

**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA
DEPARTAMENTO DE QUÍMICA ANALÍTICA
LAQQA – LABORATÓRIO DE QUIMIOMETRIA EM QUÍMICA ANALÍTICA**



**DETERMINAÇÃO DE PARÂMETROS DE QUALIDADE EM
BANANAS UTILIZANDO ESPECTROSCOPIA NO
INFRAVERMELHO E CALIBRAÇÃO MULTIVARIADA**

DISSERTAÇÃO DE MESTRADO

**AUTORA: LUCIANA VIVIANI GALLO
ORIENTADOR: Prof. Dr. RONEI JESUS POPPI**

CAMPINAS – SP, JUNHO DE 2008

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO INSTITUTO DE
QUÍMICA DA UNICAMP**

G172d	Gallo, Luciana Viviani. Determinação de parâmetros de qualidade em bananas utilizando espectroscopia no infravermelho e calibração multivariada / Luciana Viviani Gallo. -- Campinas, SP: [s.n], 2008. Orientador: Ronei Jesus Poppi. Dissertação- Universidade Estadual de Campinas, Instituto de Química. 1. Banana. 2. Quimiometria. 3. Calibração multivariada. 4. Espectroscopia no infravermelho. I. Poppi, Ronei Jesus. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.
-------	--

Título em inglês: Determination of banana quality parameters by infrared spectroscopy and multivariate analysis

Palavras-chaves em inglês: Banana, Infrared spectroscopy, Multivariate analysis, Chemometrics

Área de concentração: Química Analítica

Titulação: Mestre em Química na área de Química Analítica

Banca examinadora: Ronei Jesus Poppi (orientador), Adriana Vitorino Rossi (IQ UNICAMP), Paulo Augusto da Costa Filho (Nestlé Suíça)

Data de defesa: 18/06/2008

Dedico este trabalho em especial ao

Meu amado esposo Diogo

Meus pais Eclísio e Nanci

E minha irmã Letícia

AGRADECIMENTOS

Ao prof. Dr. Ronei Jesus Poppi pela oportunidade, paciência e orientação.

Aos meus pais, pelo incentivo ao estudo, apoio e incentivo e sem a qual nada seria possível.

A minha irmã Letícia, pela ajuda com o projeto e pela companhia.

Ao meu marido Diogo, pelo estímulo, apoio e suporte, essenciais para a realização desse trabalho.

Aos membros do grupo LAQQA pelo apoio e amizade: Alessandra, Danilo-Joãozinho, Gilmare, Jez, Luiz, Marcello, Patrícia, Paulo Henrique, Thiago, Renato, Bruno, Diórginis, Werickson e Luís.

Às técnicas do Instituto de Química, Cláudia Martelli e Márcia Antonielli pela amizade e pelo auxílio nas análises de infravermelho.

À todos os meus amigos, os quais seria impossível nomeá-los todos aqui, mas que sempre foram fundamentais.

CURRICULO

Informações Pessoais

Nome: Luciana Viviani Gallo
Data de Nascimento: 17/09/1981
Naturalidade: Osasco – SP
e-mail: luciana_81@yahoo.com

Formação

- Pós-Graduação (2004-2008): Mestrado em Química, Universidade Estadual de Campinas (UNICAMP).
- Licenciatura em Química (1999-2004); Universidade Estadual de Campinas - UNICAMP.
- Bacharel em Química (1999-2003): Universidade Estadual de Campinas - UNICAMP.

Experiência Profissional

- 01/2008 – Atual: Colégio São Bento, Vinhedo, SP: Professora no Ensino Médio
- 01/2007 – Atual: Colégio San Conrado, Souza, SP : Professora no Ensino Fundamental II
- 01/2006 – Atual: Colégio Adventista de Campinas, Campinas, SP: Professora no Ensino Médio
- 04/2007 – 02/2008: Curso e Colégio Etapa, Valinhos, SP: Plantonista de dúvidas
- 01/2004 – 12/2008: Instituto de Ensino Sant'Anna, Vinhedo, SP: Professora no Ensino Médio
- 01/2005 – 06/2005: UNICAMP, Campinas, SP: Monitoria da disciplina de Química Geral Teórica do Instituto de Química
- 01/2003 – 06/2003: UNICAMP, Campinas, SP: Monitoria da disciplina de Química Geral Teórica do Instituto de Química

Atividades Acadêmicas

- 01/2001 – 07/2002: UNICAMP - Iniciação Científica, bolsista PIBIQ/CNPq.
Orientador: Prof. Dr. Nelson Eduardo Durán Caballero – Laboratório de Produtos Naturais - Instituto de Química – UNICAMP e Laboratório de Fitoquímica – IAC

Participação em Congressos Nacionais

- VIVIANI, Luciana; NETO, Waldomiro Borges; POPPI, Ronei Jesus. Aplicação da espectroscopia vibracional e PLS na determinação de brix em Banana Prata In: 13º ENCONTRO NACIONAL DE QUÍMICA ANALÍTICA, 2005, Niterói - RJ
- VIVIANI, Luciana; VIVIANI, Leticia; BORIN, Alessandra; POPPI, Ronei Jesus. Aplicação da Espectroscopia no Infravermelho e PLS na determinação do pH em Banana Prata Anã. In: 28ª REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE QUÍMICA, 2005, Poços de Caldas - MG.
- VIVIANI, Luciana; De AZEVEDO, Mariangela de Burgos Martins; DURÁN, Nelson. Síntese da 28-homocastasterona (HCS), importante hormônio vegetal da classe dos Brassinosteróides. Estudo da toxicidade dos precursores sintéticos e HCS. In: 25ª REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE QUÍMICA, 2002, Poços de Caldas - MG

RESUMO

Autora: Luciana Viviani Gallo

Orientador: Ronei Jesus Poppi

Neste trabalho foi desenvolvida uma metodologia analítica alternativa para a determinação de parâmetros de qualidade em bananas: pH, acidez titulável e Brix utilizando as espectroscopias no infravermelho médio e próximo em conjunto com calibração multivariada. A estratégia adotada foi obter os espectros na região do infravermelho próximo e médio do fruto e estabelecer um modelo de calibração multivariada, baseada no método dos mínimos quadrados parciais, para prever os diversos parâmetros.

Foram utilizadas 54 amostras de banana, em diferentes estágios de maturação. Os espectros foram obtidos a partir da análise da polpa do fruto macerada, da solução obtida após a centrifugação da polpa, e pela técnica “Dry Extract Spectroscopy by infrared Absorption” (DESIR).

Um total de 9 modelos foram propostos para a determinação dos parâmetros de qualidade e o melhor resultado para a determinação de pH foi na região do infravermelho médio utilizando a técnica DESIR, com erro médio de 1,32%. Já para a acidez, o melhor modelo obtido foi na região do infravermelho próximo utilizando DESIR, com erro relativo médio de 6,20%. O parâmetro brix apresentou melhores resultados na região do infravermelho médio através da análise da solução, com erro médio de 3,78%. Dessa maneira, a espectroscopia vibracional possibilita viabilizar a obtenção de maior número de informações sobre a qualidade da fruta, através da relação dos espectros com várias propriedades de interesse.

ABSTRACT

Author : Luciana Viviani Gallo

Adviser : Ronei Jesus Poppi

In this work it was developed an alternative analytical methodology for determination of the banana quality parameters as pH, acidity and Brix by using infrared spectroscopy in the mid and near regions in conjunction with multivariate calibration. The strategy adopted was to obtain the fruit spectra in mid and near regions and to establish a multivariate calibration model, based on the partial least squares, to predict the several parameters. It was used 54 banana samples, in different stages of maturation. The spectra were obtained from the pulp fruit, of the solution obtained after pulp centrifugation and by technique Dry Extract Spectroscopy by Infrared Absorption (DESIR). A total of 9 models were proposed for the determination of the quality parameters and the best result for pH determination was in the mid infrared region using the DESIR technique, with mean error of 1.32%. For the acidity parameter, the best model was in the near infrared region using the DESIR technique, with mean error of 6.20%. The Brix model furnished the best result in the mid infrared region by using the solution obtained after pulp centrifugation, with mean error of 3.78%. In this way, the vibrational spectroscopy make possible to achieve large quantity of information about the fruit quality through the relationship between the spectra and several properties of interest.

Índice de Figuras

Figura 1. Sistema de Amostragem por Reflectância Difusa.	5
Figura 2. (a) Reflexão na cela; (b) Reflectância Especular na amostra; (c) Absorção e Reflectância Difusa na amostra); (d) Absorção total pela amostra.	5
Figura 3. Medida da Reflexão Total Atenuada	7
Figura 4. Representação da análise dos componentes principais (A) Pesos são os ângulos do vetor direção; (B) Escores são as projeções das amostras (1-6) na direção da componente principal. Note que os dados são centrados na média.	9
Figura 5. a) espectros originais e b) espectros com a correção com MSC	12
Figura 6. Gráfico do modelo iPLS do número de intervalos versus RMSECV.....	15
Figura 7. Gráficos do modelo iPLS: faixa escolhida para a construção do modelo PLS.	16
Figura 8. Técnica DESIR: Preparação da amostra.....	22
Figura 9. Foto de uma amostra de banana após preparação da amostra utilizando DESIR.	22
Figura 10. a) Espectros originais - MIR Pasta , b) Espectros utilizados para a construção dos modelos: faixa utilizada MIR Pasta– 800-2100 cm^{-1} (Suavização e MSC).	29
Figura 11. a) Espectros originais - MIR DESIR, b) Espectros utilizados para a construção dos modelos : faixa utilizada MIR DESIR – 500 – 5000 cm^{-1} (MSC).....	30
Figura 12. a) Espectros originais - NIR DESIR, b) Espectros utilizados para a construção dos modelos: faixa utilizada DESIR NIR – 820 – 2450 nm (MSC).	31
Figura 13. Gráficos dos valores de referência x valores previstos para os modelos: a) MIR pasta; b) MIR DESIR e c) NIR DESIR.	33
Figura 14. Espectros utilizados para a construção dos modelos MIR pasta. Faixa utilizada 770 – 1780 cm^{-1} (suavização e 1ª.derivada)	36
Figura 15. Espectros utilizados para a construção dos modelos de MIR DESIR. Faixa espectral utilizada: 500 – 5000 cm^{-1} (1ª.derivada).....	36
Figura 16. Resultados do iPLS com 5 intervalos para NIR DESIR (1ª. derivada).37	
Figura 17. Gráficos dos valores reais x valores previstos para os modelos: a) MIR pasta; b) MIR DESIR e c) NIR DESIR.....	38
Figura 18. a) Espectros MR Solução e b) Espectros utilizados para a construção dos modelos: faixa utilizada MR Solução (Derivada e iPLS);.....	41
Figura 19. a) Espectros NIR Solução e b) Espectros utilizados para a construção dos modelos e b) faixa utilizada NIR Solução – 820 – 2180 nm (Suavização)	42
Figura 20. a) Espectros NIR DESIR.b) Espectros utilizados para a construção dos modelos: faixa utilizada DESIR NIR – 1330 - 2410 (MSC).....	43
Figura 21. Gráficos dos valores de referência x valores previstos para os modelos: a) MIR Solução; b) NIR Solução e c) NIR DESIR.	45

Índice de tabelas

Tabela 1. Resultados obtidos para o parâmetro pH.....	28
Tabela 2. Resultados obtidos para o parâmetro acidez.....	35
Tabela 3. Resultados para o parâmetro Brix.....	40

Índice Geral

CAPÍTULO 1 – Introdução	1
1.1. Espectroscopia no Infravermelho.....	3
1.1.1. Espectroscopia no Infravermelho Próximo	4
1.1.2. Espectroscopia no Infravermelho Médio.....	6
1.2. Quimiometria.....	8
1.2.1. Análise de Componentes Principais (PCA).....	9
1.2.2. Método dos Mínimos Quadrados Parciais	10
1.2.3. Pré processamentos	11
1.2.3.1. Correção do Espelhamento Multiplicativo (MSC)	11
1.2.3.2. Aplicação de derivada	12
1.2.3.3. Filtro de Savitsky-Golay (Suavização)	12
1.2.3.4. Dados Centrados na Média	13
1.2.4. Cálculo dos Erros.....	13
1.2.5. PLS por Intervalos (iPLS)	14
CAPÍTULO 2 – Objetivos	17
CAPÍTULO 3 – Parte Experimental.....	19
3.1 Materiais.....	23
3.2. Equipamentos	20
3.3. Métodos de referência	20
3.3.1. Determinação de pH	21
3.3.2. Determinação da acidez titulável	21
3.3.3. Determinação do teor de sólidos solúveis	21
3.4. Medidas espectroscópicas.....	21
3.4.1. Determinação do pH	22
3.4.1.1. Medidas por Espectroscopia no Infravermelho Próximo.....	22
3.4.1.2. Medidas por Espectroscopia no Infravermelho Médio	23
3.4.2. Determinação de Acidez.....	23
3.4.3. Determinação do teor de sólidos solúveis	23
3.4.3.1. Medidas por Espectroscopia no Infravermelho Próximo.....	23
3.4.3.2. Medidas por Espectroscopia no Infravermelho Médio	24
3.5. Tratamento Quimiométrico dos dados	24
CAPÍTULO 4 – Resultados e Discussões	25
4.1. pH	34
4.2. Acidez.....	40
4.3. Brix.....	40
CAPÍTULO 5 – Conclusões.....	47
CAPÍTULO 6 – Referências Bibliográficas	49

CAPÍTULO 1 – Introdução

1. Introdução

A banana, fruto de considerável importância sócio-econômica nos países tropicais, constitui-se numa fonte de calorias, vitaminas e minerais, além de baixo custo, e elevado consumo nas diversas camadas da sociedade. Devido às suas agradáveis peculiaridades de aroma e sabor, é consumida tanto na forma *in natura* como processada, por populações de baixo e alto poder aquisitivo¹. É cultivada em mais de 80 países tropicais e é o quarto alimento mais consumido em todo o mundo.

