UNIVERSIDADE TECNOLÓGICA
FEDERAL DO PARANÁ
Programa de Pós-Graduação em Tecnologia de Alimentos

DAYANE APARECIDA DOS SANTOS

DESENVOLVIMENTO DE MODELO DE CALIBRAÇÃO MULTIVARIADA MULTIPRODUTO PARA QUANTIFICAÇÃO DE ACIDEZ E VITAMINA C EM NÉCTAR DE FRUTAS E BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA

DISSERTAÇÃO

CAMPO MOURÃO
2015

DAYANE APARECIDA DOS SANTOS

DESENVOLVIMENTO DE MODELO DE CALIBRAÇÃO MULTIVARIADA MULTIPRODUTO PARA QUANTIFICAÇÃO DE ACIDEZ E VITAMINA C EM NÉCTAR DE FRUTAS E BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA

Dissertação apresentada ao programa de Pós Graduação em Tecnologia de Alimentos da Universidade Tecnológica Federal do Paraná, como parte dos requisitos para obtenção do título de mestre em Tecnologia de Alimentos.

CAMPO MOURÃO

2015

Dados Internacionais de Catalogação na Publicação

S237d Santos, Dayane Aparecida dos.

Desenvolvimento de modelo de calibração multivariada multiproduto para quantificação de acidez e vitamina C em néctar de frutas e bebidas industrializadas a base de soja. / Dayane Aparecida dos Santos. – 2015.

64 f. : il. ; 30 cm.

Orientadora: Patrícia Valderrama.

Co-orientador: Paulo Henrique Março.

Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologia de Alimentos. Medianeira, 2015.

Inclui bibliografias.

1. Vitamina C. 2. Alimentos – Análise. 3. Alimentos – Dissertações. I. Valderrama, Patrícia, oriente. II. Março, Paulo Henrique, co-orient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologia de Alimentos. IV. Título.

CDD: 664.02 - Ed. 22

Biblioteca Câmpus UTFPR Medianeira

Marci Lucia Nicodem Fischborn 9/1219

Ministério da Educação
Universidade Tecnológica Federal do Paraná
Programa de Pós-Graduação em Tecnologia de Alimentos

TERMO DE APROVAÇÃO

DESENVOLVIMENTO DE MODELO DE CALIBRAÇÃO MULTIVARIADA MULTIPRODUTO PARA QUANTIFICAÇÃO DE ACIDEZ E VITAMINA C EM NÉCTAR DE FRUTAS E BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA

POR

DAYANE APARECIDA DOS SANTOS

Essa dissertação foi apresentada às nove horas, do dia vinte e sete de março de dois mil e quinze, como requisito parcial para obtenção do titulo de Mestre em Tecnologia de Alimentos, Linha de Pesquisa: Ciência e Tecnologia de Produtos Alimentícios, no Programa de Pós-graduação em Tecnologia de Alimentos, da Universidade Tecnológica Federal do Paraná. A candidata foi arguida pela Banca Examinadora composta pelos professores abaixo assinados. Após a deliberação, a Banca Examinadora considerou o trabalho APROVADO.

_____

Profª. Drª. Patricia Valderrama (Orientadora – PPGTA)

_____

Prof. Dr. Evandro Bona (Membro Interno – UTFPR)

_____

Prof. Dr. Ronei J. Poppi (Membro Externo – UNICAMP)

Orientador

Professora Dr. Patrícia Valderrama

Coorientador

Professor Dr. Paulo Henrique Março

“O sucesso é uma consequência e não um objetivo.”

Gustave Flaubert.

AGRADECIMENTOS

Agradeço primeiramente a Deus, que nos momentos difíceis me encorajou a seguir em frente.

À minha família, minha mãe Rosa Maria de Oliveira Santos, meu pai Revail Pereira dos Santos, ao meu irmão Alexandre Aparecido dos Santos e minha irmã Madalena Aparecida dos santos, pelo amor, confiança, conselhos e apoio aos quais me fizeram chegar até aqui.

