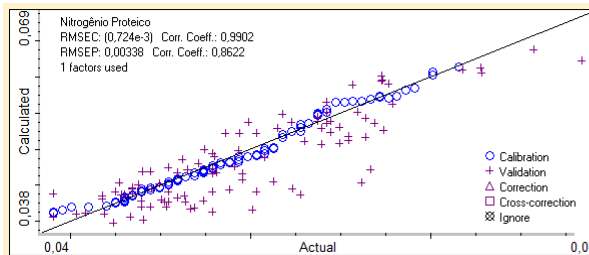


TRATAMENTO SNV

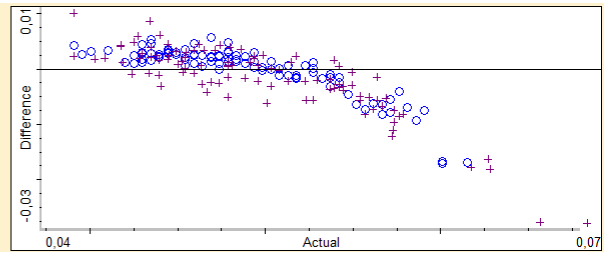
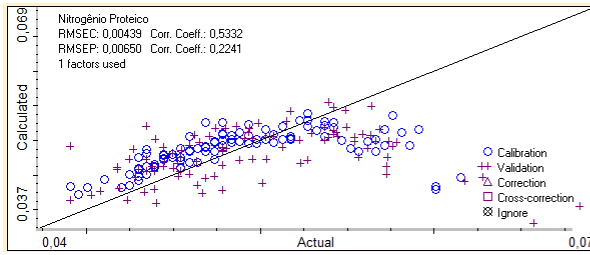
ABCDEF



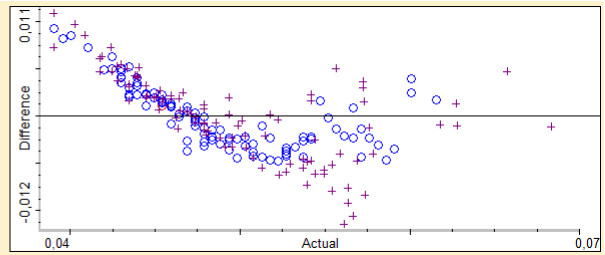
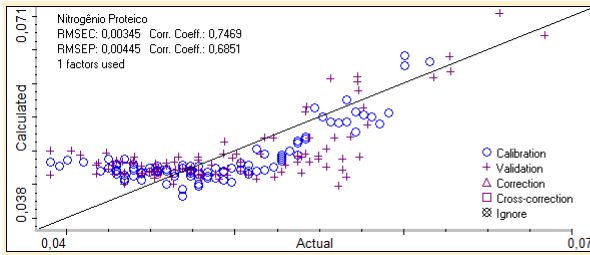
A



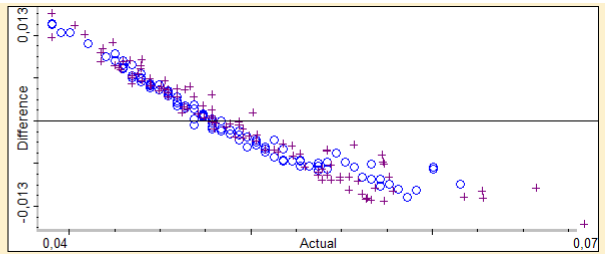
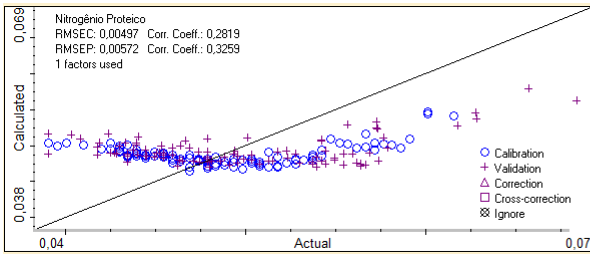
B



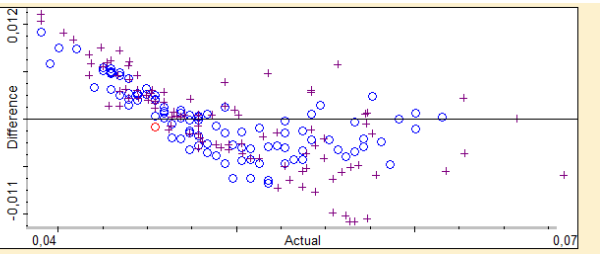
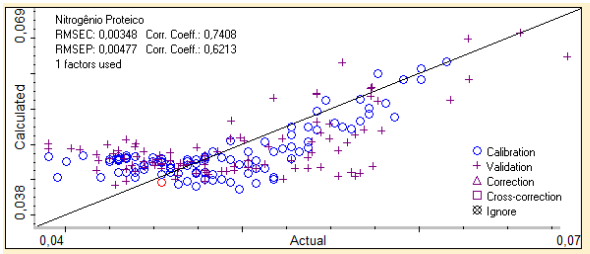
C



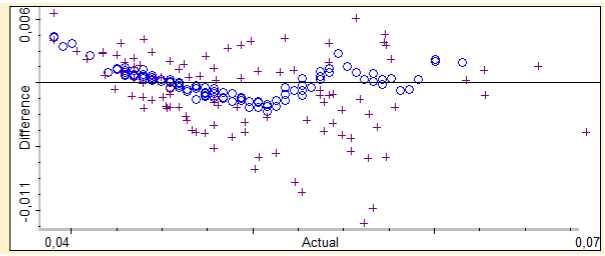
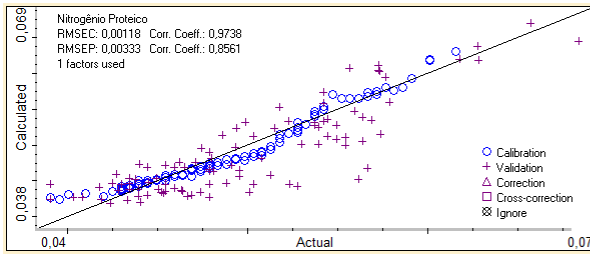
D



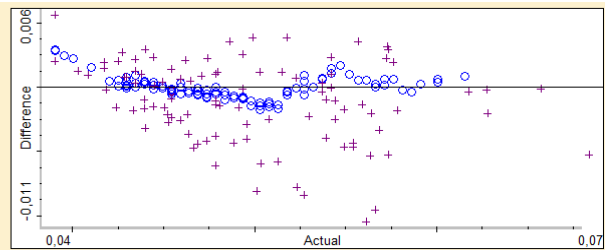
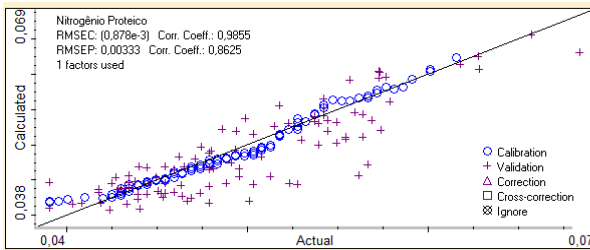
E



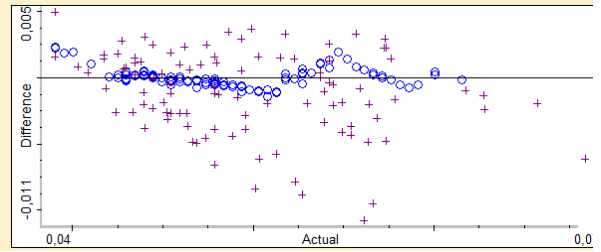
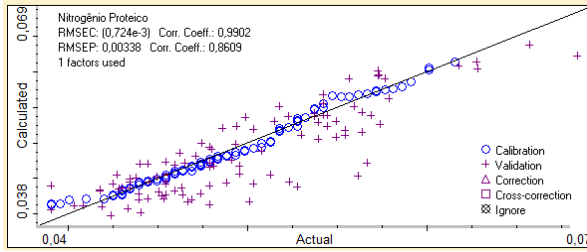
F



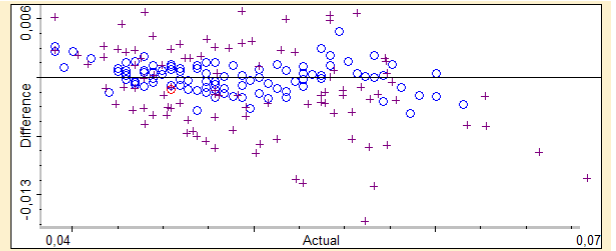
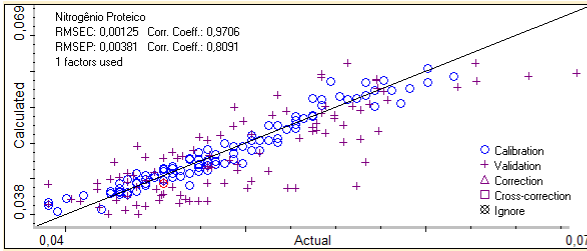
AA+BB+CC



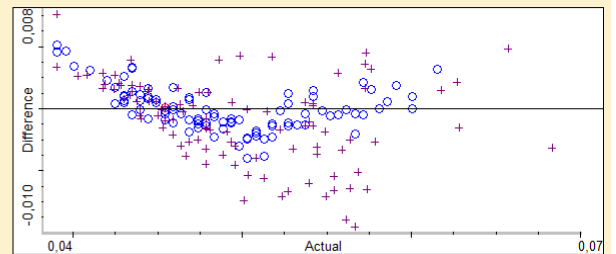
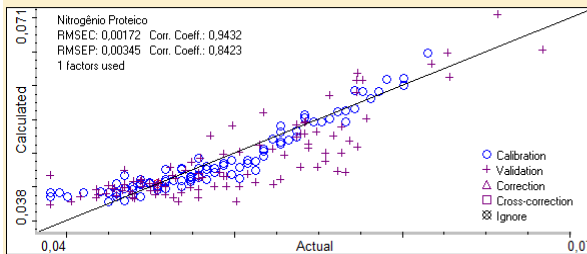
AA



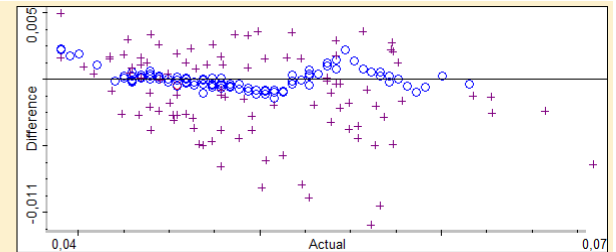
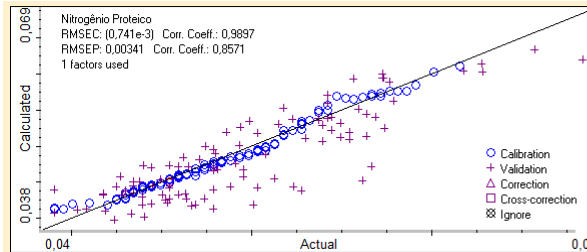
BB