O Brasil é o terceiro maior produtor mundial e a banana é a segunda fruta mais consumida no país (cerca de 6,9 milhões de toneladas no ano de 2007)². O consumo médio anual no Brasil é estimado em 24 kg per capita e é consumido principalmente *in natura*, em razão do preço geralmente acessível³. A região Nordeste é responsável por 30% da produção de bananas, sendo considerada a maior produtora, seguida pela região Norte 27,2%, Sudeste 25,5%, Sul 9,6% e Centro-Oeste 7,7%^{4,5}.

A exportação ocorre apenas para os produtos de qualidade superior, sendo que os principais parâmetros para avaliar a qualidade física da banana são a aparência (forma), peso e relação polpa/casca. Dentre os índices físico-químicos mais utilizados para avaliar a qualidade pós-colheita da banana estão o pH, acidez titulável, sólidos solúveis, relação sólidos solúveis/acidez, açúcares redutores, açúcares não redutores, açúcares totais, substâncias pécicas, teor de amido⁶.

Durante o amadurecimento a transformação mais relevante que se observa é a transformação de amido em açúcares pelo mecanismo de hidrólise enzimática. O teor de amido decai de 20 a 23% para 1 a 2% e, simultaneamente, a percentagem de açúcares solúveis, aumenta e 1 para 20%, sendo que estes valores variam conforme o cultivar^{7,8}. Na polpa da banana os principais açúcares encontrados são glicose, frutose e sacarose. A quantidade de amido encontrado é igual a 17,4% no fruto verde e 5,2% no fruto maduro, sendo que em banana Prata

madura foi encontrado 1,8% de glicose, 1,7% de frutose e 16% de sacarose. Em cultivares Prata-Anã o teor de amido no fruto verde foi de 18,7% e 2,5% no fruto maduro, sendo que os teores de glicose, frutose e sacarose encontrados foram 2,4%, 2,4% e 9,1%, respectivamente⁹.

Atualmente, principalmente no Brasil, as análises químicas das bananas utilizadas para o acompanhamento do amadurecimento de bananas são realizadas por métodos via-úmida, em muitos laboratórios de controle de alimentos e de pesquisa⁹. Essas análises, são demoradas e difíceis de realizar. Por isso, tem-se buscado alternativas instrumentais que as tornem mais práticas e simples. Nesse contexto, insere-se a espectroscopia no Infravermelho. Nos últimos anos, a espectroscopia no Infravermelho tem sido utilizada cada vez mais para o controle de qualidade em produtos alimentícios e agrícolas, para avaliação dos mais diferentes parâmetros, como quantidade de açúcares totais, individuais¹⁰, redutores, amido, sólidos solúveis e acidez¹¹.

1.1. Espectroscopia no Infravermelho

A condição para que ocorra absorção da radiação infravermelha é que haja variação do momento de dipolo elétrico da molécula como consequência de seu movimento vibracional ou rotacional. O momento de dipolo é determinado pela magnitude das diferenças de carga e distância entre dois centros de carga. Somente nestas circunstâncias, o campo elétrico alternante da radiação incidente interage com a molécula, originando espectros de absorção. De outra forma, pode-se dizer que o espectro de absorção no infravermelho tem origem quando a radiação eletromagnética incidente tem um componente com valor de número de onda correspondente a uma transição entre dois níveis vibracionais.

As vibrações podem ser de duas categorias: estiramento e deformação. As vibrações de estiramento envolvem uma mudança contínua na distância interatômica ao longo do eixo da ligação entre dois átomos. Vibrações de deformação são caracterizadas pela mudança no ângulo entre duas ligações.

A região do infravermelho abrange a faixa de 12800 a 10 cm^{-1} ou de $0,78$ a $1000\text{ }\mu\text{m}$. Esta faixa é dividida em radiação no infravermelho próximo, médio e distante. A maior parte das aplicações tem sido realizada no Infravermelho médio e próximo, amplamente utilizadas para análises qualitativas e quantitativas.¹²

1.1.1. Espectroscopia no Infravermelho Próximo

O campo da espectroscopia no infravermelho próximo (NIR) está crescendo rapidamente e cada vez mais aplicações são propostas em laboratórios acadêmicos e industriais, por exemplo, no controle de qualidade de produtos agrícolas e alimentícios¹³. A técnica é baseada na interação da radiação do infravermelho próximo (região situada entre 750 e 2650 nm), com a matéria, observando-se, essencialmente, sinais provenientes de sobretons e combinações vibracionais de C-H, O-H e N-H.

Esta crescente utilização deve-se ao fato de que praticamente todos os compostos orgânicos absorvem no infravermelho próximo, e amostras líquidas usualmente seguem a Lei de Beer-Lambert. Em relação aos métodos tradicionais de análise, a espectroscopia NIR possui como vantagens o baixo tempo de processamento da análise, baixo custo (quando muitas amostras são analisadas), não necessidade de tratamento da amostra (reduzindo gastos com reagentes e tratamento de resíduos), pequenas quantidades de amostra utilizada, estado físico de agregação indiferente.

Pode-se medir a absorção da radiação pela matéria como a reflectância difusa proveniente da sua penetração nas camadas das partículas. Espectros de reflectância difusa podem ser obtidos diretamente de amostras sólidas, facilitando muito a manipulação e obtenção dos espectros. A Figura 1 mostra um esquema de uma medida por reflectância difusa.

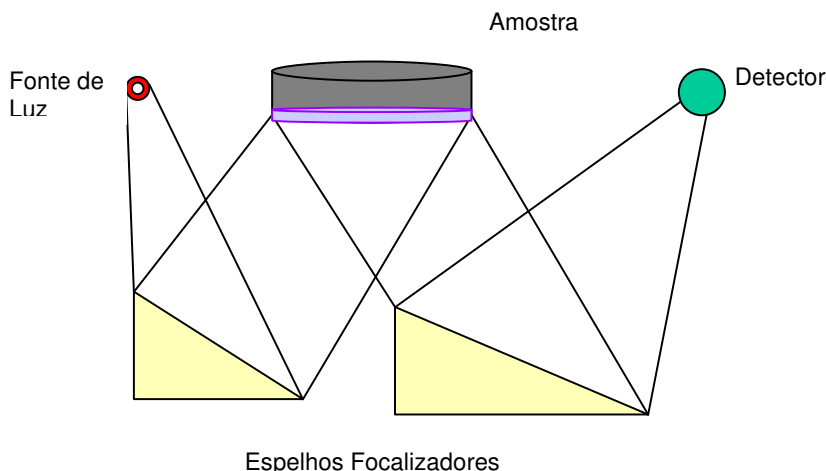


Figura 1. Sistema de Amostragem por Reflectância Difusa.

Para análise por reflectância difusa é necessário um sistema que a separe da reflectância especular. A reflectância especular está associada à imagem refletida pela lâmpada de tungstênio (fonte de luz), enquanto que a reflectância difusa está associada com a detecção do sinal após a radiação ter interagido com a amostra. Também podem ocorrer reflexão da radiação pelas paredes da cela e absorção total da radiação pela amostra quando a radiação atinge a amostra sólida. No caso da reflectância difusa, após a radiação incidir na amostra sólida, ocorre pequena penetração e absorção em comprimentos de onda específicos que depende da amostra. A radiação que é espalhada em todas as direções chega até o detector. A Figura 2 mostra os diferentes processos que podem ocorrer, quando a radiação incide sobre uma amostra sólida.

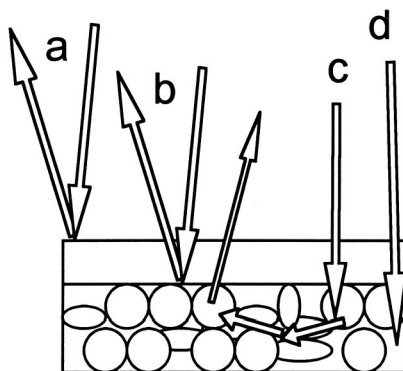


Figura 2. (a) Reflexão na cela; (b) Reflectância Especular na amostra; (c) Absorção e Reflectância Difusa na amostra; (d) Absorção total pela amostra.

A sensibilidade, utilizando-se a espectroscopia NIR, é geralmente adequada para amostras com concentração de 0,5% ou mais, embora muitas referências reportam as análises de quantidades mais baixas de analito.

Para análise de amostras sólidas, é necessário que as amostras estejam uniformes em termos de distribuição de constituintes e de tamanho de partícula. A importância da uniformidade do tamanho da partícula afeta a quantidade de reflectância, contribuindo para a observação do espectro.

O uso da espectroscopia NIR como ferramenta quantitativa se deve ao uso de procedimentos estatísticos multivariados, ou quimiométricos, principalmente aqueles baseados na calibração multivariada.

Como aplicações podemos citar o uso de espectroscopia NIR para determinação de açúcares individuais¹⁴ e totais¹⁵, Brix^{16,17} em frutas *in natura* e em sucos¹⁸.

1.1.2. Espectroscopia no Infravermelho Médio

A espectroscopia no infravermelho médio também tem sido utilizada em determinações quantitativas, principalmente com o avanço de técnicas de amostragem como reflexão total atenuada (ATR).

A técnica ATR permite obter espectros no infravermelho médio de amostras como sólido de baixa solubilidade, filmes, adesivos, pós, pastas, entre outros.

O princípio deste tipo de espectroscopia baseia-se no fato de que quando um feixe de radiação, passando por um meio mais denso (cristal ATR), transparente ao infravermelho, para um meio menos denso (amostra) ocorre reflexão. A fração do feixe incidente que é refletida aumenta conforme o ângulo de incidência, e quando excede um determinado ângulo crítico (α) a reflexão é completa. No ponto de reflexão, o feixe é atenuado como se penetrasse a uma pequena distância dentro da amostra. A profundidade de penetração (d_p) é dada pela equação 1:

$$d_p = \frac{\lambda}{2\pi n_1 \sqrt{\sin^2 \theta - (n_2 / n_1)^2}} \quad (1)$$

onde:

n_2 : maior índice de refração (cristal ATR);

n_1 : menor índice de refração (amostra);

θ : é o ângulo de incidência;

λ : é o comprimento de onda da radiação.

Assim, a profundidade da penetração, depende do comprimento de onda, do índice de refração dos dois materiais e do ângulo do feixe em relação a interface. A radiação penetrada é chamada de onda evanescente. Se o meio menos denso absorve a radiação evanescente, a atenuação do feixe ocorre nos comprimentos de onda de absorção de bandas. Este fenômeno é chamado de Reflexão Total Atenuada (ATR).

A Figura 3 mostra um esquema para medida da reflexão total atenuada.

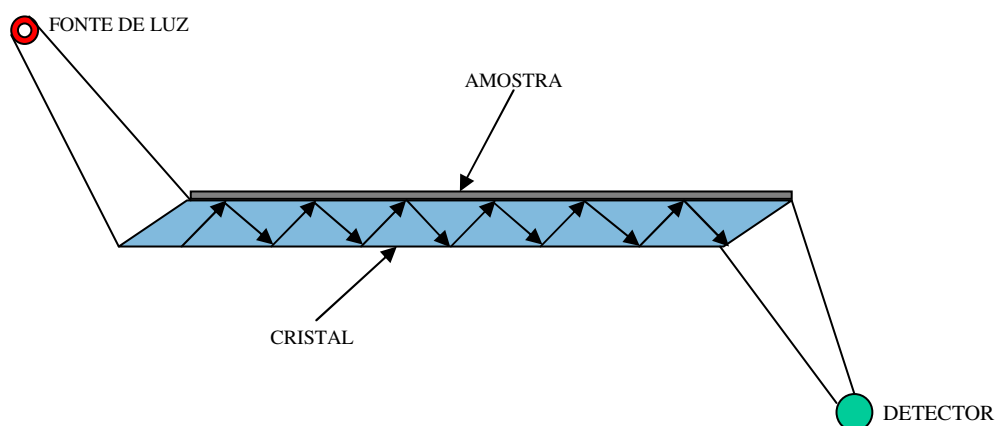


Figura 3. Medida da Reflexão Total Atenuada

Os espectros de ATR são similares, mas não idênticos aos espectros de absorção comum. Em geral, enquanto os mesmos picos são observados, suas intensidades relativas diferem. As absorbâncias, embora dependentes do ângulo de incidência, são independentes da espessura da amostra, pois a radiação interage apenas poucos micrômetros com a amostra.

Uma das maiores vantagens da espectroscopia de reflectância interna é que o espectro de absorção é facilmente obtido para uma ampla variedade de amostras com um mínimo de preparação.^{12,19-21}

1.2. Quimiometria

Há mais de 20 anos, a quimiometria²⁰ vem aplicando conhecimentos estatísticos matemáticos e computacionais na Química, com o objetivo de planejar e otimizar experimentos e procedimentos, bem como, extrair o maior número de informações possível a partir dos dados químicos. Dentro da Quimiometria, uma área muito aplicada atualmente é a calibração multivariada. A calibração multivariada¹² é baseada em desenvolver um modelo matemático que possa correlacionar uma propriedade de interesse (**y**) com muitas medidas instrumentais (**X**), de um espectro, por exemplo, simultaneamente.

A calibração multivariada consiste na execução de dois passos principais: calibração e validação.

Para a calibração são necessárias múltiplas medidas instrumentais, formando uma matriz **X**, obtidas para inúmeros padrões de uma ou mais espécies de interesse. Os valores obtidos poderiam ser, por exemplo, a absorbância de vários padrões a vários comprimentos de onda. A partir desses dados, cria-se um modelo matemático que melhor possa reproduzir a matriz de resposta **Y** a partir dos dados obtidos para a matriz **X**.