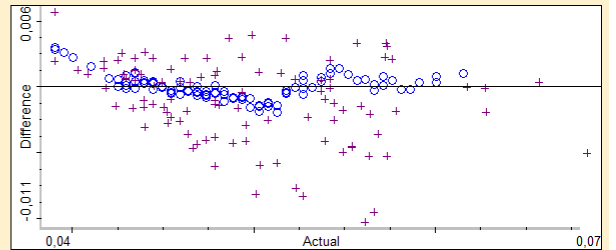
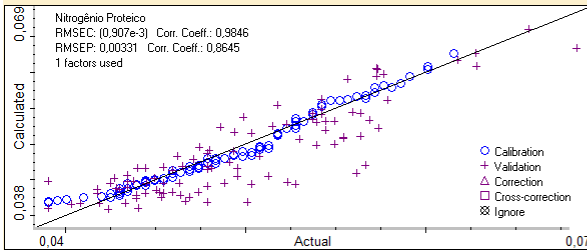
CC



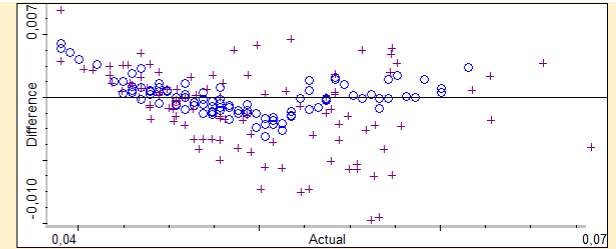
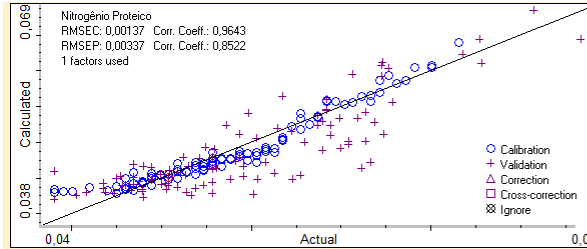
AA+BB



AA+CC

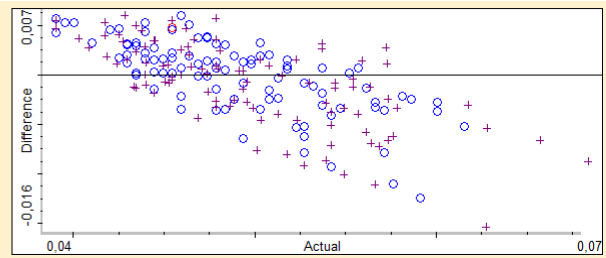
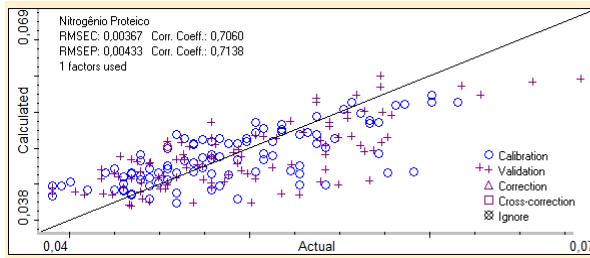


BB+CC

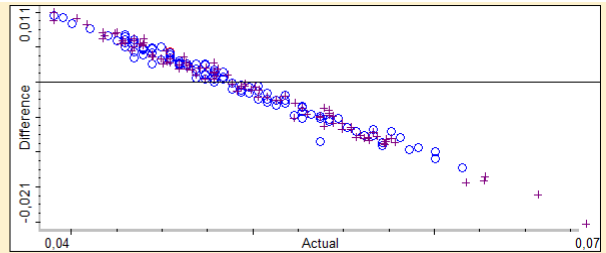
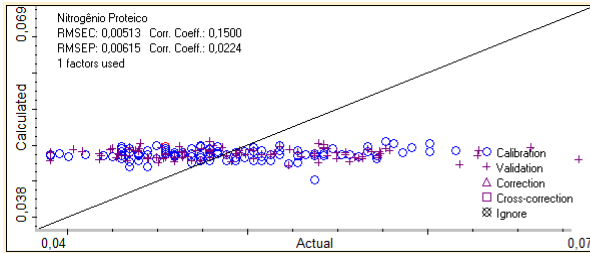


TRATAMENTO 1DER

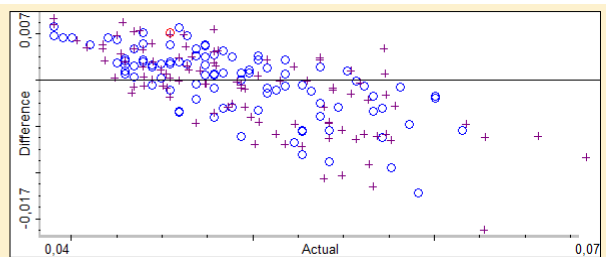
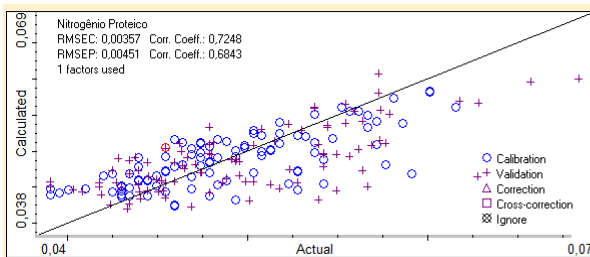
REGIÕES A



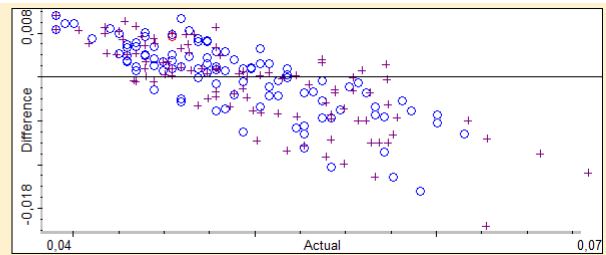
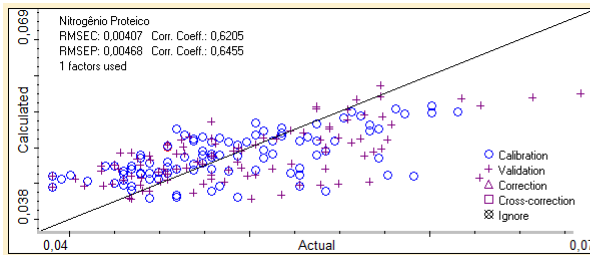
Tratamento 1 der B



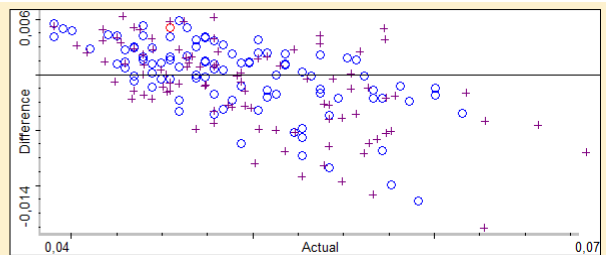
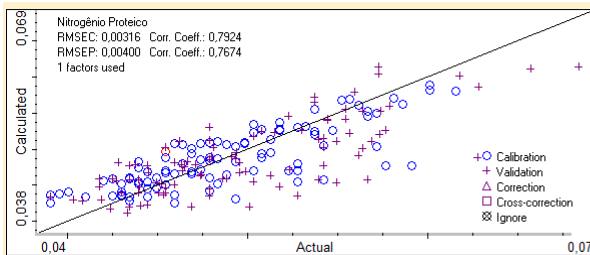
C



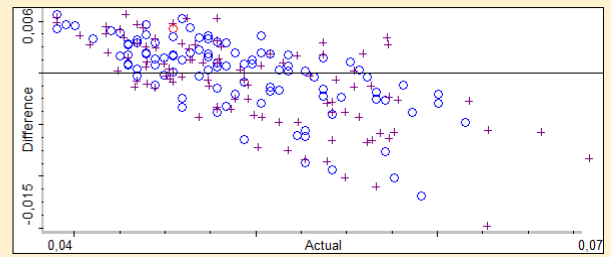
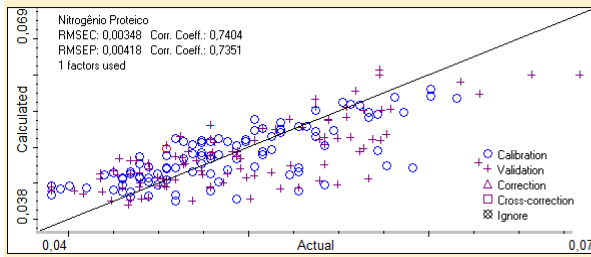
D



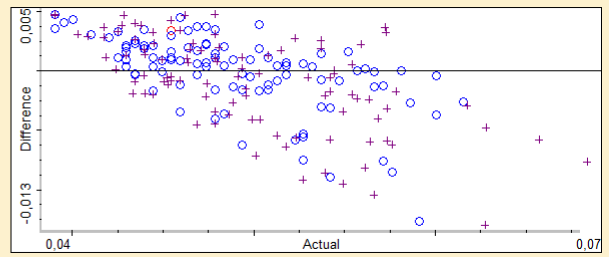
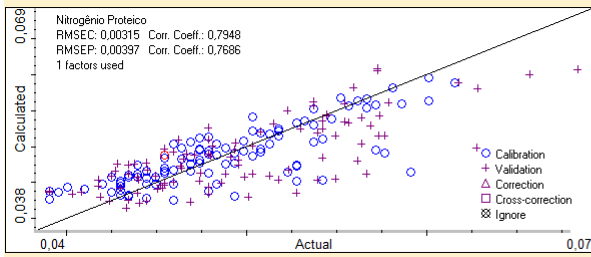
E



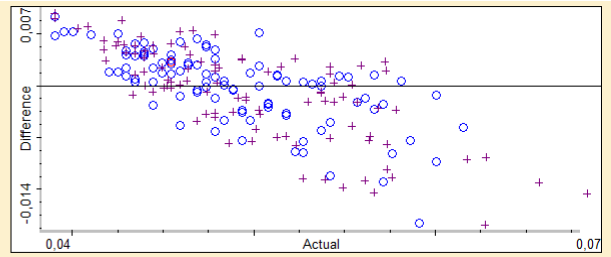
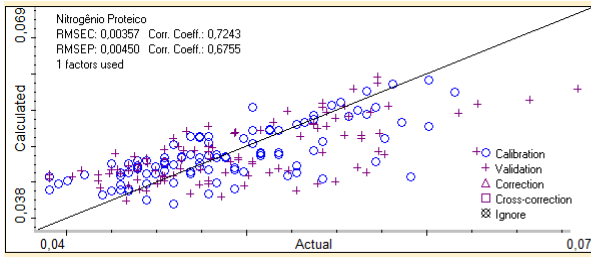
F



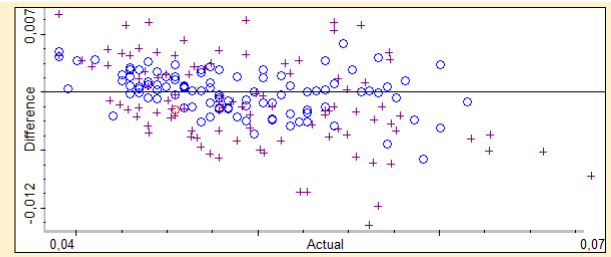
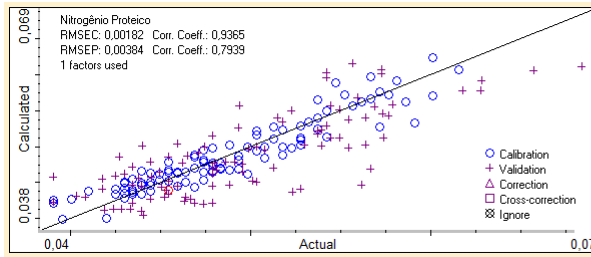
G



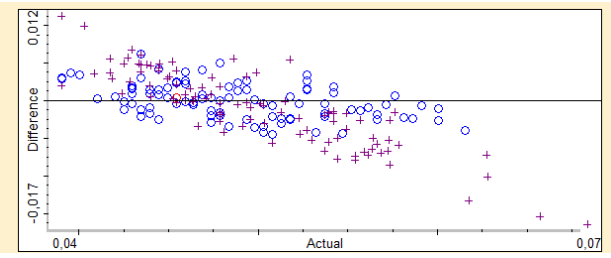
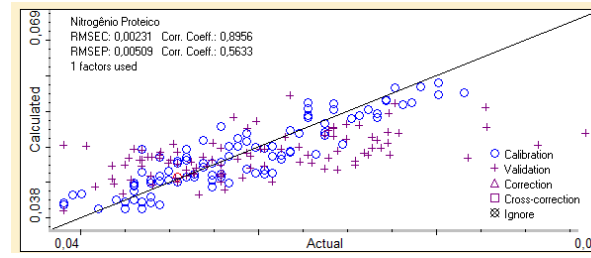
H



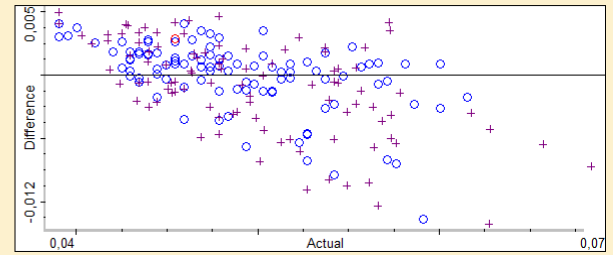
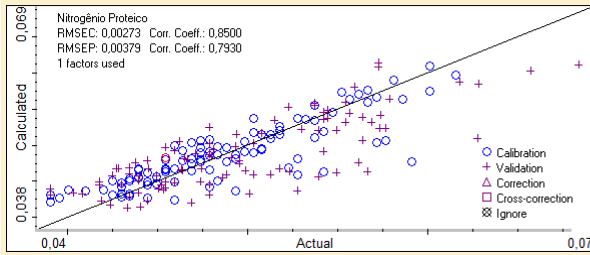
I



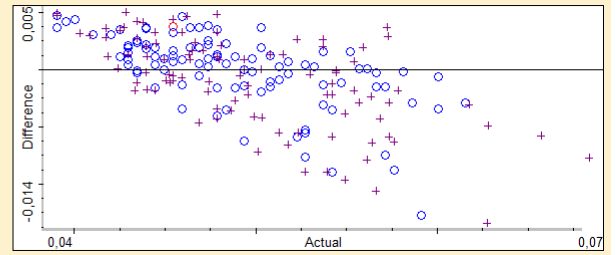
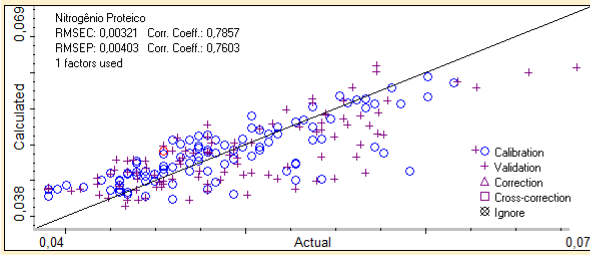
J



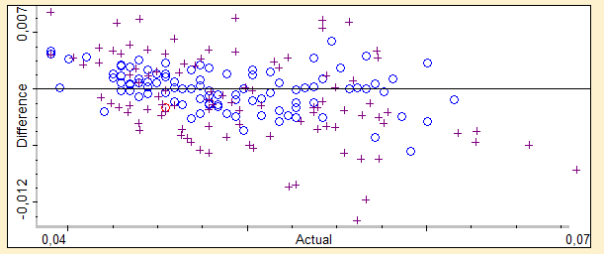
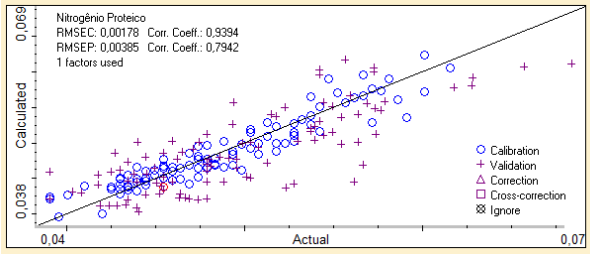
Tratamento 1 der, AA+BB+CC



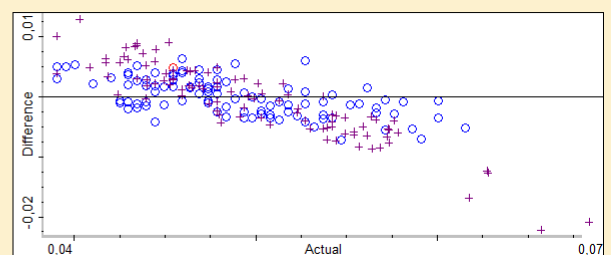
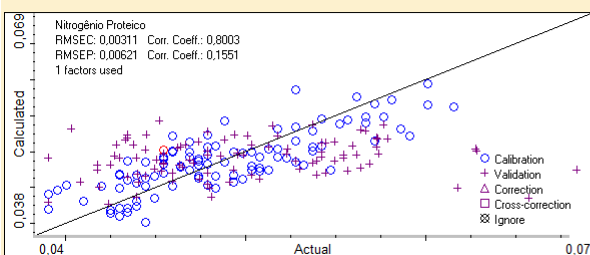
AA



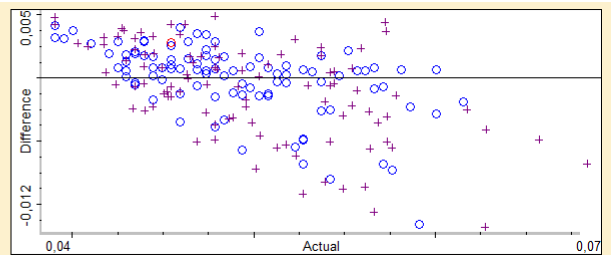
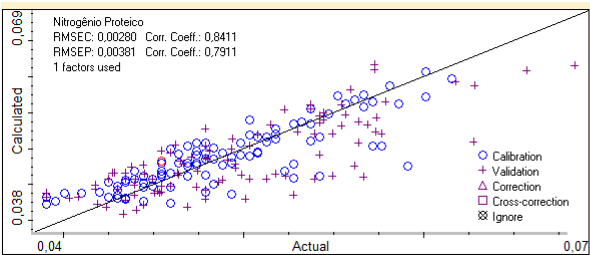
BB



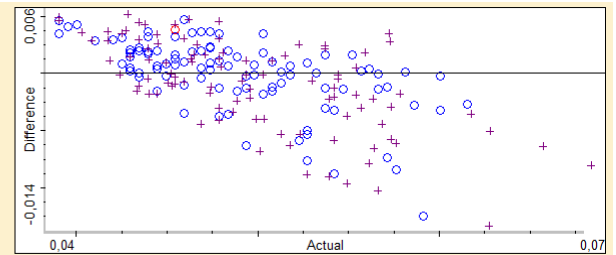
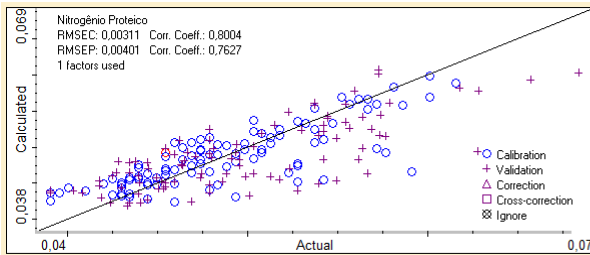
CC



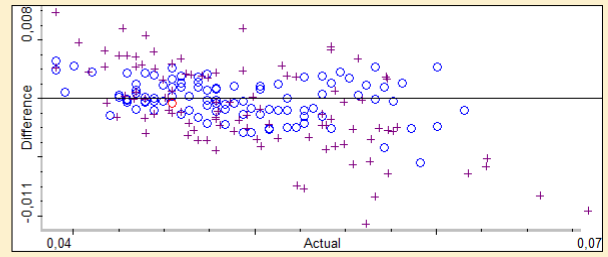
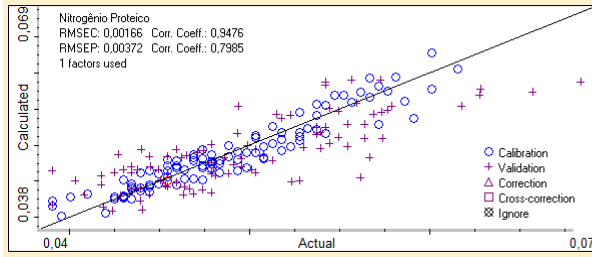
AA BB



AA CC

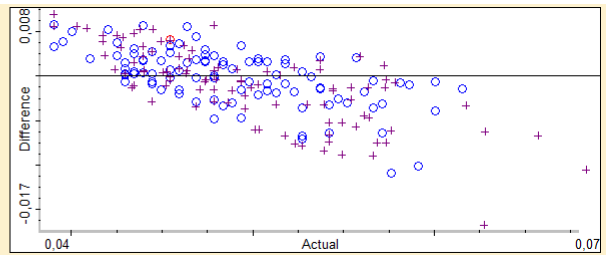
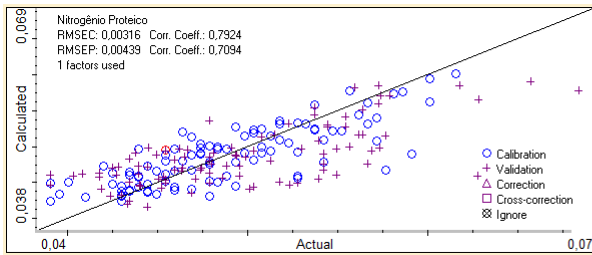