O método de calibração multivariada mais utilizado e atualmente considerado como padrão na área, é o método dos mínimos quadrados parciais ou PLS²¹ (do inglês “Partial Least Squares”).

Na previsão, a amostra de interesse é medida instrumentalmente e o modelo desenvolvido utilizando o PLS é utilizado para a quantificação do analito.

1.2.1. Análise de Componentes Principais (PCA)

A análise de componentes principais (PCA – Principal Component Analysis) consiste na aproximação de uma tabela de dados, ou seja, de uma matriz **X**, em termos da soma de várias matrizes **M** de posto 1 (do inglês rank), no qual posto significa um número que expressa a verdadeira dimensionalidade da matriz.^{22,23}

$$\mathbf{X} = \mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 + \dots \mathbf{M}_a \quad (2)$$

Essas matrizes de posto 1 podem ser escritas como produtos de dois vetores: o escore **t** e o peso **p**:

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1 + \mathbf{t}_2 \mathbf{p}_2 + \dots \mathbf{t}_a \mathbf{p}_a \quad (3)$$

outra maneira de escrever a equação acima é:

$$\mathbf{X} = \mathbf{T} \mathbf{P}^T \quad (4)$$

onde **T** é a matriz de todos os valores **t** e **P** é a matriz de todos os valores **p**.

Na Figura 4 tem-se a representação gráfica da Análise de Componentes Principais.

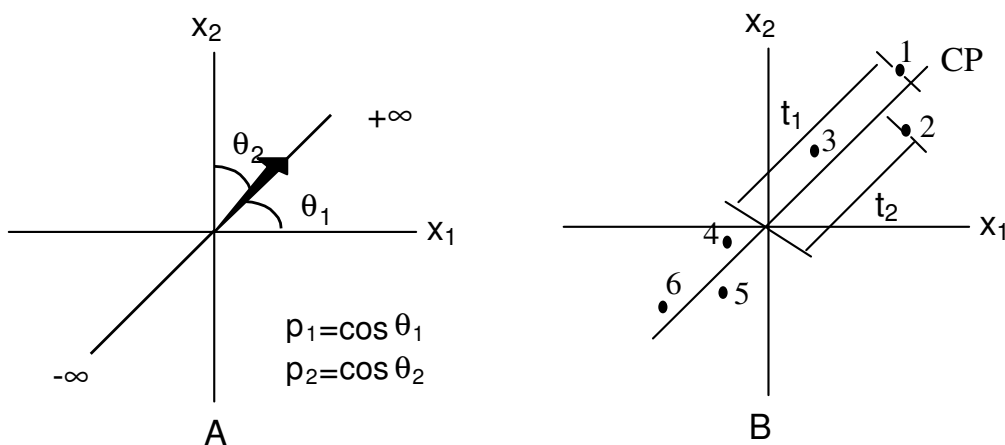


Figura 4. Representação da análise dos componentes principais (A) Pesos são os ângulos do vetor direção; (B) Escores são as projeções das amostras (1-6) na direção da componente principal. Note que os dados são centrados na média.

Os principais objetivos desta técnica são o de encontrar relações entre objetos, e o de classificá-los de acordo com suas similaridades. Associado a isto está a detecção de amostras anômalas, que não pertencem a nenhuma das categorias conhecidas. Outro objetivo importante é a redução de dados, que se torna muito útil quando grandes quantidades de informações necessitam ser manipuladas.

1.2.2. Método dos Mínimos Quadrados Parciais

O PLS é um método matemático baseado em uma relação linear entre as variáveis instrumentais ($\mathbf{x}$) e as variáveis de interesse ($\mathbf{y}$).^{24,25}

As matrizes destas duas variáveis ($\mathbf{X}$ e $\mathbf{Y}$) podem ser representadas pela Análise de Componentes Principais:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (5)$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} \quad (6)$$

onde $\mathbf{T}$ e $\mathbf{U}$ são as matrizes de escores, $\mathbf{P}$ e $\mathbf{Q}$ são as matrizes de pesos das matrizes $\mathbf{X}$ e $\mathbf{Y}$, respectivamente, e $\mathbf{E}$ e $\mathbf{F}$ são os resíduos.

Uma relação linear entre os dois blocos pode ser realizada correlacionando os escores para cada componente de cada vez, utilizando-se um modelo linear:

$$\mathbf{u}_h = b_h \mathbf{t}_h \quad (7)$$

onde “h” é o número do componente principal.

Sendo que:

$$b_h = \mathbf{u}_h \mathbf{t}_h / \mathbf{t}_h^T \mathbf{t}_h, \quad (8)$$

representa coeficiente de regressão linear para o modelo.

1.2.3. Pré processamentos

A espectroscopia no infravermelho para análises quantitativas está sendo muito utilizada atualmente, visto o grande número de trabalhos publicados. Porém, apesar da espectroscopia no infravermelho fornecer um grande número de dados, parte destes não possuem informações correlacionadas diretamente com o(s) analito(s) de interesse, o que pode ocasionar distorções ao modelo, e conseqüentemente a conclusões errôneas nas análises²⁶.

Como solução para minimizar este problema têm sido utilizados diversos métodos de pré-processamento, como correção de espalhamento multiplicativo, cálculo da derivada²⁷ e seleção de variáveis, o qual permite eliminar os termos que não são relevantes na modelagem, gerando um subconjunto com um menor número de variáveis, e que apresente maior sensibilidade e linearidade para o(s) analito(s) de interesse²⁸.

1.2.3.1. Correção do Espalhamento Multiplicativo (MSC)

Uma alternativa para a correção da linha da base dos espectros é a correção do espalhamento multiplicativo do inglês “Multiplicative scatter correction, MSC”²⁹. Esta correção faz uma regressão de um espectro medido contra um espectro de referência (geralmente o espectro médio) e corrige o espectro medido usando a inclinação e coeficiente linear da reta obtida. Para cada espectro individual x_i , parâmetros a_i e b_i são estimados por regressão em x (espectro de referência/médio). Os parâmetros estimados a_i e b_i são utilizados para realizar a correção do espectro da seguinte maneira:

$$x = a_i + x_i b_i \quad (9)$$

$$x_{\text{corrigido}} = \frac{(x_i - a_i)}{b_i} \quad (10)$$

A Figura 5 mostra o efeito da correção com o MSC sobre espectros de infravermelho com diferentes linha-base.

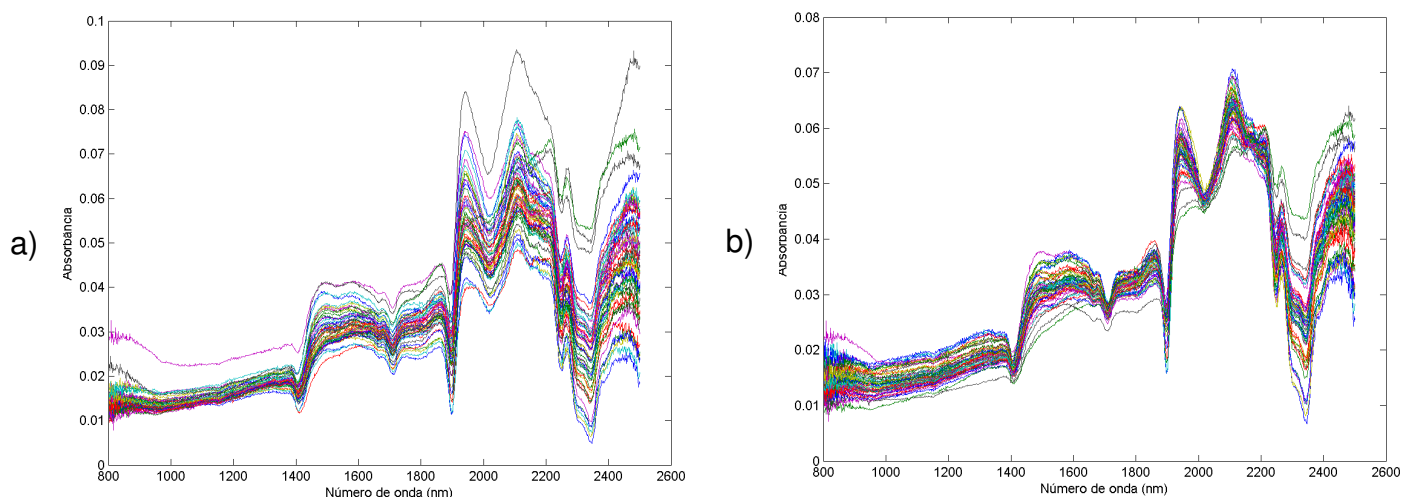


Figura 5. a) Espectros originais e b) Espectros com a correção com MSC

1.2.3.2. Aplicação de derivada

A aplicação da primeira ou segunda derivada sobre os dados espectrais brutos é um procedimento que pode destacar ombros espectrais, bem como minimizar o efeito de inclinações provocadas na linha de base dos espectros, devido à morfologia das partículas.^{27,28,30,31}

Ao aplicarmos as operações de derivação aos espectros, as informações contidas ao longo dos diferentes comprimentos de onda são geralmente acentuadas. Não só os sinais espectrais, mas também os ruídos tornam-se acentuados, portanto, deve-se ter cuidado com a qualidade dos espectros com os quais se deseja aplicar o cálculo das derivadas³¹.

1.2.3.3. Filtro de Savitsky-Golay (Suavização)

Basicamente, o filtro de Savitsky-Golay ajusta um polinômio a certa janela com um número fixo de pontos do espectro. O valor dado pelo polinômio no ponto de interesse corresponde ao valor filtrado. Em seguida, há um deslocamento de um ponto nesta janela, ajustando-se outro polinômio (da mesma ordem), e assim sucessivamente ocorre o deslocamento até o final do espectro.²⁶

Os filtros exponenciais e de média deslizante consideram aproximações lineares ao sinal. Contudo alguns sinais podem ser mais bem modelados com aproximações quadráticas ou cúbicas²⁶.

1.2.3.4. Dados Centrados na Média

Consiste em fazer com que para cada variável seus valores tenham média zero. Para centrar os dados na média, obtêm-se para cada coluna o valor médio e, em seguida, subtrai-se este valor de cada variável dessa mesma coluna. Desta forma, ocorre a mudança do sistema de coordenadas para o centro dos dados. A equação 11 é utilizada para centrar os dados na média^{26, 30}.

$$X_{(i,j)cm} = X_{(i,j)} - \bar{X}_j \quad (11)$$

onde, $X_{(i,j)cm}$ corresponde ao valor centrado na média para a variável j na amostra i ; $X_{(i,j)}$ é o valor da variável j na amostra i ; e $\bar{X}_j$ é a média das amostras na coluna j , calculado pela equação 12²⁶:

$$\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{(i,j)} \quad (12)$$

Na equação 12, n é o número de amostras.

1.2.4. Cálculo dos Erros

A eficiência dos modelos de calibração multivariada pode ser avaliada pelo cálculo dos valores de erro quadrático médio (RMSE – do inglês, *root mean square error*) e erro relativo (E).²⁷

Tais valores expressam a exatidão do modelo, ou seja, a proximidade entre o valor calculado pelo modelo (y_{prev}) e o valor verdadeiro ou obtido por um método de referência (y_{real}). Os erros são definidos como:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_{prev} - y_{real})^2}{N}} \quad (13) ,$$

sendo N o número de amostras.

$$E = \sqrt{\frac{\sum_{i=1}^N (y_{prev} - y_{real})^2}{(y_{real})^2}} \times 100\% \quad (14)$$

RMSEC é o RMSE calculado a partir das amostras de calibração, isto é, uma medida do erro na modelagem. RMSECV a partir das amostras da validação cruzada e RMSEP das amostras do conjunto de previsão²⁷.

1.2.5. PLS por Intervalos (iPLS)

O método iPLS é uma extensão iterativa para o PLS, que desenvolve modelos locais PLS em subintervalos eqüidistantes de toda a região do espectro.

O principal objetivo deste método é prever informação relevante nas diferentes subdivisões do espectro global, de forma a remover as regiões espectrais cujas variáveis se apresentam como supostamente de menor relevância e/ou interferentes. A partir deste ponto, um novo modelo PLS é construído a partir das variáveis selecionadas.³²

Os modelos locais PLS para cada região do espectro subdividido são comparados com o modelo PLS desenvolvido para todo o espectro (modelo global) em um gráfico baseado no parâmetro de validação RMSECV (erro quadrático médio da validação cruzada), calculado para cada um dos modelos.

Outros parâmetros são também utilizados como R^2 (coeficiente de correlação) e inclinação da reta do gráfico dos valores reais e previstos pelo modelo. Amostras e/ou medidas anômalas detectadas pelo PLS devem ser geralmente removidas antes da aplicação do iPLS³².

As regiões espectrais com valores de parâmetro de validação menor que o modelo global são selecionadas para a construção de modelos PLS que são comparados com a eficiência do modelo PLS global.

As Figuras 6 e 7 exemplificam para o modelo de calibração de amostras de banana, para o parâmetro acidez na região do NIR, os gráficos do iPLS. Nessa seleção de variáveis foi utilizado o programa iPLS (Lars Norgaard, Chemometrics Group - KVL, Dinamarca, 2004) do iToolbox para Matlab 6.1 (MathWorks).