BB CC

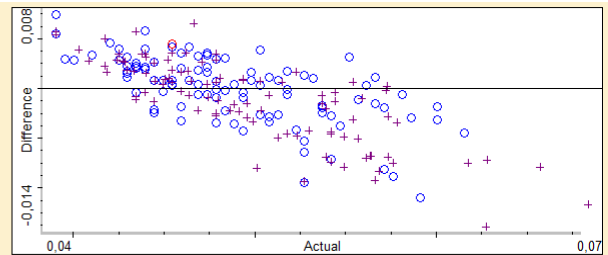
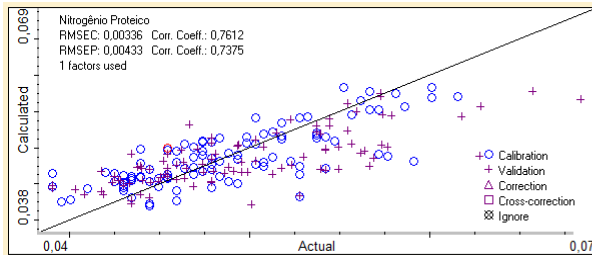


TRATAMENTO 2DER

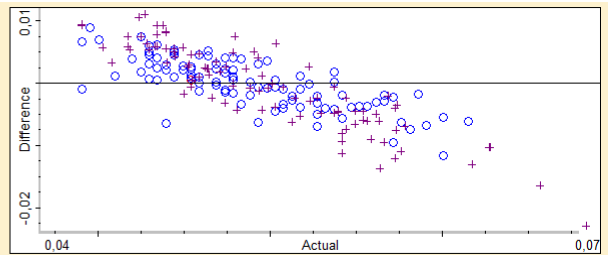
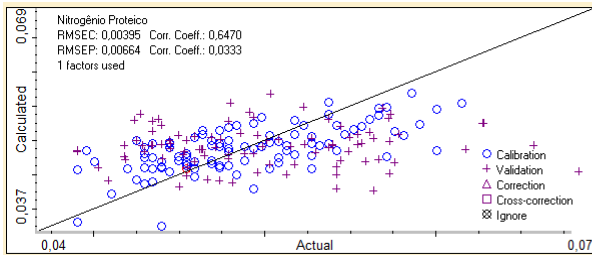
A



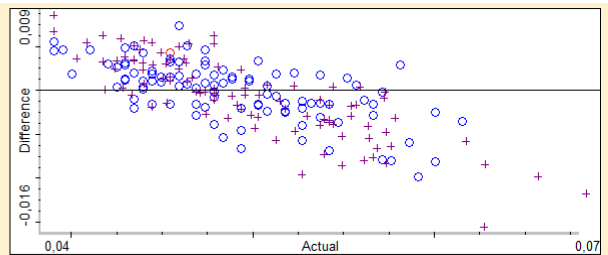
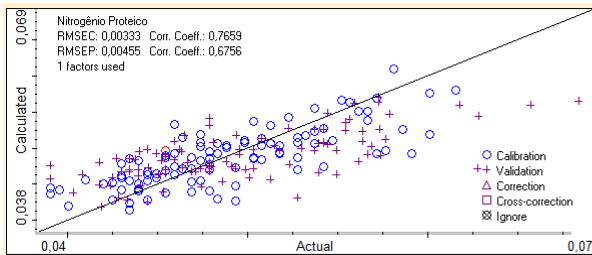
B



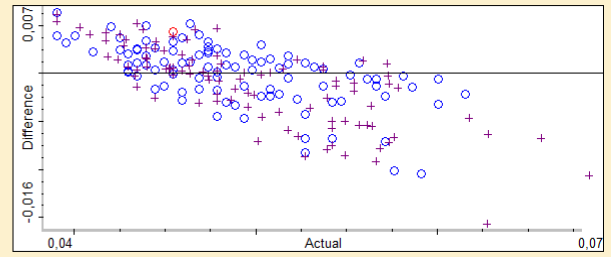
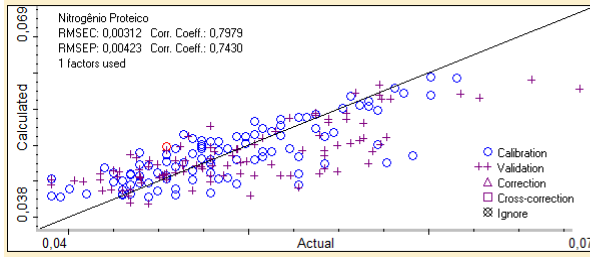
C



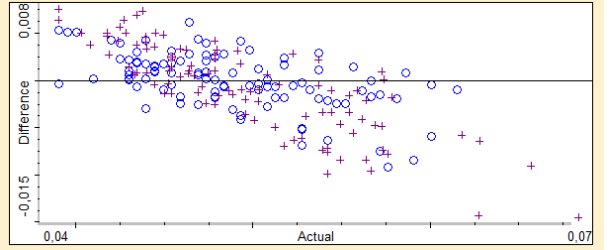
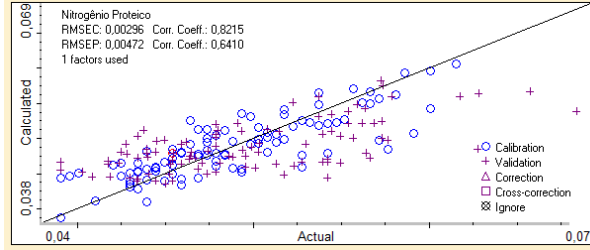
D



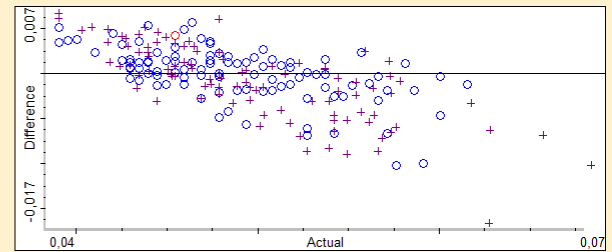
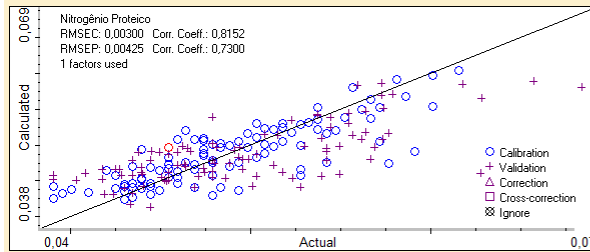
AB



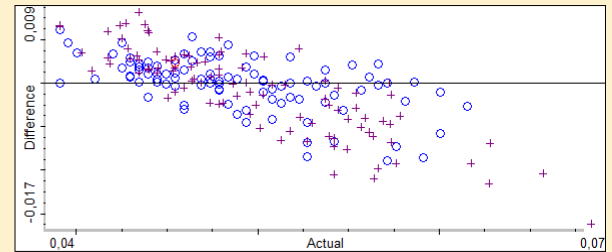
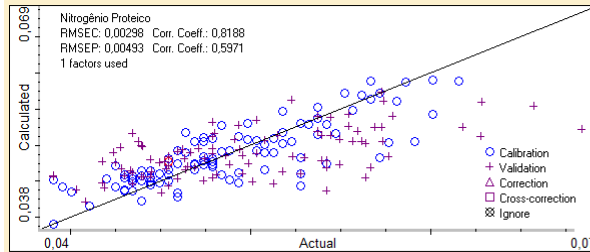
AC



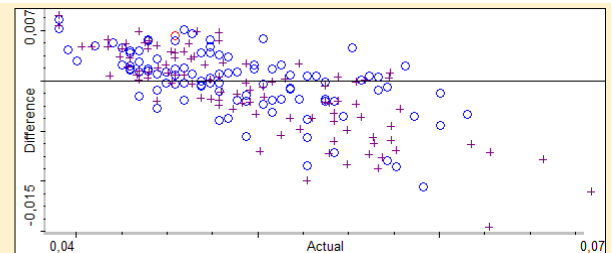
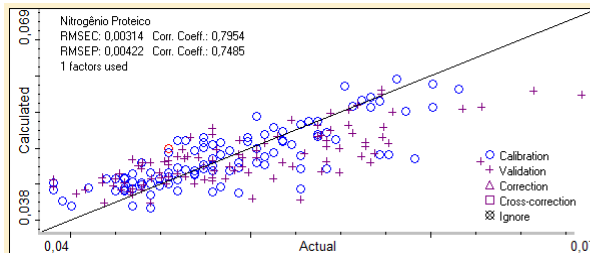
AD



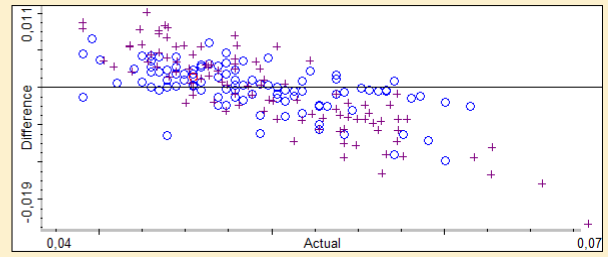
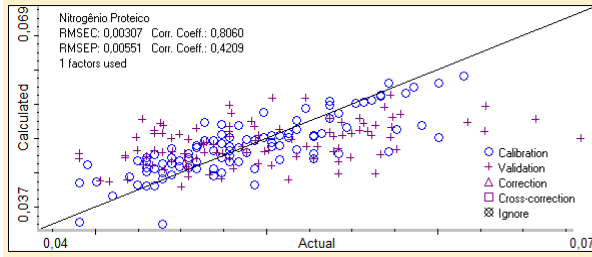
BC



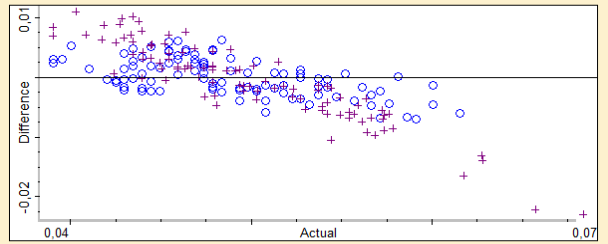
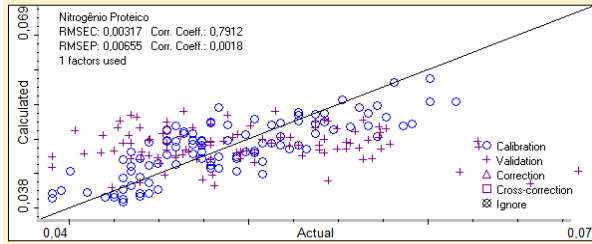
DB



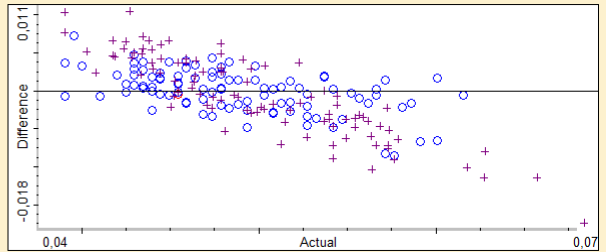
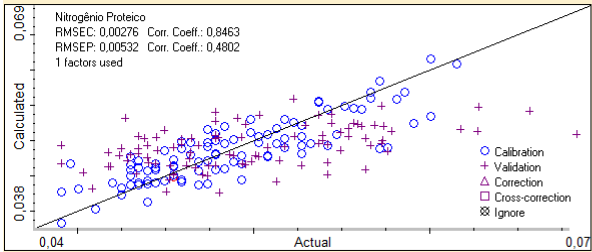
DC



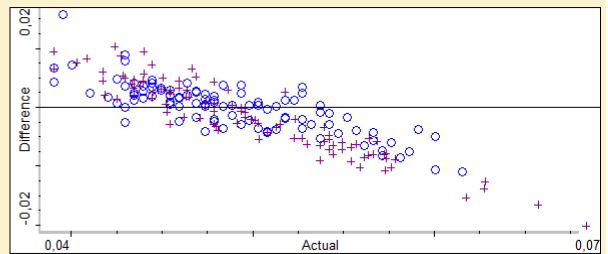
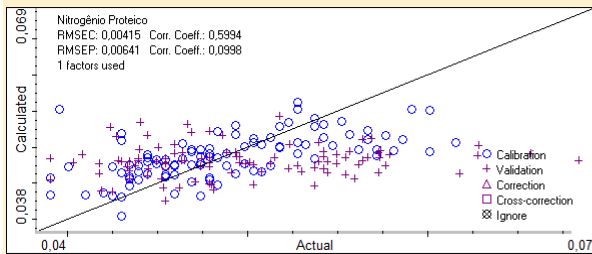
Tratamento 2 der , AA BB CC



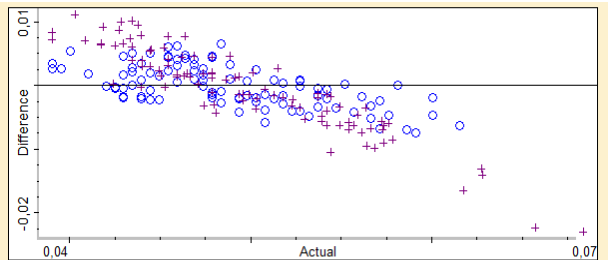
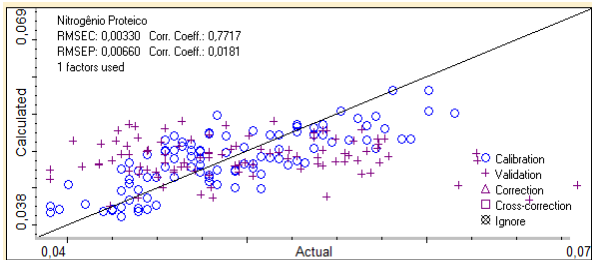
AA



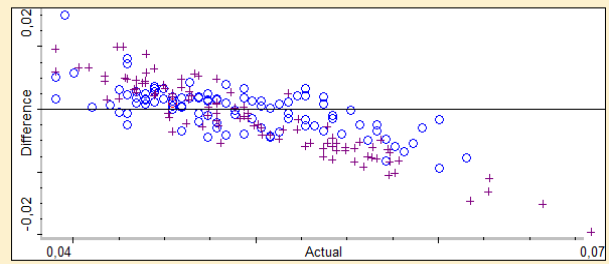
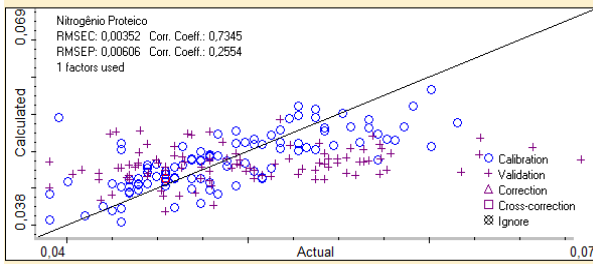
BB



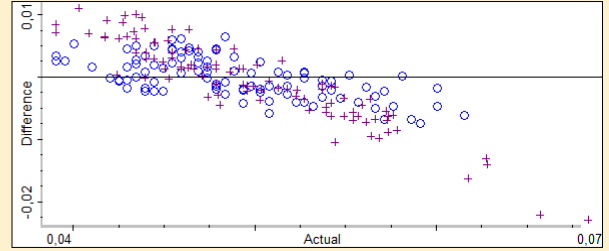
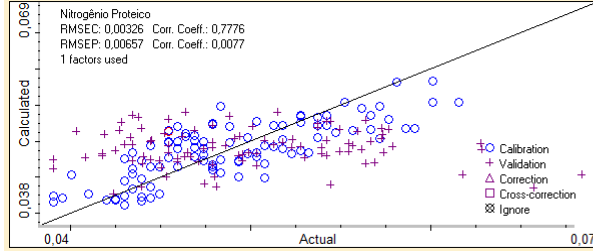
CC



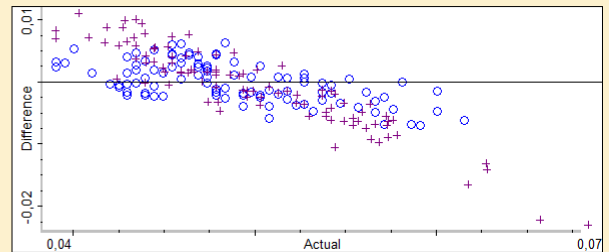
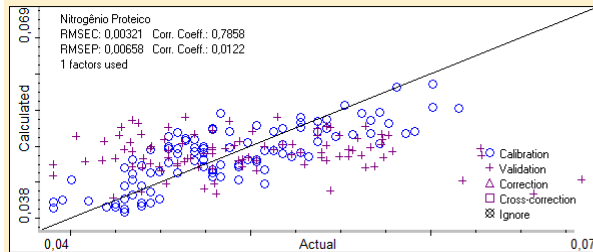
AA BB



AA CC

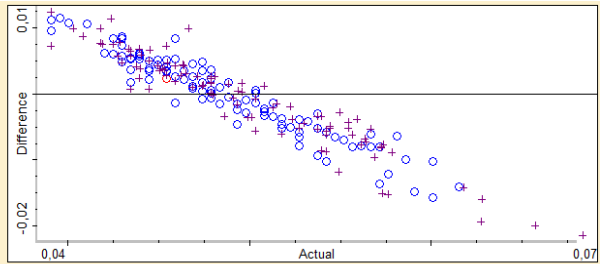
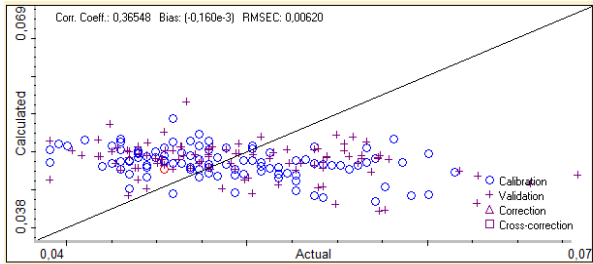


CC BB

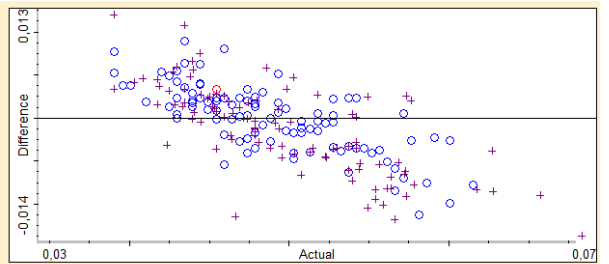
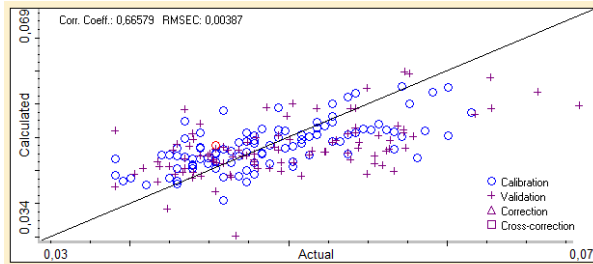


Calibração em todo o espectro

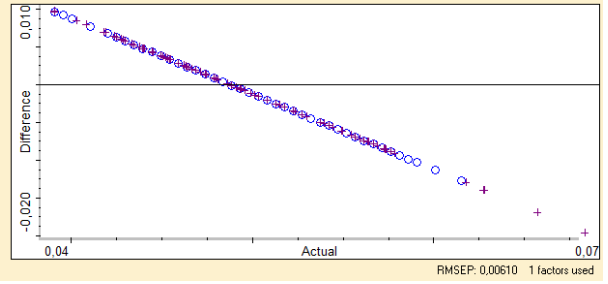
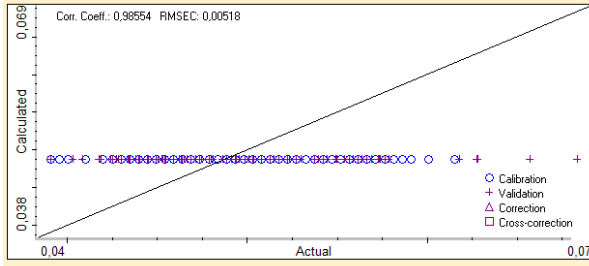
PLS



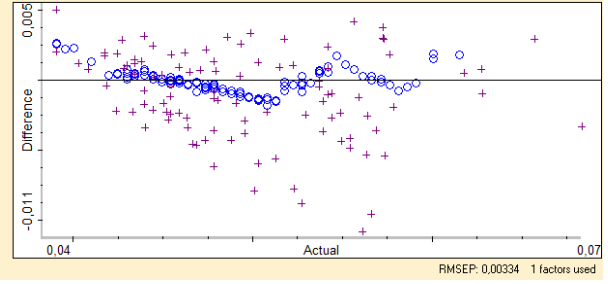
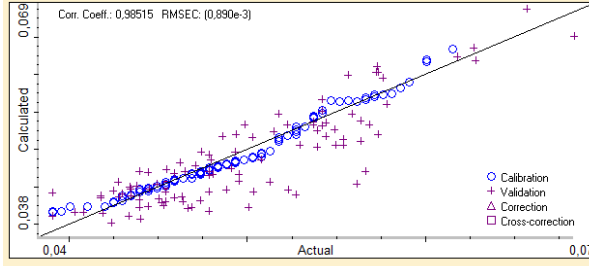
PLS/CM