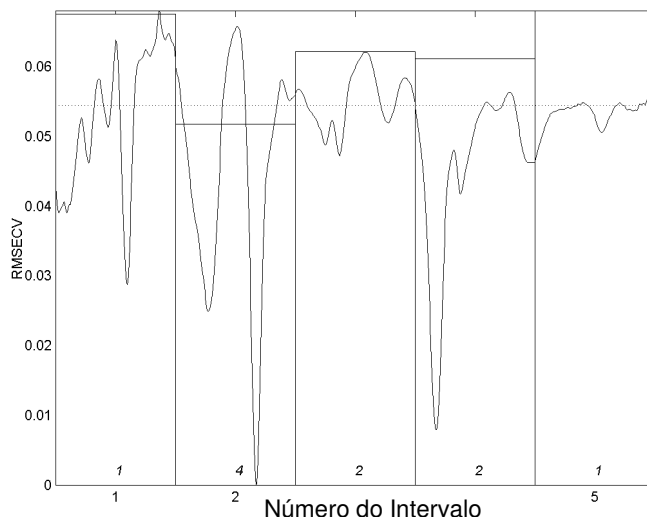


Figura 6. Gráfico do modelo iPLS do número de intervalos versus RMSECV.

Na Figura 6, é apresentada a divisão dos dados em 5 subgrupos. Para cada subgrupo é desenvolvido um modelo de PLS, e os resultados de RMSECV são mostrados como barras na Figura. Os números no interior das faixas no gráfico da Figura 6, são os números de LVs indicados pelo modelo iPLS para serem usados na construção dos modelos PLS com as variáveis selecionadas em cada intervalo. A linha horizontal tracejada indica o valor de RMSECV para o modelo global (utilizando todas as variáveis). Pode-se notar que o 2º intervalo apresenta o menor valor de RMSECV e por isso selecionado para o desenvolvimento do modelo de calibração. Na Figura 7, é apresentada a faixa das variáveis escolhidas (200 a 380).

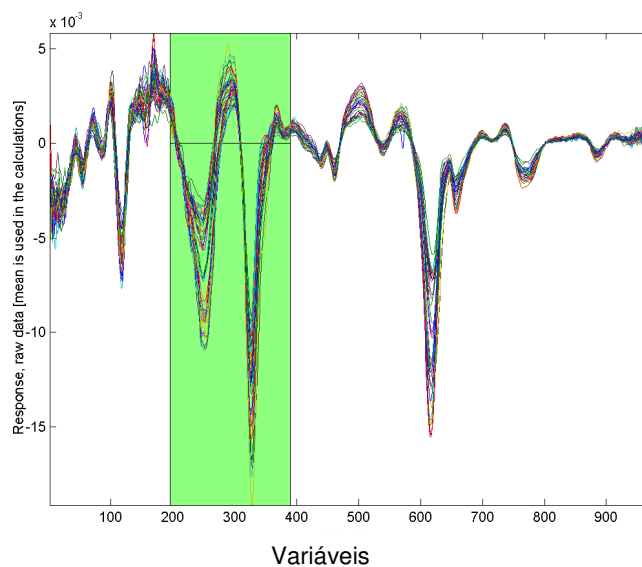


Figura 7. Gráficos do modelo iPLS: faixa escolhida para a construção do modelo PLS.

Assim, no exemplo do modelo de calibração de amostras de banana, para o parâmetro acidez na região do NIR, o modelo PLS deve ser construído a partir de 180 variáveis do espectro original (variáveis 200 a 380), ao invés das 900 variáveis iniciais.

CAPÍTULO 2 – Objetivos

2. Objetivos

Nesta dissertação foi desenvolvida uma metodologia analítica alternativa para a determinação de parâmetros de qualidade em bananas: pH, acidez titulável e Brix utilizando espectroscopia no infravermelho médio e próximo, calibração multivariada e seleção de variáveis. A estratégia adotada foi obter os espectros no infravermelho próximo e médio do fruto e estabelecer um modelo de calibração multivariada utilizando-se o PLS e iPLS para prever os diversos parâmetros.

CAPÍTULO 3 – Parte Experimental

3. Parte Experimental

3.1. Materiais

Os frutos foram adquiridos da Central de Abastecimento de Campinas - CEASA – provenientes de produtores do Estado de São Paulo, região de Avaré. A variedade utilizada foi a banana Prata Anã. Para cada parâmetro analisado foram utilizadas 54 amostras, sendo que cada amostra foi preparada a partir de uma mesma banana.

3.2. Equipamentos

Para as análises no Infravermelho médio utilizou-se um espectrômetro BOMEM, MB-100, equipado com acessório de reflectância total atenuada horizontal de Seleneto de Zinco; e para o Infravermelho próximo, foi utilizado um espectrofotômetro UV-Vis-NIR CARY 5G dispersivo da marca VARIAN, equipado com acessório de esfera de integração para medidas de reflectância difusa.

Os espectros no infravermelho médio foram obtidos na faixa de 600 a 5000 cm^{-1} , com resolução de 4 cm^{-1} e 64 varreduras por espectro. Para a região do infravermelho próximo, os espectros foram obtidos na faixa de 800 a 2500 nm, com resolução de 0,1 nm e média de 10 espectros. Dependendo da aplicação, foram selecionadas diferentes regiões espectrais que serão especificadas na seção de Resultados e Discussões.

3.3. Métodos de referência

Para o desenvolvimento dos modelos de calibração multivariada, foram necessários na fase de calibração, amostras de banana analisadas pelos métodos padrão (ou de referência).

As metodologias empregadas nas análises de pH, acidez titulável e sólidos solúveis estão descritas a seguir.

3.3.1. Determinação de pH

O pH³³ foi determinado através do uso de pHmetro digital. Como o método é baseado na determinação hidrogeniônica (pH) utilizando o potenciômetro, foi necessário calibrá-lo usando soluções tampão de pH igual 4,00 e 7,00 a 20 °C. O resultado é expresso em unidade de pH, acompanhado de décimos, de acordo com a escala do potenciômetro. O eletrodo foi imerso na pasta macerada de banana, e a leitura foi realizada após a estabilização do valor medido.

3.3.2. Determinação da acidez titulável

A acidez titulável³⁴ foi determinada e calculada como volume em mililitros de NaOH 0,1000M requeridos para titular a solução filtrada de 10g de polpa de banana macerada diluída em 90 mL de água, expresso também como porcentagem de ácido málico.. .

3.3.3. Determinação do teor de sólidos solúveis

A determinação de sólidos solúveis³³ foi realizada por refratometria, utilizando refratômetro de Abbe. Após a maceração, as amostras foram centrifugadas e colocou-se duas gotas do sobrenadante obtido entre os prismas do refratômetro e fez-se a leitura em temperatura constante de 20 °C, em duplicata. Para as leituras feitas em temperaturas diferentes da citada, utilizou-se a tabela de correção¹².

3.4. Medidas espectroscópicas

Para a obtenção de espectros sem interferência relacionada à presença de água, as amostras foram preparadas de acordo com a técnica “Dry Extract Spectroscopy by infrared Absorption” (DESIR)^{11, 25} onde uma pequena alíquota de solução preparada com a polpa da banana é aplicada em membrana de fibra de

vidro 0,45 μm e seca em estufa, para eliminação de água, por aproximadamente 1 hora a 60°C (Figuras 8 e 9).

Realizou-se também a análise direta da polpa macerada (pH e acidez titulável) e da solução obtida da centrifugação da polpa (sólidos solúveis).

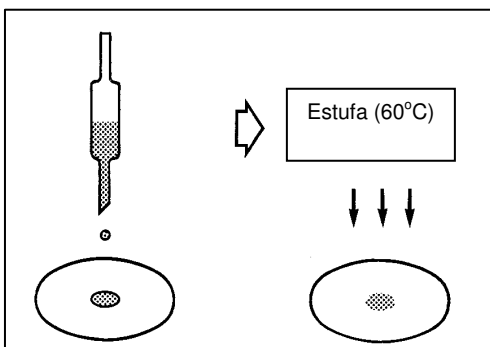


Figura 8. Técnica DESIR: Preparação da amostra

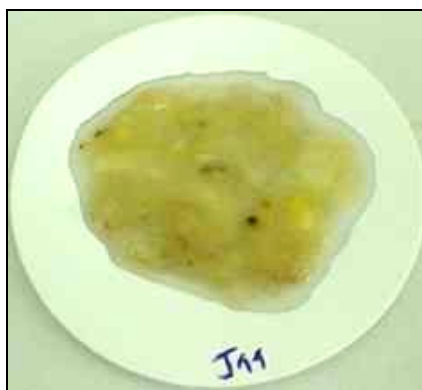


Figura 9. Foto de uma amostra de banana após preparação da amostra utilizando DESIR.

3.4.1. Determinação do pH

3.4.1.1. Medidas por Espectroscopia no Infravermelho Próximo

DESIR: Foram obtidos espectros de reflectância difusa das amostras.

3.4.1.2. Medidas por Espectroscopia no Infravermelho Médio

Polpa: As medidas de absorção foram realizadas com a disposição da polpa de banana macerada sobre o cristal de ATR.

DESIR: Foram obtidos espectros de reflectância difusa das amostras. Para isso, recortou-se a membrana de fibra de vidro de acordo com o tamanho necessário para a análise.

3.4.2. Determinação de Acidez

As análises referentes à determinação da acidez titulável e pH foram feitas a partir das mesmas amostras, sendo assim os procedimentos para preparação e análise das amostras são os mesmos que os descritos anteriormente.

3.4.3. Determinação do teor de sólidos solúveis

3.4.3.1. Medidas por Espectroscopia no Infravermelho Próximo

Líquido: As medidas de absorção foram realizadas com a disposição do sobrenadante entre placas chanfradas de quartzo, de caminho óptico 1 mm.

DESIR: Após a preparação da amostra, foram obtidos espectros de reflectância difusa.

3.4.3.2. Medidas por Espectroscopia no Infravermelho Médio

Líquido: As medidas de absorção foram realizadas com a disposição do líquido sobre o cristal de ATR.

3.5. Tratamento Quimiométrico dos dados

Para o tratamento quimiométrico dos dados foi utilizado o pacote computacional “PLS – Toolbox 3.5 for use with MATLAB” da empresa “Eigenvectors Research” – USA.

CAPÍTULO 4 – Resultados e Discussões

4. Resultados e Discussões

Os resultados apresentados foram obtidos a partir de 54 amostras, sendo que cada amostra consistia de uma banana com diferentes estágios de amadurecimento. Medidas em duplicata foram obtidas para o parâmetro Brix e em triplicata para os parâmetros pH e acidez. As análises de pH e acidez foram realizadas simultaneamente.

Para os parâmetros pH e acidez, modelos de calibração multivariada baseados no PLS foram desenvolvidos para espectros no infravermelho médio obtidos diretamente da polpa e com a técnica DESIR e para espectros no infravermelho próximo obtidos com a técnica DESIR. Nesse caso, com o espectrofotômetro utilizado, espectros da polpa não foram possíveis de serem obtidos com reprodutibilidade adequada. Para o parâmetro Brix, devido à natureza da propriedade que estava sendo monitorada, espectros foram tomados nas regiões do infravermelho próximo e médio de soluções obtidas da centrifugação da polpa de cada banana. Também espectros na região do infravermelho próximo das amostras obtidas pela técnica DESIR foram correlacionados com o parâmetro Brix.

Para cada conjunto de espectros devido aos desvios que ocorrem por ruídos, variação da linha base e espalhamento da radiação, aplicou-se ferramentas de pré-processamento espectral (suavização, 1° derivada ou MSC) e seleção de variáveis (iPLS ou seleção manual), com a finalidade de obter o melhor modelo de calibração/validação. Os conjuntos de calibração/validação foram determinados através do algoritmo Kennard-Stone³⁵, onde as amostras são escolhidas baseadas nas distâncias euclidianas na dimensão das variáveis espectrais. Primeiro é escolhida a amostra mais distante do centro dos dados. A próxima amostra a ser escolhida é a que tem a maior distância em relação a esta última. O processo segue até que sejam selecionadas todas as amostras da calibração.

As tabelas 1, 2 e 3 trazem as informações relativas aos melhores modelos obtidos, sendo elas: os pré-processamentos utilizados, o número de variáveis

latentes, número de amostras de calibração e validação utilizadas para a construção dos modelos e valores do coeficiente de correlação da reta obtida graficando-se o valor esperado contra o valor previsto pelo PLS (R^2). A análise dos modelos baseou-se nos valores de erro médio (%), da raiz quadrada do erro quadrático médio da previsão (RMSEP) e do R^2 .

4.1. pH

A faixa de pH analisada variou de 4,01 a 4,66 e a modelagem dos dados para esse parâmetro foi difícil de ser realizada para a previsão de amostras com valores de pH mais baixos em relação ao conjunto, uma vez que poucas amostras apresentavam esses valores. Assim foi necessário para o desenvolvimento de modelos de calibração aceitáveis a exclusão de valores de pH abaixo de 4,10. Também quando na modelagem dos dados obtidos, principalmente na região do infravermelho próximo, várias amostras (espectros) tiveram que ser descartadas por apresentarem-se muito discrepantes em relação aos demais e por possuírem altos ruídos.

A Tabela 1 mostra as condições utilizadas na construção dos modelos e os melhores resultados obtidos; as Figuras 10a, 11a e 12a mostram os espectros originais MIR e NIR; e as Figuras 10b, 11b e 12b os espectros após o pré-processamento e com as faixas espectrais utilizadas para a construção dos modelos de calibração. Para espectros MIR – pasta, a região espectral encontrada que produziu os melhores espectros foi de 800 a 2100 cm^{-1} . Nessa região são encontradas bandas de absorção relativas a carboidratos (1400 - 800 cm^{-1}), proteínas (1700 - 1600 cm^{-1}) e água (1650 cm^{-1}).³⁶

Tabela 1. Resultados obtidos para o parâmetro pH

	Pré processamentos	LV ³	Calibração ⁴	Validação ⁵	Erro médio (%)	RMSEP	(R ²) ⁶
MIR	Suavização ¹						
Pasta	e MSC ²	3	34	17	1,17	0,06	0,70
MIR	Suavização ¹						
DESIR	e MSC ²	6	36	14	1,32	0,07	0,82
NIR							
DESIR	MSC ²	4	29	12	1,21	0,06	0,74

¹Suavização – filtro digital baseado no algoritmo Savitz-Golay

²MSC – correção de espalhamento multiplicativo

³LV – número de variáveis latentes

⁴Calibração – número de amostras utilizadas no conjunto de calibração

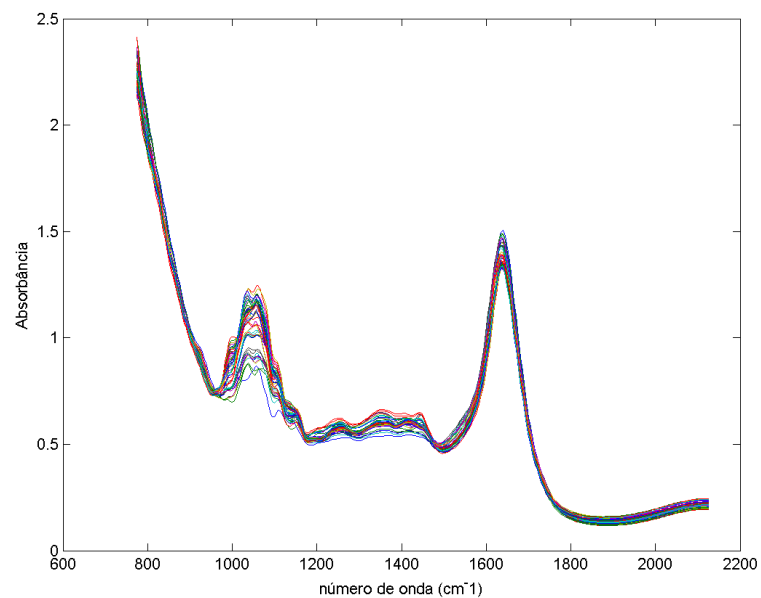
⁵Validação – número de amostras utilizadas no conjunto de validação

⁶R² – Coeficiente de correlação da reta obtida graficando-se o valor esperado contra o valor previsto pelo PLS

Para os espectros MIR e NIR obtidos a partir das amostras processadas pela técnica DESIR, não foi encontrada nenhuma faixa espectral que produzisse resultados superiores ao da utilização do espectro inteiro, somente para o NIR, foram eliminadas regiões iniciais (início em 820 nm) e finais do espectro (final em 2450 nm). Analisando-se o espectro inteiro na região do MIR, pode ainda destacar bandas de absorção relativas a carboidratos e gorduras na faixa de 2980 a 2800 cm⁻¹. Na Figura 11b, a região contínua de 2000 a 3700 cm⁻¹, devido a forte absorção da água, foi eliminada para a construção do modelo.

Em relação aos espectros NIR devido à alta sobreposição espectral é difícil uma atribuição das bandas espectrais encontradas, mas podem-se destacar bandas devido a proteínas (1020 nm, 1510 nm, 1980 nm, 2050 nm e 2180 nm), amido (990 nm, 1540 nm, 2000 nm, 2100 nm e 2460 nm), sacarose (1440 nm e 2080 nm) e glicose (1480 nm)¹⁴.

a)



b)

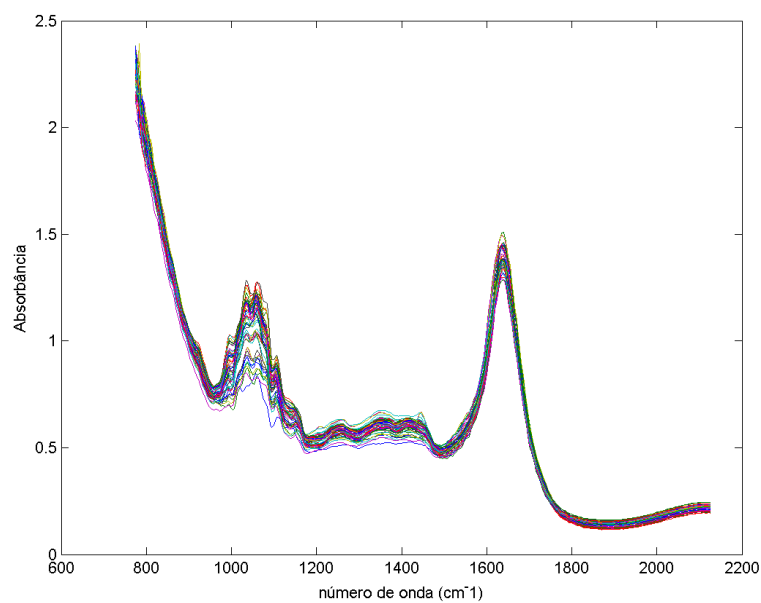
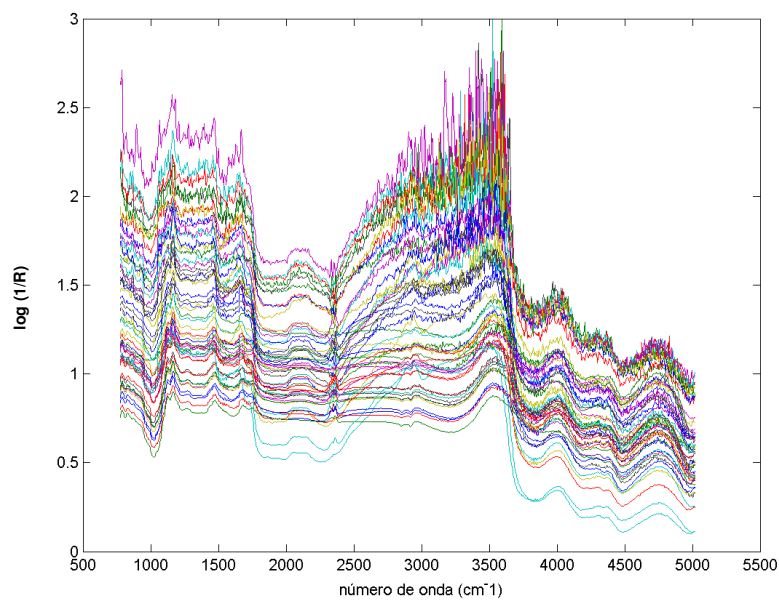


Figura 10. a) Espectros originais - MIR Pasta , b) Espectros utilizados para a construção dos modelos: faixa utilizada MIR Pasta– 800-2100 cm^{-1} (Suavização e MSC).

a)



b)

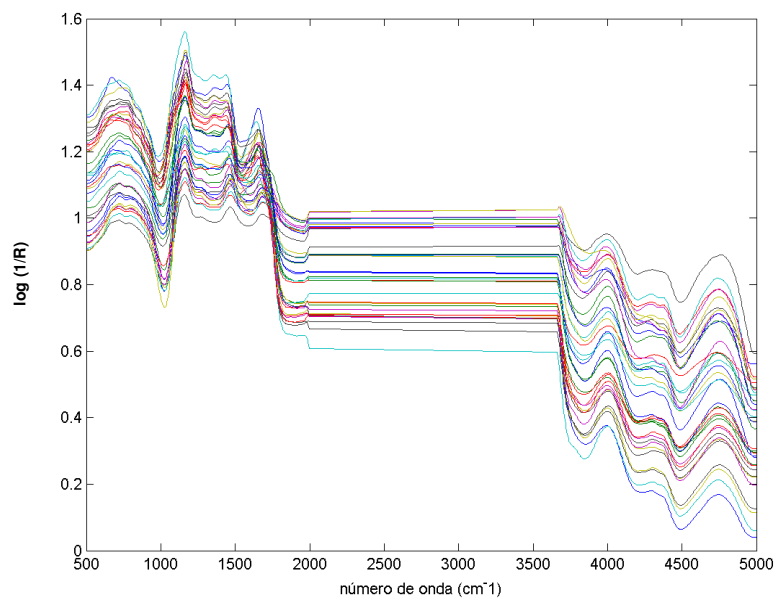
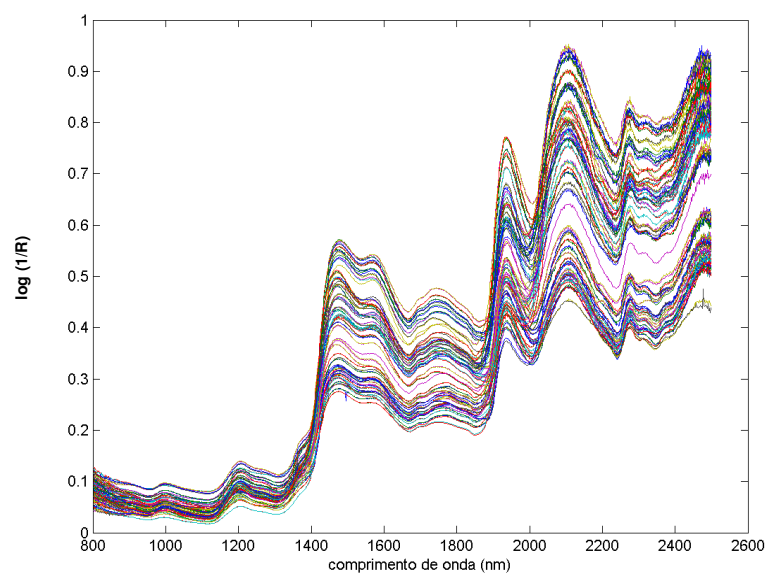


Figura 11. a) Espectros originais - MIR DESIR, b) Espectros utilizados para a construção dos modelos : faixa utilizada MIR DESIR – 500 – 5000 cm^{-1} (MSC).

a)



b)

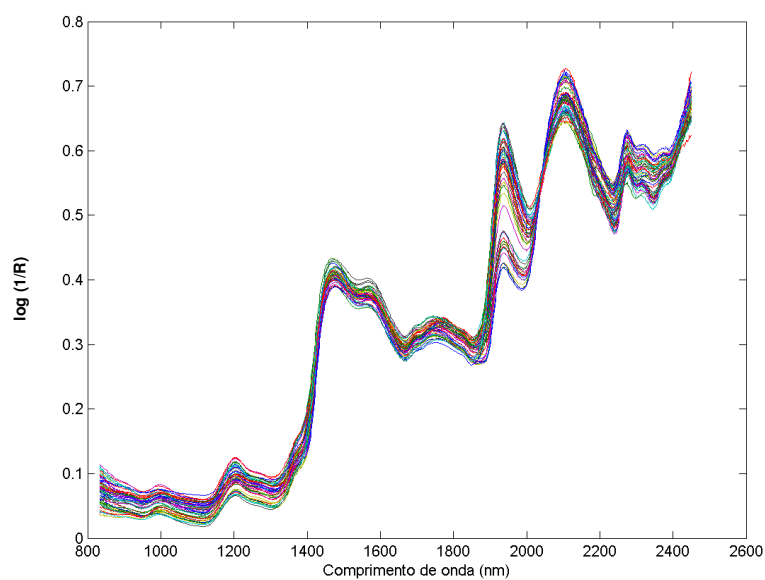
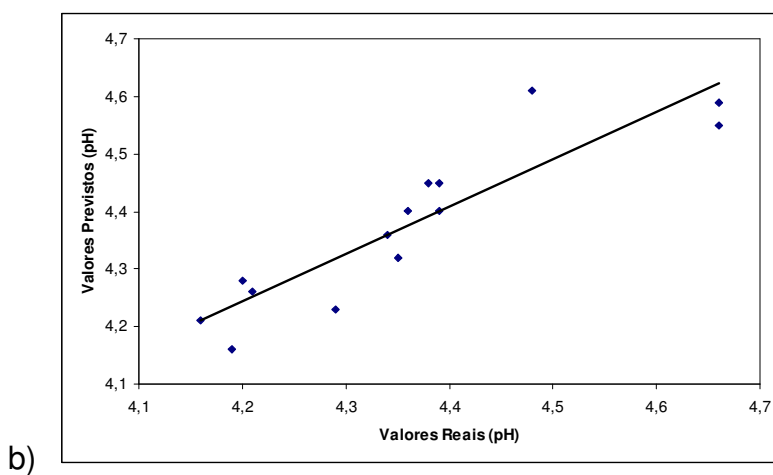
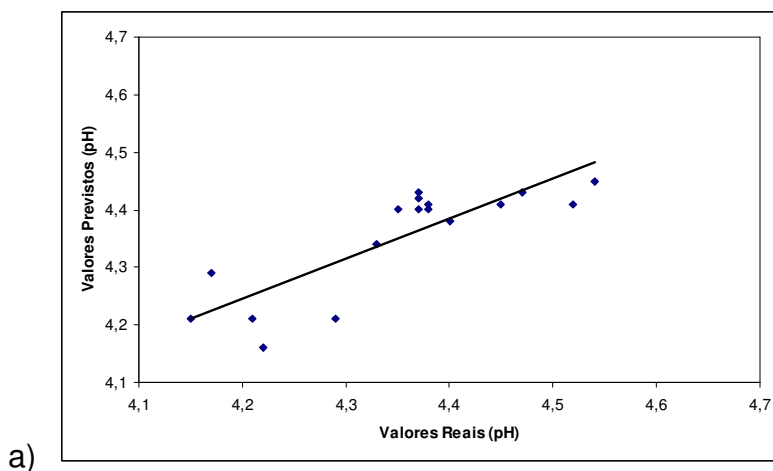


Figura 12. a) Espectros originais - NIR DESIR, b) Espectros utilizados para a construção dos modelos: faixa utilizada DESIR NIR – 820 – 2450 nm (MSC).

As tabelas 2a), 2b) e 2c) mostram os valores reais e previstos e as figuras 13a), 13b) e 13c) os gráficos relativos aos dados citados anteriormente.