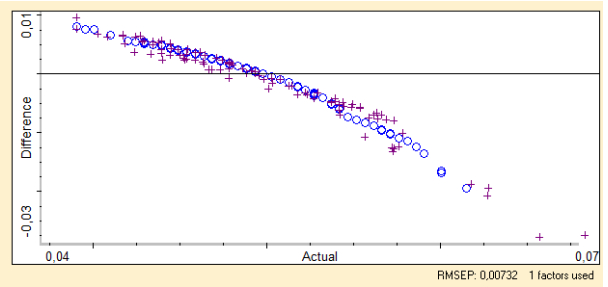
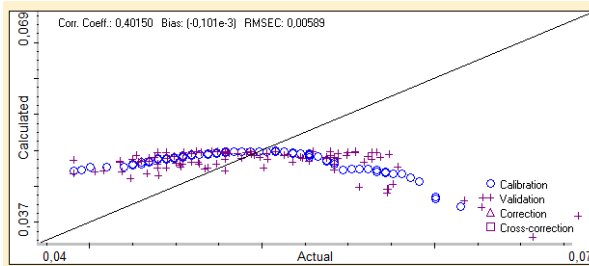
PLS/MSC



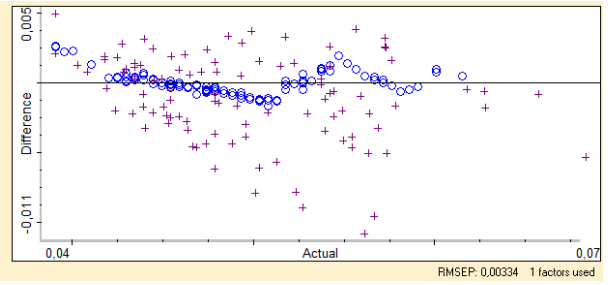
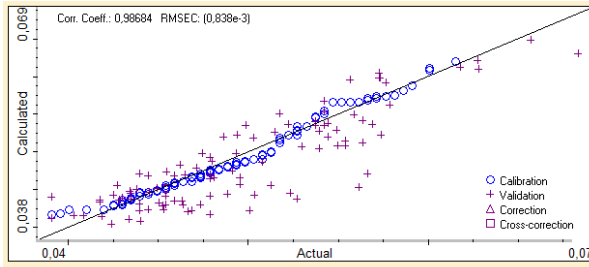
PLS/MSC/CM



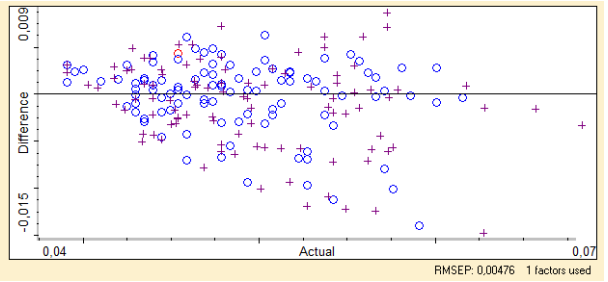
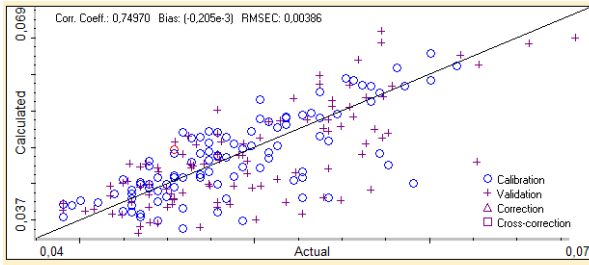
PLS/SNV



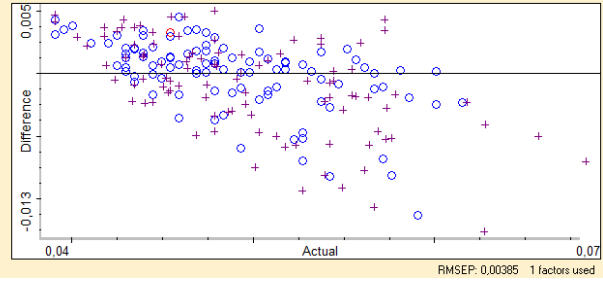
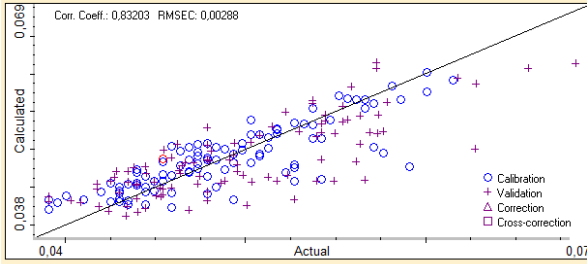
PLS/SNV/CM



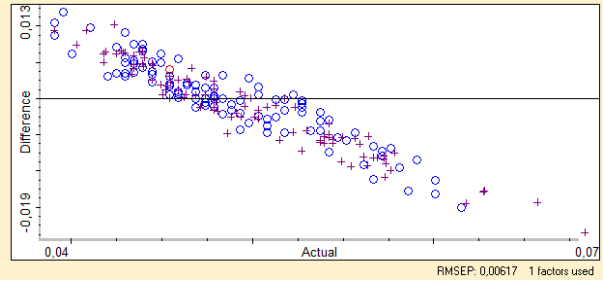
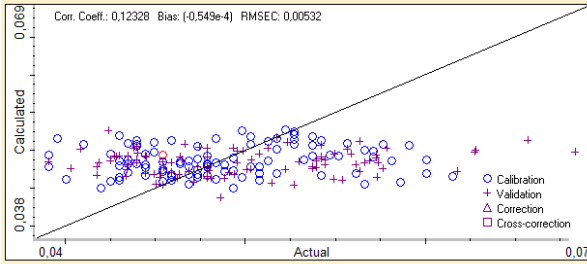
PLS/1DER/SG



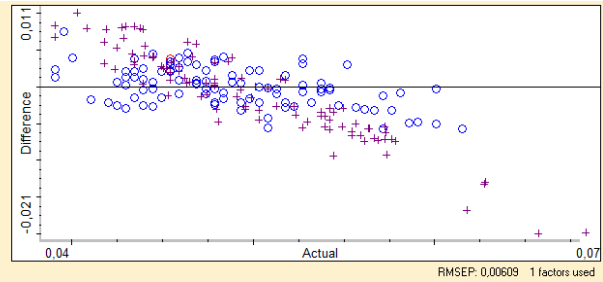
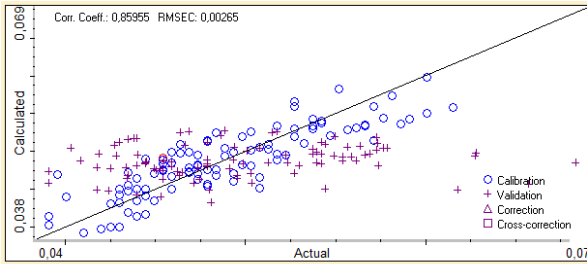
PLS/1DER/SG/CM



PLS/2DER/SG

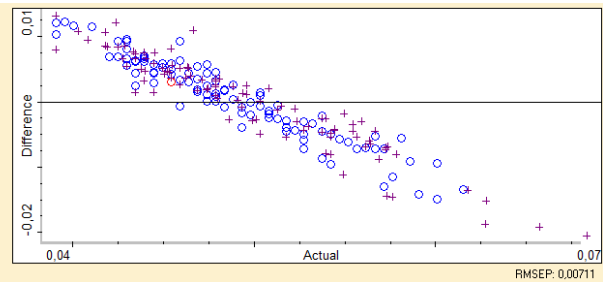
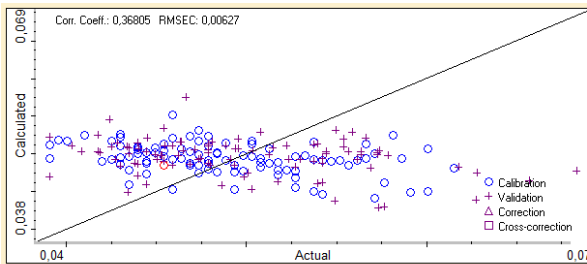


PLS/2DER/SG/CM

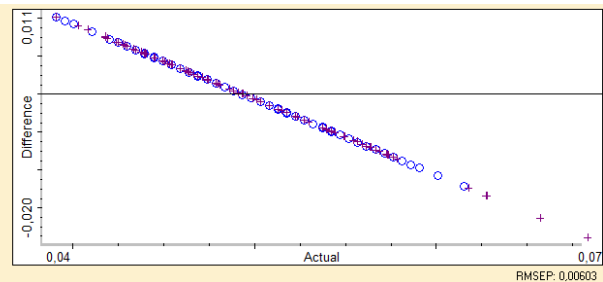
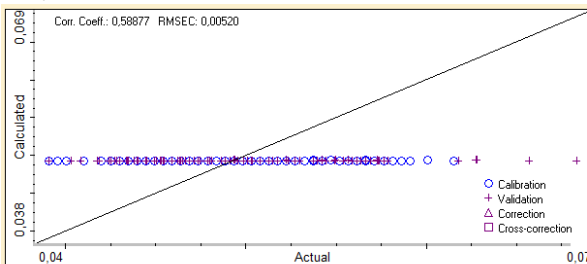


MÉTODOS-Modelos avaliados com 99 valores para calibração e 99 para a validação, utilizando as 3 regiões sugeridas pela literatura.

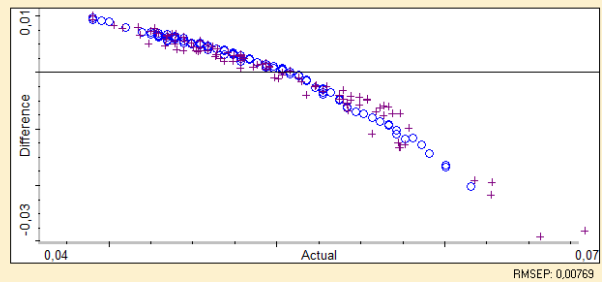
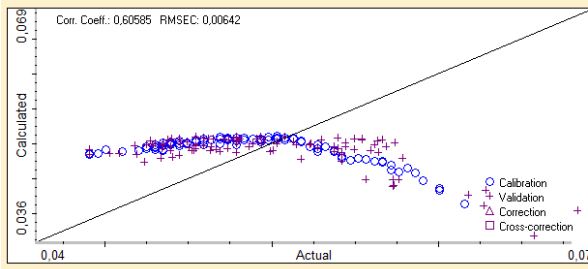
CLS



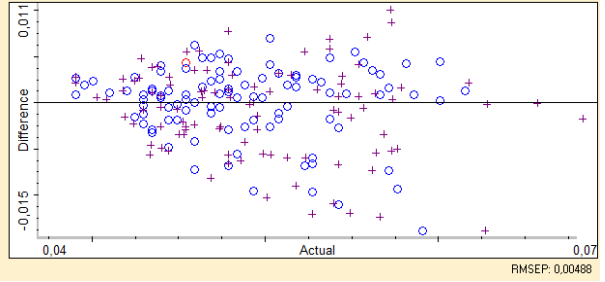
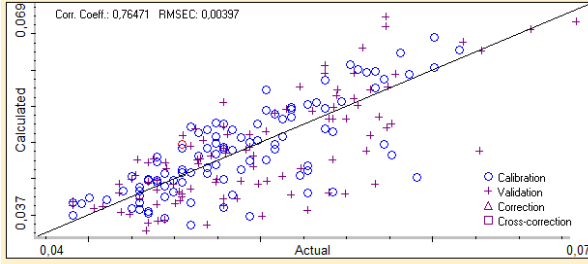
CLS/MSC



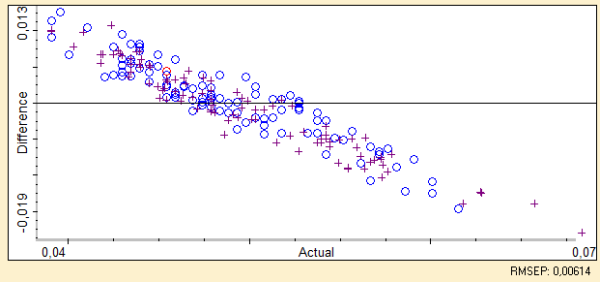
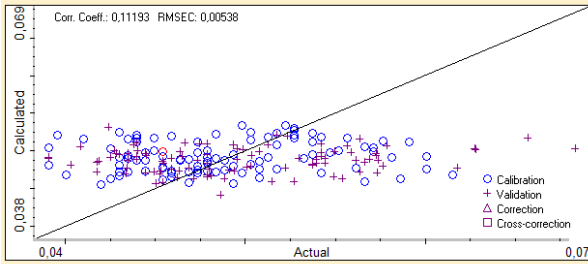
CLS/SNV



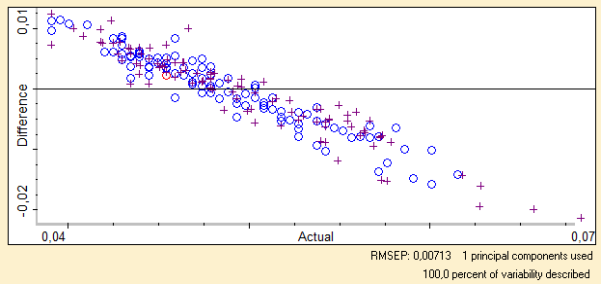
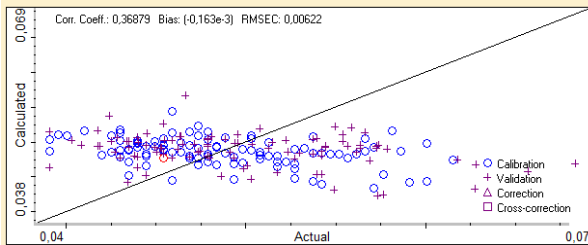
CLS/1DER/SG95P 2ORD)



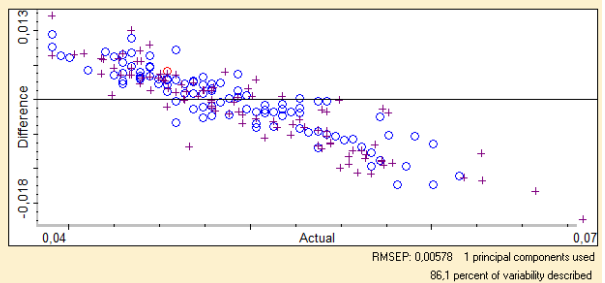
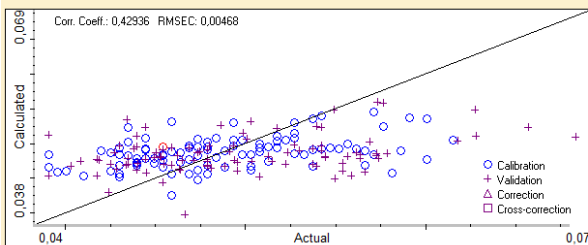
CLS/2DER/SG 5p 2ord



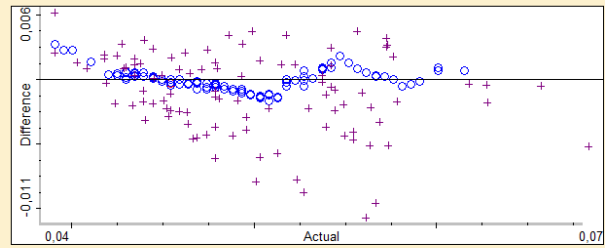
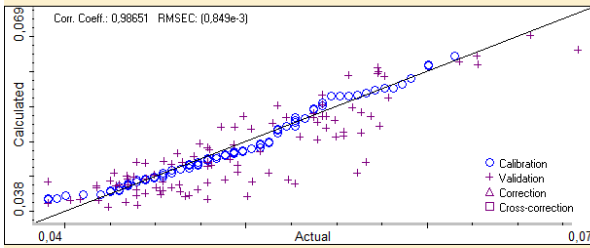
PCR



PCR/CM

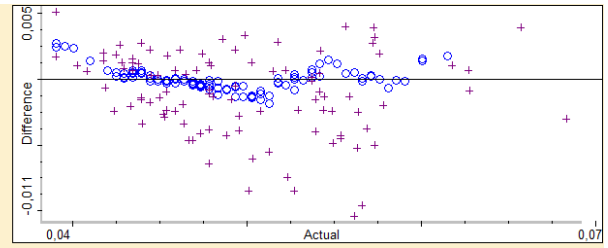
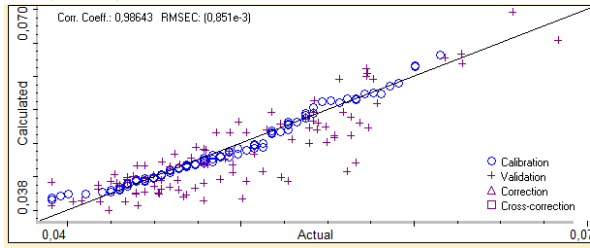


PCR/SNV/CM



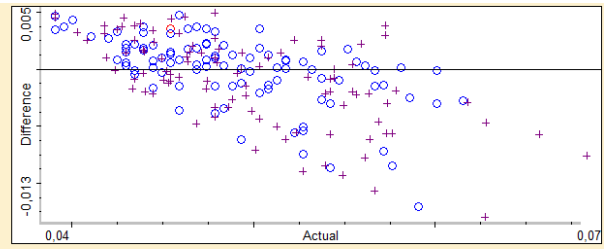
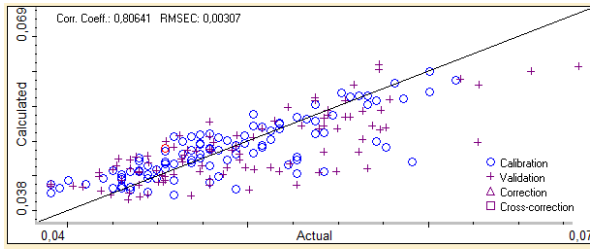
RMSEP: 0.00333 1 principal components used
96,8 percent of variability described

PCR/MSC/CM



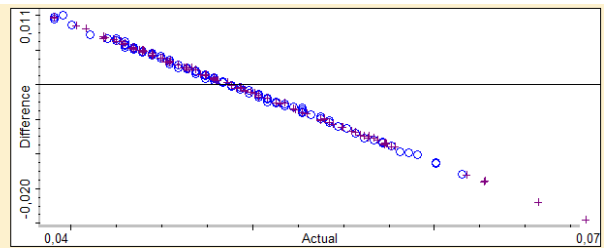
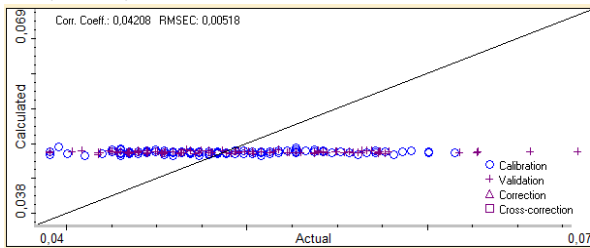
RMSEP: 0.00333 10 principal components used
100,0 percent of variability described

PCR/1DER/CM



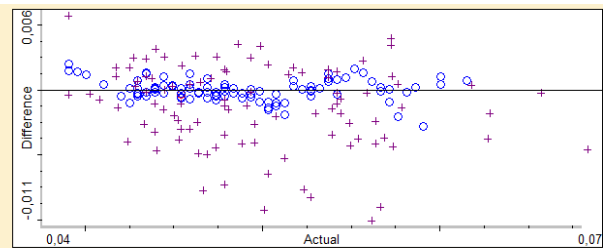
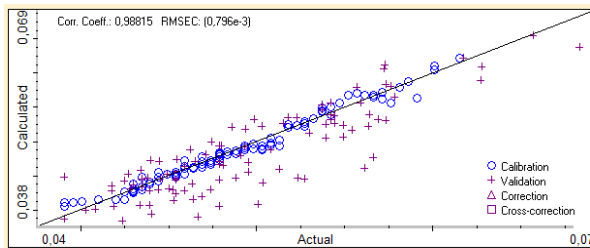
RMSEP: 0.00392 1 principal components used
27,5 percent of variability described

PCR/2DER/CM



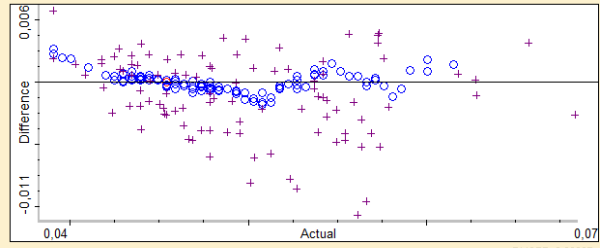
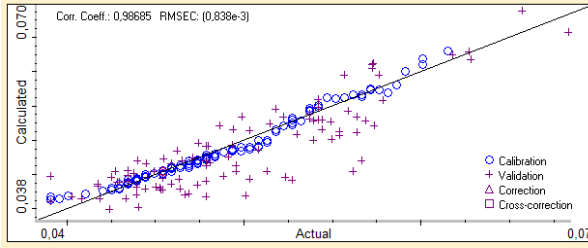
RMSEP: 0.00610 1 principal components used
3,5 percent of variability described

SMRL



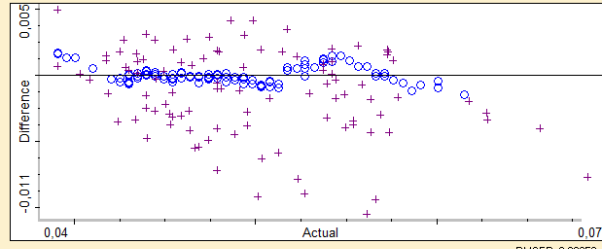
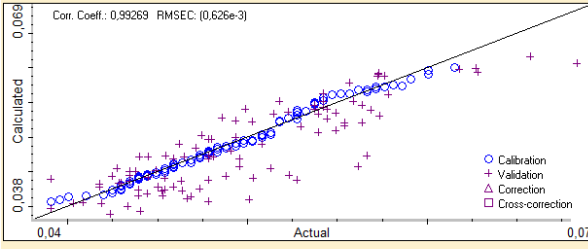
RMSEP: 0.00344
97,6 percent of variability described