Tabela 2. Valores reais e previstos: a) MIR pasta; b) MIR DESIR e c) NIR DESIR.

a)	real	previsto	b)	real	previsto	c)	real	previsto
	4,29	4,21		4,29	4,23		4,31	4,29
	4,22	4,16		4,66	4,59		4,36	4,26
	4,17	4,29		4,39	4,45		4,20	4,27
	4,15	4,21		4,21	4,26		4,38	4,40
	4,21	4,21		4,66	4,55		4,16	4,25
	4,52	4,41		4,35	4,32		4,26	4,31
	4,45	4,41		4,48	4,61		4,17	4,24
	4,47	4,43		4,16	4,21		4,19	4,25
	4,37	4,43		4,34	4,36		4,45	4,42
	4,35	4,40		4,19	4,16		4,40	4,39
	4,37	4,42		4,38	4,45		4,25	4,28
	4,37	4,40		4,36	4,40		4,45	4,52
	4,38	4,41		4,39	4,40			
	4,38	4,40		4,20	4,28			
	4,40	4,38						
	4,33	4,34						
	4,54	4,45						



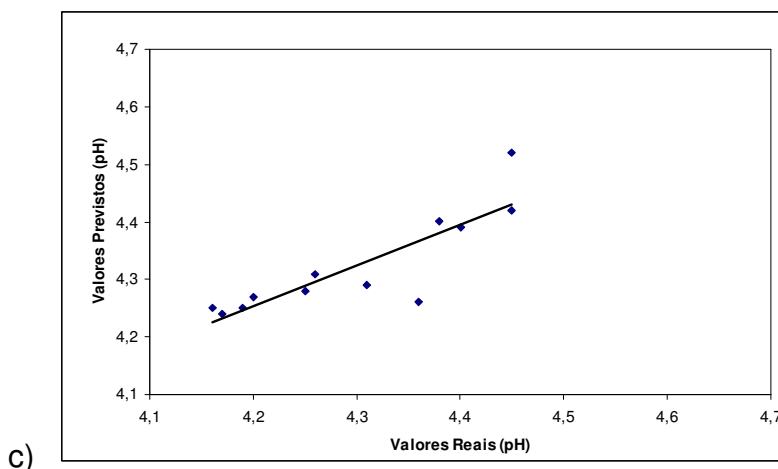


Figura 13. Gráficos dos valores de referência x valores previstos para os modelos: a) MIR pasta; b) MIR DESIR e c) NIR DESIR.

Os três modelos desenvolvidos apresentaram resultados muito próximos em relação ao erro médio (%) com valores em torno de 1% e ao RMSEP com valores de 0,06/0,07 unidades de pH, sendo que erros sempre menores que 2,9% foram obtidos em todos os casos. Porém a partir da análise dos coeficientes de correlação (R^2) das retas obtidas entre o valor de referência e o valor previsto, pode-se notar que para o modelo no MIR, com medidas pela técnica DESIR (Figura 13b), existe um melhor ajuste da reta, ou seja, previsões com erros mais parecidos foram obtidas. Analisando-se os espectros talvez seja possível encontrar uma explicação para esse fato, pois no caso do MIR – DESIR nota-se claramente uma maior diferenciação entre as amostras, além de ter sido eliminada a banda de água. Também nesse caso, devido ao procedimento empregado no DESIR, existe uma concentração do material sólido para a realização das medidas.

4.2. Acidez

O método padrão é baseado na titulação com solução padrão de NaOH, monitorada com eletrodo indicador de pH de vidro. Muitas vezes a polpa da banana fica sobre a membrana de vidro, apesar da agitação constante, ocasionando grandes erros na quantificação da acidez, prejudicando a modelagem dos dados. A faixa de acidez variou de 0,300 a 0,613 g de ácido málico para 100 g de amostra. A Tabela 2 mostra as condições e resultados dos melhores modelos desenvolvidos e as Figuras 14, 15 e 16 os espectros com as faixas espectrais utilizadas para a construção dos modelos de calibração. Nesse caso, devido a maior dificuldade no desenvolvimento dos modelos de calibração, foi necessário um estudo prévio para seleção da melhor faixa espectral, visto que a utilização de todo o espectro não produzia bons resultados. No caso dos espectros na região do infravermelho próximo, foi possível com a utilização do iPLS a seleção de um excelente intervalo espectral que produziu resultados muito favoráveis. Para os espectros da região do infravermelho médio, não foi possível obter resultados favoráveis com a utilização do iPLS, mas houve a seleção manual da melhor faixa espectral por tentativa e erro.

Para dados de MIR pasta, a melhor região espectral escolhida foi de 770 – 1780 cm^{-1} que pode ser atribuída a bandas devido a carboidratos (800–1400 cm^{-1})³⁶.

Tabela 3. Resultados obtidos para o parâmetro acidez.

Pré Processamentos		LV¹	Calibração¹	Validação¹	Erro médio (%)	RMSEP	(R²)¹
MIR Pasta	Suavização ¹ e derivada	6	34	17	10,50	0,05	0,05
MIR DESIR	Suavização e Derivada	4	36	14	9,38	0,05	0,58
NIR DESIR	Derivada e iPLS	5	24	13	6,20	0,03	0,93

¹Vide Tabela 1

Na Figura 16, pode-se observar os resultados do iPLS, onde nesse caso o espectro foi dividido em 5 partes iguais. Foi estudada a divisão de espectro em diferentes partes, mas o melhor resultado é o que está aqui apresentado. A linha pontilhada nessa Figura indica o RMSECV obtido quando são utilizadas todas as variáveis no modelo, enquanto cada barra vertical indica o RMSECV do intervalo. Na barra vertical preenchida (escura) está apresentada a região do espectro escolhida para a modelagem que tem o menor valor de RMSECV, correspondente ao segundo intervalo. Este intervalo corresponde aos comprimentos de onda de 1141 a 1480 nm relacionados principalmente a açúcares e amido (estiramento O-H, 1º. sobretom)¹⁴.

Os conjuntos de espectros originais foram os mesmos obtidos para o parâmetro pH, podendo ser observados nas Figuras 10, 11 e 12.

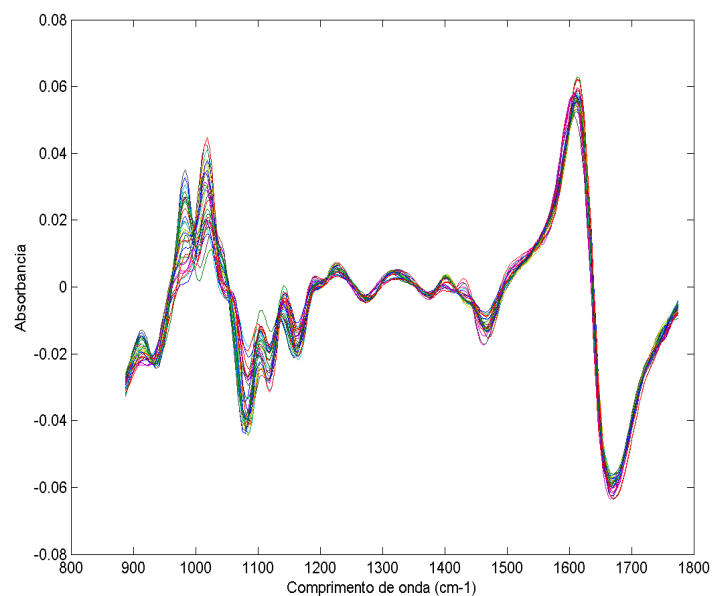


Figura 14. Espectros utilizados para a construção dos modelos MIR pasta. Faixa utilizada 770 – 1780 cm^{-1} (suavização e 1ª.derivada)

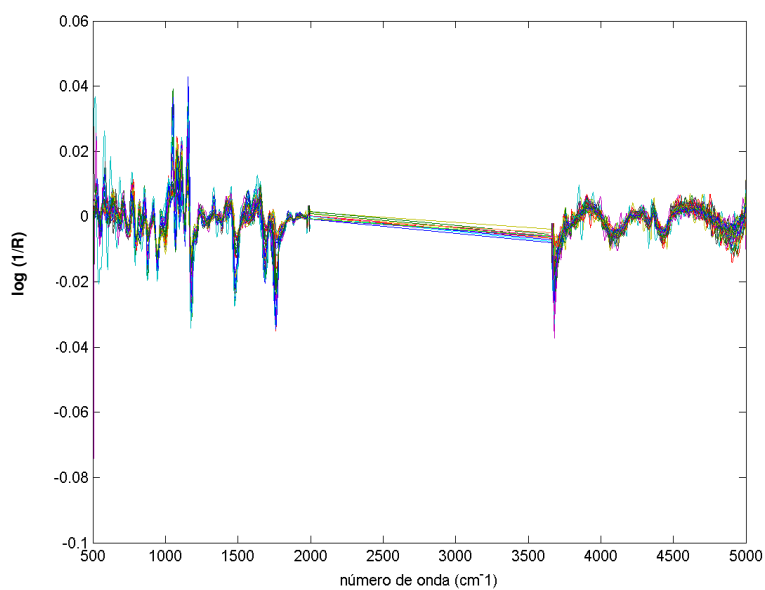


Figura 15. Espectros utilizados para a construção dos modelos de MIR DESIR. Faixa espectral utilizada: 500 – 5000 cm^{-1} (1ª.derivada)

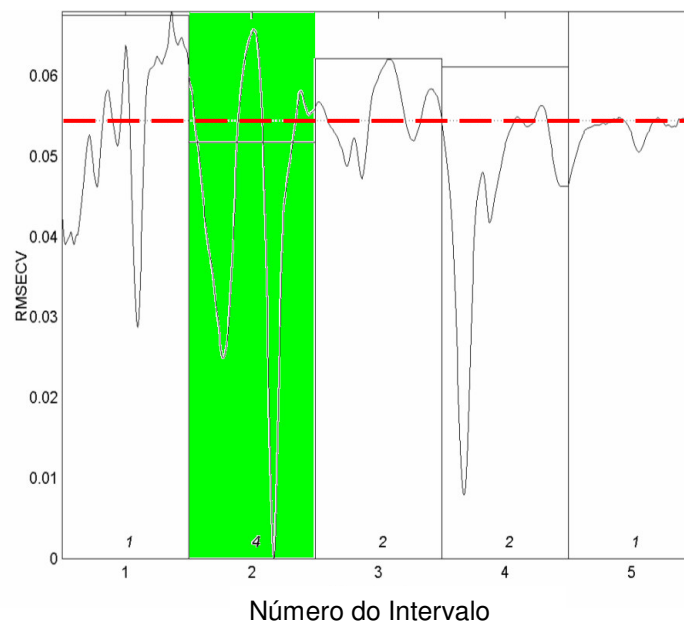


Figura 16. Resultados do iPLS com 5 intervalos para NIR DESIR (1ª. derivada).

As tabelas 4a), 4b) e 4c) mostram os valores reais e previstos e as figuras 17a), 17b) e 17c) as retas obtidas quando são graficados os valores de referência contra os valores previstos pelo modelo relativos à previsão da acidez.

Tabela 4. Valores reais e previstos: a) MIR pasta; b) MIR DESIR e c) NIR DESIR.

a)	real	previsto	b)	real	previsto	c)	real	previsto
	0,45	0,45		0,54	0,49		0,54	0,57
	0,52	0,52		0,52	0,53		0,61	0,63
	0,37	0,43		0,61	0,72		0,51	0,50
	0,38	0,46		0,50	0,42		0,31	0,31
	0,38	0,45		0,53	0,46		0,41	0,38
	0,39	0,44		0,50	0,45		0,47	0,46
	0,51	0,44		0,45	0,48		0,38	0,31
	0,50	0,43		0,42	0,45		0,37	0,34
	0,45	0,44		0,49	0,41		0,41	0,37
	0,41	0,42		0,45	0,41		0,45	0,44
	0,49	0,43		0,49	0,46		0,49	0,54
	0,49	0,42		0,41	0,42		0,45	0,42
	0,41	0,42		0,35	0,38		0,40	0,38
	0,36	0,42		0,44	0,41			
	0,45	0,43						
	0,43	0,38						
	0,44	0,37						

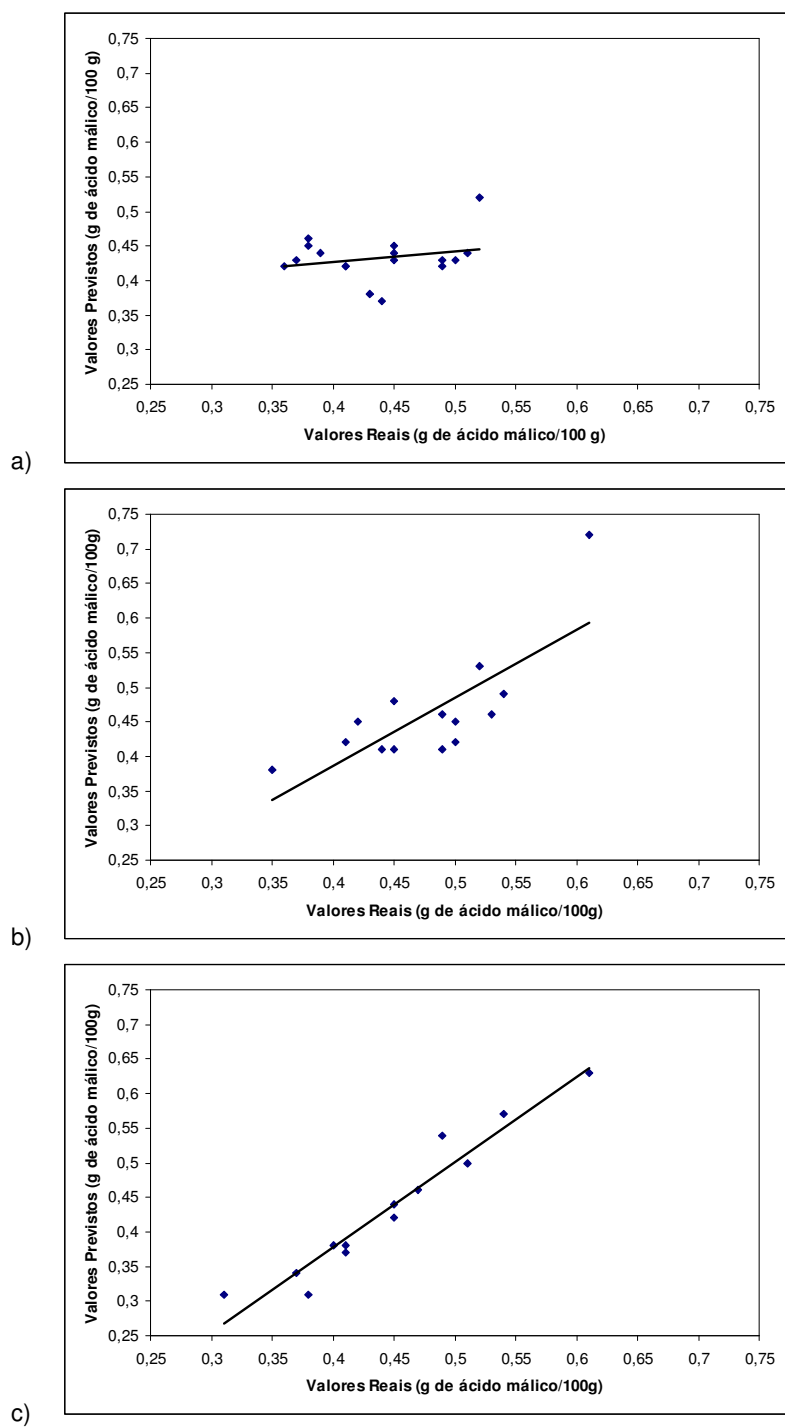


Figura 17. Gráficos dos valores reais x valores previstos para os modelos: a) MIR pasta; b) MIR DESIR e c) NIR DESIR.

Os modelos obtidos baseados na região do infravermelho próximo apresentaram os melhores resultados com menores erros médios de previsão de 6% e melhor ajuste entre os valores de referência e previstos pelo modelo. Para os modelos construídos com espectros na região do infravermelho médio, erros de até 20% foram obtidos, além de erros médios da ordem de 10%. Também se verifica claramente na Figura 17 que o melhor ajuste é obtido pelo NIR. Pode-se afirmar nesse caso que modelos aceitáveis foram obtidos somente com espectroscopia NIR, fato até certo ponto não esperado. Assim como para o parâmetro pH, a técnica DESIR produziu melhores resultados, provavelmente pela maior concentração do material para a medida e ausência de água. Porém nesse caso, medidas na região do NIR com a região selecionada pelo iPLS produziram melhores resultados pela seleção adequada da região espectral.

4.3. Brix

A faixa de Brix variou de 6,5 a 24,5° para as bananas analisadas, o que deve garantir uma variabilidade adequada para o desenvolvimento dos modelos de calibração. A Tabela 3 apresenta as condições utilizadas na construção dos modelos e os melhores resultados obtidos; as Figuras 18a, 19a e 20a os espectros originais e as Figuras 18b, 19b e 20b as regiões espectrais utilizadas para MIR da solução e NIR da solução e DESIR.

As tabelas 6a), 6b) e 6c) mostram os valores reais e previstos e as figuras 21a), 21b) e 21c) os gráficos relativos aos modelos.

Tabela 5. Resultados para o parâmetro Brix.

	Pré Processamentos	LV ¹	Calibração ¹	Validação ¹	Erro médio (%)	RMSEP	(R ²) ¹
MIR Solução	Derivada e iPLS	6	28	14	3,78	0,60	0,98
NIR Solução	Suavização ¹	5	23	11	5,32	1,11	0,83
NIR DESIR	MSC ¹	5	26	11	13,16	2,18	0,50

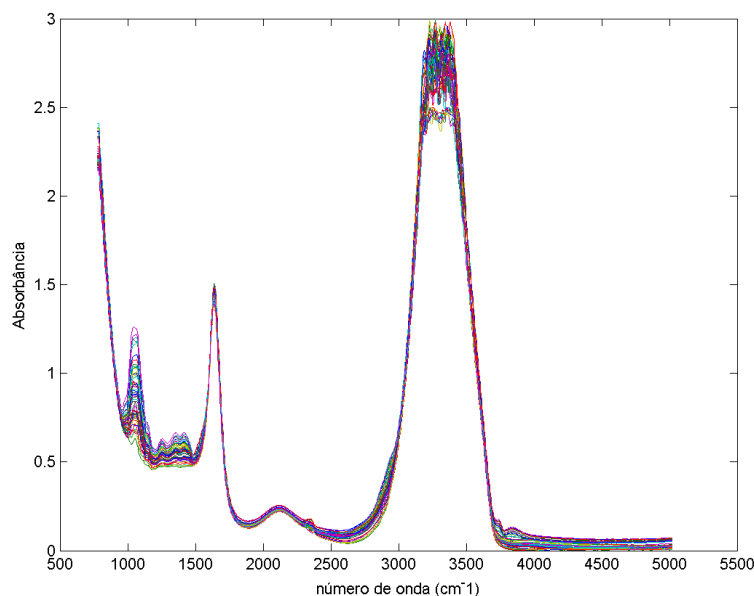
¹Vide Tabela 1.

Nesse caso, para os espectros MIR em solução foi possível com o iPLS encontrar uma região espectral mais adequada para o desenvolvimento dos modelos de calibração. Para os espectros na região do infravermelho próximo não foi possível obter com o iPLS modelos superiores aos apresentados. Para o NIR, a região escolhida para o desenvolvimento dos modelos foi realizada por tentativa e erro.

No desenvolvimento do iPLS para espectros MIR da solução, novamente os melhores resultados foram obtidos quando o espectro foi dividido em 5 partes, sendo que o segundo intervalo apresentou valor de RMSECV menor que os outros e menor que na utilização de todo o espectro. Na Figura 18b está apresentado o

intervalo escolhido que corresponde à região de 1481 a 2360 cm^{-1} relacionados a bandas de carboidratos (800 – 1400 cm^{-1}), proteínas (1600 – 1700 cm^{-1}) e água (1650 cm^{-1})³⁶.

a)



b)

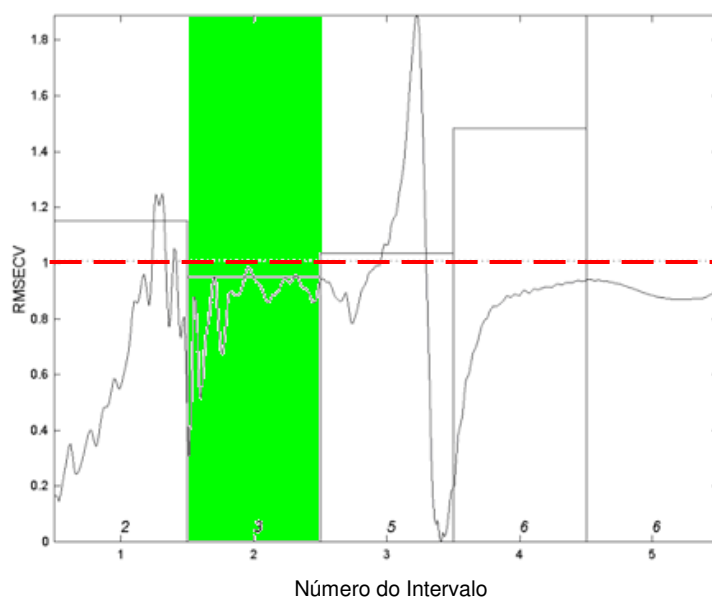
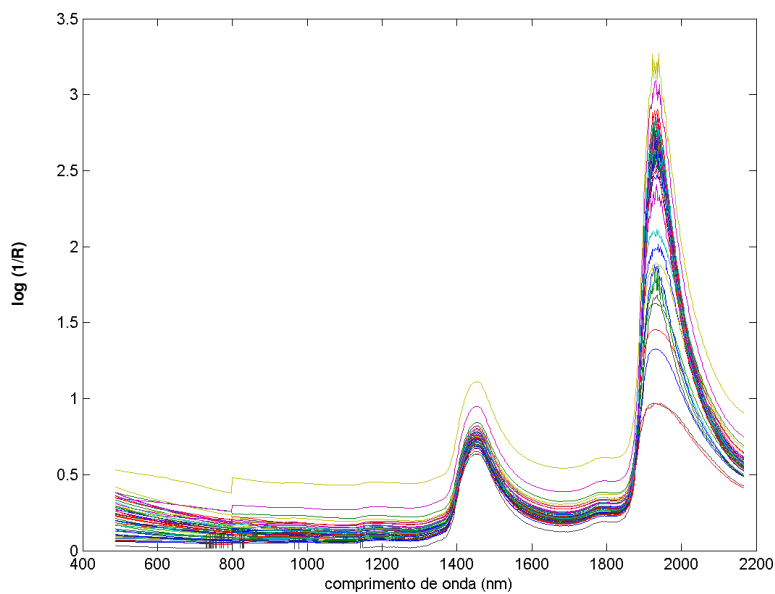


Figura 18. a) Espectros MIR Solução e b) Espectros utilizados para a construção dos modelos: faixa utilizada MIR Solução (Derivada e iPLS);

a)



b)

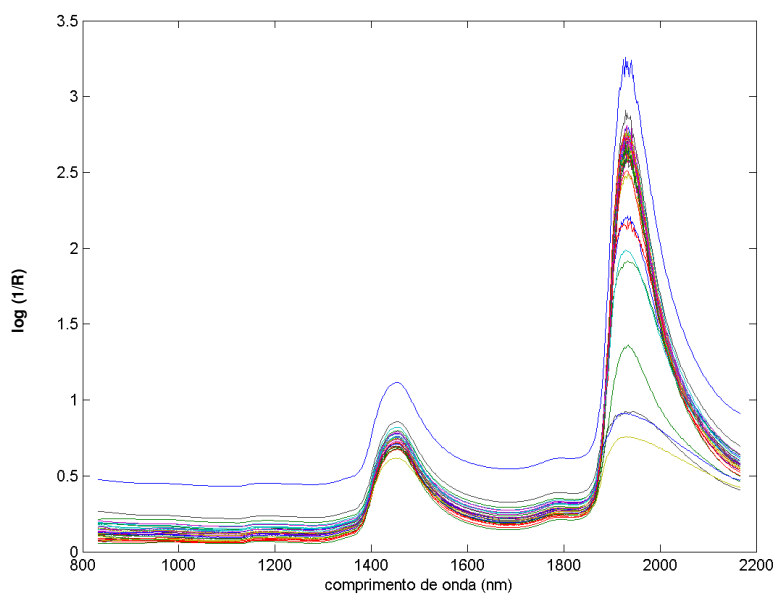
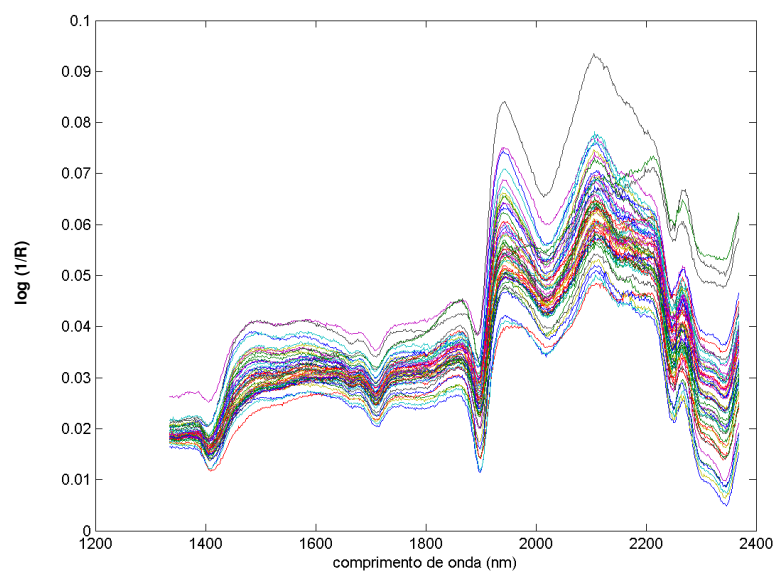


Figura 19. a) Espectros NIR Solução e b) Espectros utilizados para a construção dos modelos e b) faixa utilizada NIR Solução – 820 – 2180 nm (Suavização)

a)



b)

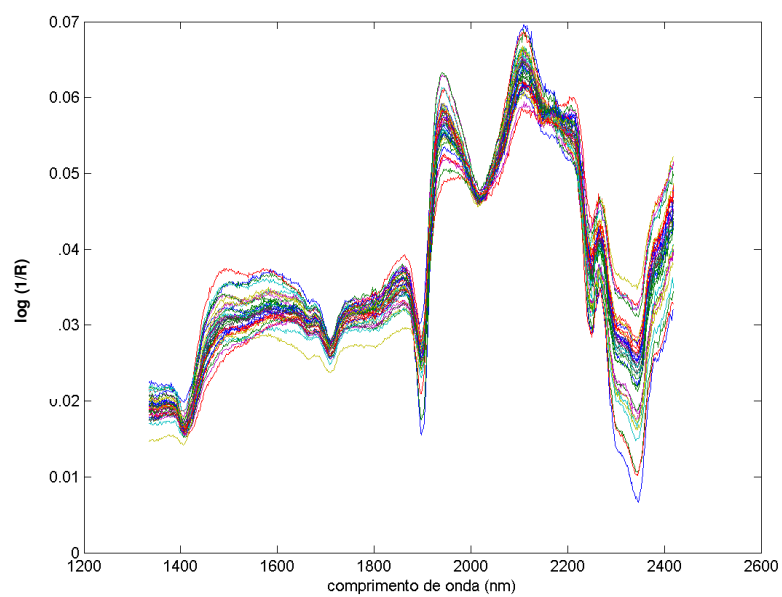


Figura 20. a) Espectros NIR DESIR. b) Espectros utilizados para a construção dos modelos: faixa utilizada DESIR NIR – 1330 - 2410 (MSC).