SMRL/MSC



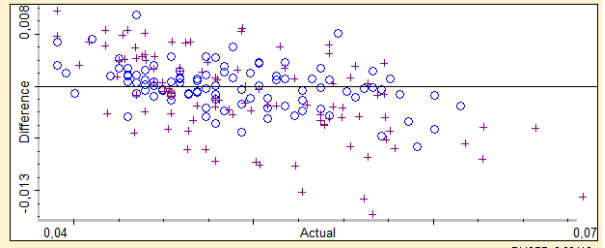
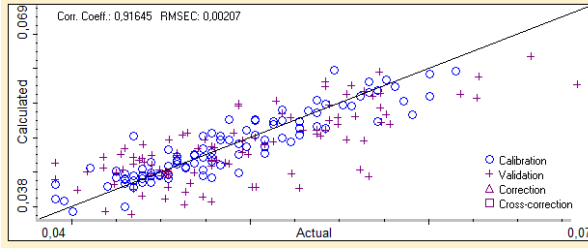
RMSEP: 0.00337
97,4 percent of variability described

SMRL/SNV



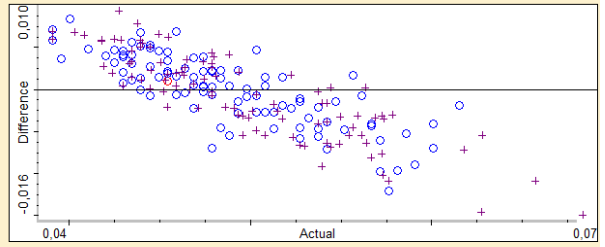
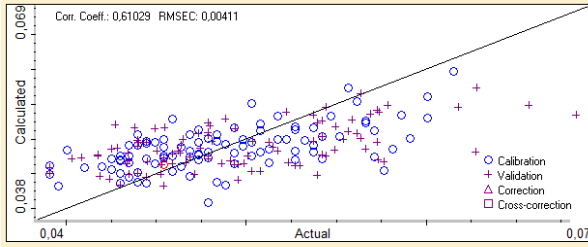
RMSEP: 0.00352
98,5 percent of variability described

SMRL/1DER



RMSEP: 0.00419
84,0 percent of variability described

SMRL/2DER

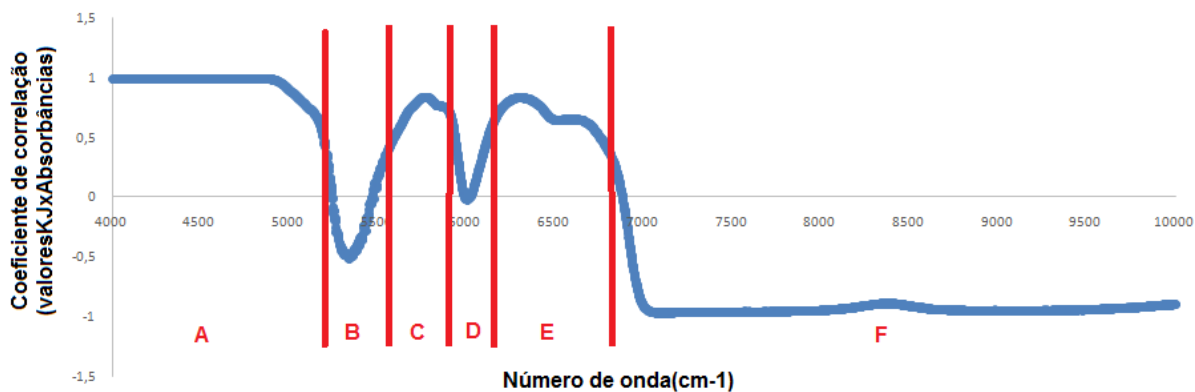


RMSEP: 0.00511
37,2 percent of variability described

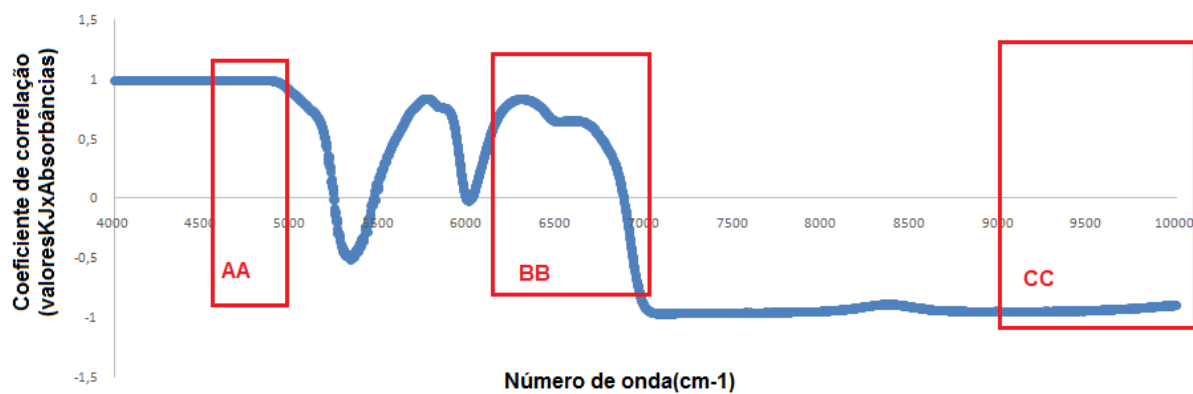
ANEXO C-ESTUDO GRÁFICO DAS REGIÕES

TRATAMENTO SNV

REGIAO A(4000-5195)B(5196-5500)C(5501-5938)D(5939-6090)E(6091-6796)F(6797-10000)

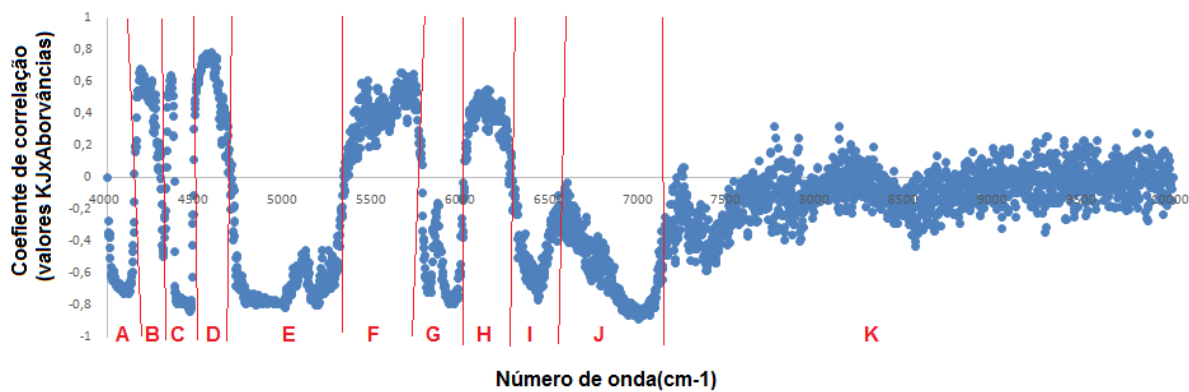


REGIAO AA(4545-5000)BB(6250-7143)CC(9091-10000) regiões da literatura onde encontra-se os estiramentos de NH.

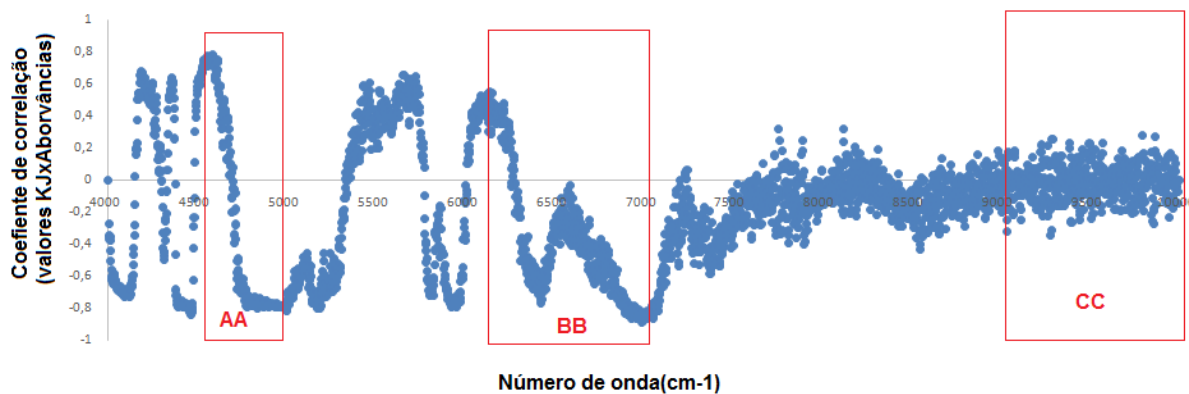


TRATAMENTO 1ª DERIVADA

REGIAO A(4000-4200)B(4201-4300)C(4301-4500)D(4501-4700)E(4701-5500)F(5501-5800)G(5801-6000)H(6001-6250)I(6251-6500) J(6501-7000) K(7001-10000)

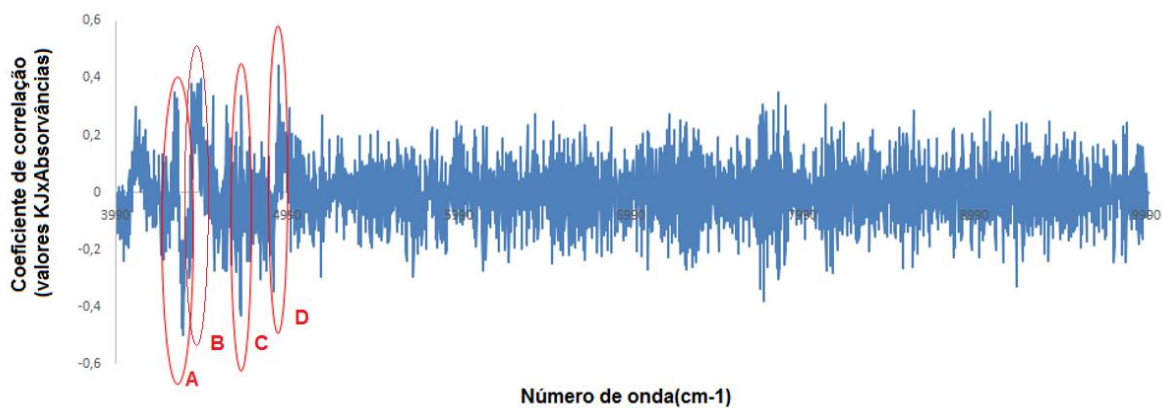


REGIAO AA(4545-5000)BB(6250-7143)CC(9091-10000) regiões da literatura onde encontra-se os estiramentos de NH.



TRATAMENTO 2ª DERIVADA

REGIAO A(4250-4450)B(4451-4550)C(4650-4750)D(4850-4950)



REGIAO AA(4545-5000)BB(6250-7143)CC(9091-10000) regiões da literatura onde encontra-se os estiramentos de NH.