Tabela 6. Valores reais e previstos: a) MIR Solução; b) NIR Solução e c) NIR DESIR.

a)	real	previsto	b)	real	previsto	c)	real	previsto
	14,5	13,8		15,5	15,8		15,3	13,2
	11,5	12,8		15,0	15,9		9,5	12,1
	13,5	13,6		12,0	12,9		16,5	16,6
	15,0	15,5		17,5	15,4		17,0	17,0
	16,5	17,2		20,0	20,2		12,5	16,1
	9,5	10,4		13,0	13,2		17,5	16,6
	17,5	17,8		16,5	14,7		18,0	16,7
	16,1	15,2		17,0	17,3		9,0	10,1
	20,0	20,2		11,0	10,8		13,0	14,6
	23,2	22,8		15,4	13,4		17,5	13,1
	8,0	8,1		18,0	17,8		13,1	11,5
	11,2	10,9						
	10,9	10,6						
	13,1	12,8						

Para os espectros NIR em solução a região espectral escolhida foi de 820 a 2180 nm que corresponde a bandas de proteína (1020 nm, 1510 nm, 1980 nm, 2050 nm e 2180 nm), amido (990 nm, 1540 nm, 2000 nm, 2100 nm), sacarose (1440 nm e 2080 nm) e glicose (1480 nm)¹⁴. Nesse espectro são muito destacadas as bandas relativas à água em 1450 nm e 1940 nm¹⁴.

Por outro lado, para espectros NIR DESIR a região selecionada foi de 1330 a 2410 nm. Nesse caso, pode-se destacar bandas relativas a proteína (1510 nm, 1980 nm e 2180 nm), amido (1540 nm, 1900 nm, 2000 nm, 2100 nm e 2252 nm), sacarose (1440 nm e 2080 nm) e glicose (1480 nm)¹⁴.

As Figuras 21a, 21b e 21c apresentam os gráficos das retas obtidas quando são graficados os valores de referência contra os valores previstos pelo modelo relativos à previsão do parâmetro Brix.

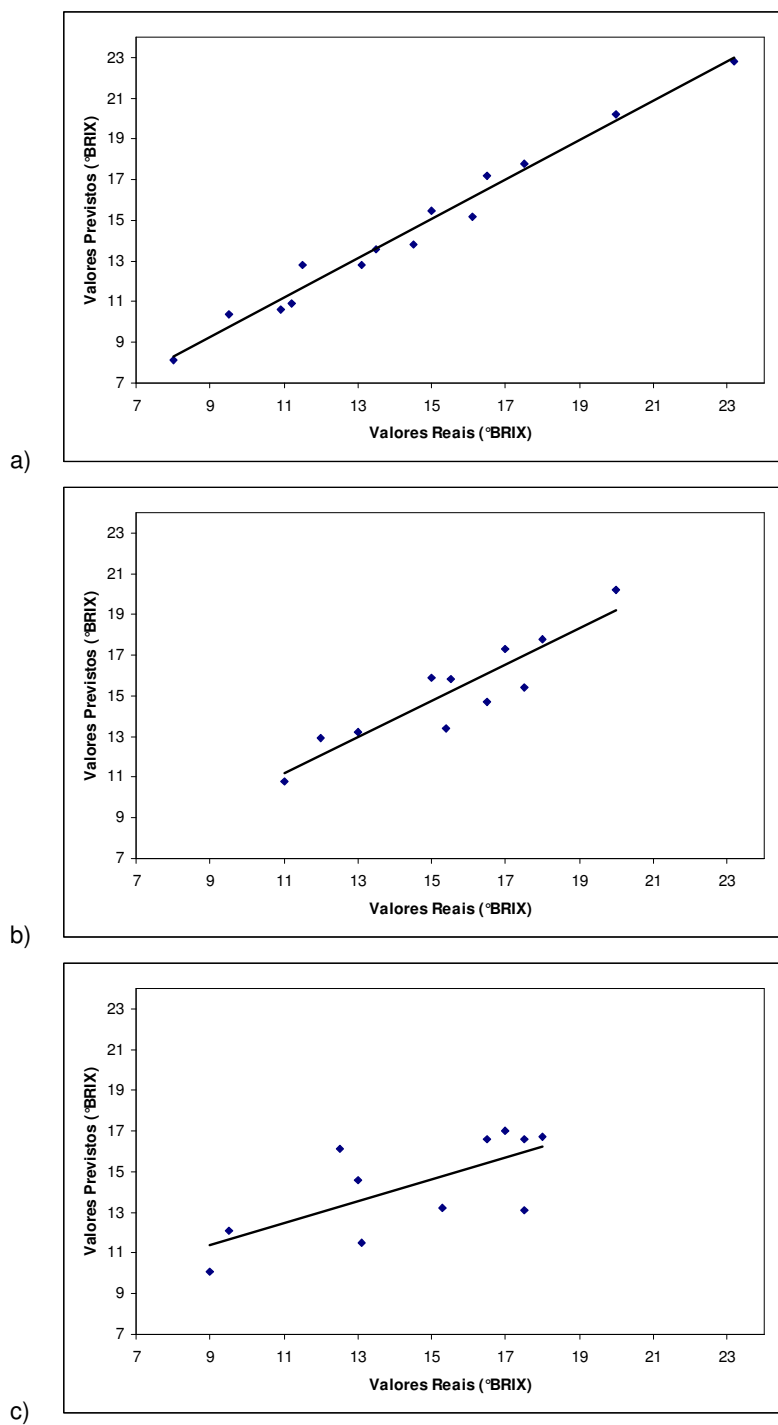


Figura 21. Gráficos dos valores de referência x valores previstos para os modelos:
a) MIR Solução; b) NIR Solução e c) NIR DESIR.

Os melhores resultados foram obtidos utilizando espectroscopia no infravermelho médio da solução. A utilização da solução obtida a partir da centrifugação da polpa macerada de banana gerou dificuldades na obtenção dos espectros no NIR, pois devido à viscosidade da mesma não foi possível filtrá-la, restando alguns pequenos resíduos sólidos, impedindo que a solução se apresentasse límpida, necessária para a análise. Já no MIR, a análise foi realizada pela disposição da solução em um cristal de ATR, minimizando o mesmo efeito e apresentando os melhores valores resultados de erro médio, RMSEP e R^2 . Esse fato justifica a menor correlação dos dados do NIR quando comparada ao modelo obtido no MIR, como pode ser observado na tabela 3, independente da técnica utilizada.

A técnica DESIR não obteve bons resultados para esse parâmetro, uma vez que esta análise relaciona-se com a quantidade de sólidos em solução. Como com a técnica DESIR ocorre a eliminação de água, fica difícil a sua determinação, ocasionando baixa correlação entre os valores reais e previstos e altos valores de erro médio(%) e RMSEP.

CAPÍTULO 5 – Conclusões

5. Conclusões

A utilização da espectroscopia no infravermelho, tanto no MIR como no NIR é aplicável para a determinação dos parâmetros que determinam a qualidade do fruto proposto. Os métodos de exposição da amostra, solução, pasta e DESIR atenderam as expectativas e responderam positivamente na construção dos modelos, sendo que para medidas em solução o NIR não apresentou resultados satisfatórios.

Para a determinação do pH os modelos foram eficientes em ambas regiões do infravermelho e independente da técnica utilizada. Pode-se observar a limitação na modelagem dos dados com magnitudes de pH muito pequenas em relação ao conjunto.

Na determinação da acidez, o método padrão não foi eficiente para a obtenção dos valores de referência, prejudicando assim a modelagem, sendo necessário utilizar um grande número de amostras para minimizar o erro gerado. Nesse caso, foi verificado que modelos aceitáveis foram obtidos somente com espectroscopia NIR com seleção de variáveis pelo iPLS, fato até certo ponto não esperado.

Para a determinação de sólidos solúveis, a análise na região do infravermelho médio, utilizando a reflectancia total atenuada da solução foi mais eficiente do que os métodos utilizados na região do infravermelho próximo, devido a maneira de exposição da amostra e a relação com a propriedade de interesse.

CAPÍTULO 6 – Referências Bibliográficas

6. Referências Bibliográficas

1. VILAS BOAS, E.V. *Modificações Pós-Colheita de banana "Prata" (Musa acuminata x Musa balbisiana Grupo AAB) γ -irradiada*. Universidade Federal de Lavras, Dissertação de Mestrado, Lavras, 1995.
2. SOUZA, A.T. *Banana - Expectativa é de recuperação do setor em 2008*. Centro de Socioeconomia e Planejamento Agrícola – Governo do Estado de Santa Catarina, <http://cepa.epagri.sc.gov.br>, acessado em março de 2007.
3. SOTO-BALLESTERO, M. *Bananos: cultivo y comercialización*, 2° Ed., Litografia e Imprenta LIL, San José, 1992.
4. NOGUEIRA, L. ed. *Cresce Minas: Um projeto Brasileiro. Norte de Minas é pólo de pesquisa*. Minas faz ciência. **5**, 12-15, 1999.
5. AGRIANUAL, *Anuário da agricultura brasileira*. FNP – Consultoria e Comércio, São Paulo, 2007.
6. MOTA, R. V. *Cienc. Tecnol. Aliment.* **25**, 560 – 563, 2005
7. CHITARRA, M.I.F.; CHITARRA, A.B. *Pós colheita de frutas e hortaliças. Fisiologia e manuseio*, ESAL/FAEPE, Lavras, 1990.
8. FORSUTH, W.G.C. *Banana and Plantain. In: Tropical and subtropical fruits*, AVI Publishing, Westport, 1980.
9. MOTA, R. V.; LAJOLO, F. M. ; CORDENUNSI, B. R. *Cienc. Tecnol. Aliment.* **17**, 94-97, 1997.

10. BOLLARD, E.C. *The Biochemistry of Fruits and their Product*, Academic Press, New York, 1971.
11. ASKAR, A.; TREPTOW, H. *Quality assurance in tropical fruit processing*, Springer , Weinheim, 1993.
12. SKOOG, D.A.; LEARY, J.J. *Principles of instrumental Analysis*. 4^o Ed., Saunders College Publishing, New York, 1992.
13. TARKOSOVA, J.; COPIKOVA, J. *J. Near Infrar. Spectrosc.* **8**, 21-26, 2000.
14. OSBORNE, B.G.; FEARN, T.; HINDLE, P.H. *Pratical NIR spectroscopy with applications in food and beverage analysis*, 2Ed., Longman Scientific & Technical, London, 1993.
15. DRENNEN, J.K.; KRAEMER, E.G.; and LODDER, R.A. *Critical Rev. Anal. Chem.* **22**, 443-475, 1991
16. GIANGIACOMO, R.; MAGEE, J.B.; BIRTH, G.S.; DULL, G.G. *J. Food Sci.* **46**, 531-534, 1981.
17. LANZA, E.; LI, B.W. *J. Food Sci.* **49**, 995-998, 1984.
18. SARANWONG, S., SORNSRIVICHAI, J.; KAWANO, S. *J. Near Infrar. Spectrosc.* **9**, 287-295, 2001.
19. GUTHRIE, J., WALSH, B.W.K. *J. Near Infrar. Spectrosc.* **6**, 259-265, 1998.
20. SEGTMAN, V. H.; ISAKSSON, T. J. *J. Near Infrar. Spectrosc.* **8**, 109-116, 2000.

21. LISHENG U.; SCHLUP, J.R. *Appl. Spectrosc.* **50**, 109-114, 1996.
22. HARRICK, N.J. *Anal. Chem.* **36**, 188-191, 1964.
23. BRUNS, R.E.; FAIGLE, J.F.G. *Química Nova* **18**, 84 –99, 1995.
24. THOMAS, E.V. *Anal. Chem.*, **66**, 795A-814A, 1994.
25. GELADI, P.; KOWALSKI, B.R. *Anal. Chim. Acta* **185**, 1-17, 1986.
26. POPPI, R. J. *Quantificação de picos cromatográficos superpostos por métodos de calibração multivariada* – Dissertação de Mestrado, UNICAMP, Campinas, 1989.
27. MARTENS, H.;TORMOD, N. *Multivariate Calibration*, John Wiley & Sons, New York, 1989.
28. OTTO, M. *Chemometrics*, Wiley-VCH, Weinheim, 1999.
29. MALEKI, M. R.; MOUAZEN, A. M.; RAMON, H. *Biosyst. Eng.* **96**, 427-433, 2007
30. THOMAS, E.V. *Anal. Chem.* **66**, A795-A804, 1994.
31. THOMAS, E.V.; HAALAND, D.M. *Anal. Chem.* **62**, 1091 – 1099, 1990.
32. NORGAARD, L.; SAUDLAND, A.; WAGNER, J.; NIELSEN, J. P.; MUNCK, L; ENGELSEN, B. *Appl. Spectrosc.* **54**, 413-419, 2000.
33. CECCHI, H. M. *Fundamentos teóricos e práticos em análise de alimentos*, 2ª. Ed., Editora da Unicamp, Campinas, 2003.

34. Official Methods of Analysis of AOAC International 942.15 ,*Acidity (Titrable) of fruit Products*, 16th edition, AOAC International, Gaithersburg, 1997.
35. KENNARD, R. W.; STONE, L. A. *Technometr.* **11**, 137-148, 1969.
36. CHALMERS, J. M.; GRIFFITHS, P. R. *Handbook of Vibrational Spectroscopy*, vol.5, Pharmaceutical and Natural Sciences, Wiley, 2002.