À minha orientadora Profa. Dra. Patricia Valderrama, que não mediu esforços para me orientar e sempre esteve presente e sempre estar presente nos momentos de duvida nessa caminhada, por confiar no meu trabalho, pela oportunidade de aprendizado e por me acalmar nos meus momentos de desespero.

Ao Prof. Dr. Paulo Henrique Março pela coorientação, por sua atenção, tranquilidade e compreensão ao ensinar, por compartilhar seu conhecimento com o intuito de enriquecer o trabalho.

Aos professores do Programa de Pós-graduação em Tecnologia de Alimentos pelos conhecimentos transmitidos durante o curso.

A minha amiga Karen Priscila de Lima, pela ajuda em todos os trabalhos.

A todos meus amigos, pelo apoio em todos os momentos.

Aos amigos que fiz durante o curso e que sempre terei no coração Jorge L. Sanchez, Franciele Viel, Ana Paula de Oliveira Pizzoli, Vanessa Rodrigues, Vanessa Fernandes Fonseca, Mariana Oliveira, Maristela Angel.

À Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) pelo auxílio financeiro.

À Metrohm Pensalab pelo empréstimo do equipamento Vis-NIR.

Quero agradecer a todos que auxiliaram de alguma forma no desenvolvimento deste trabalho, muito obrigada!

Ao senhor Deus, porque nele encontro força e refúgio, aos meus pais Revail e Rosa, que são as pessoas mais importantes da minha vida. A minha irmã Madalena e ao meu irmão Alexandre, por sempre estarem ao meu lado me apoiando. Ao meu namorado Jonas pelo apoio, compreensão e carinho.

RESUMO GERAL

SANTOS, Dayane Aparecida. Desenvolvimento de modelo de calibração multivariada multiproduto para quantificação de acidez e vitamina c em néctar de frutas e bebidas industrializadas a base de soja. 2015. 61 f. Dissertação – Programa de Pós-Graduação em Tecnologia de Alimentos, Universidade Tecnológica Federal do Paraná. Campo Mourão, 2015.

A determinação da acidez total e vitamina C empregando espectroscopia ultravioleta, visível, infravermelho próximo e calibração multivariada multiproduto é apresentada como uma alternativa ao método convencional da titulação potenciométrica e da titulação por oxidação-redução, respectivamente. No modelo desenvolvido foram incluídos diferentes produtos (néctar de frutas de diferentes sabores, néctar de frutas light, bebidas industrializadas a base de soja e a base de soja light). Os conjuntos de calibração e validação foram otimizados através da identificação de amostras anômalas. O método linear de mínimos quadrados parciais (PLS) foi empregado na construção do modelo. A exatidão, representada pelos valores dos erros médio quadráticos de calibração (RMSEC) e previsão (RMSEP), foi determinada. A elipse de confiança e o ajuste confirmam a exatidão alcançada enquanto que, os resíduos apresentaram comportamento aleatório, indicando linearidade para os modelos desenvolvidos. Outras figuras de mérito, tais como, sensibilidade, sensibilidade analítica, relação de desempenho do desvio, limites de detecção e quantificação foram calculadas. A metodologia proposta apresenta como vantagens rapidez, não emprega reagentes/solventes e não gera resíduos tóxicos, sendo uma alternativa aos métodos convencionais baseados em potenciometria e titulação de oxidação-redução.

PALAVRAS-CHAVE: Vitamina C, Acidez Total, Calibração multivariada, Multiproduto

GENERAL ABSTRACT

SANTOS, Dayane Aparecida. Multiproduct multivariate calibration model development to quantify the acidity and vitamin C in industrialized fruit nectar and soy juices. 2015. 61 f. Dissertação – Programa de Pós-Graduação em Tecnologia de Alimentos, Universidade Tecnológica Federal do Paraná. Campo Mourão, 2015.
Development multiproduct multivariate calibration model to quantify the acidity and vitamin C in fruit nectar beverage and industrial soybean base.

The determination of the total acid and vitamin C using ultraviolet spectroscopy, visible near infrared spectroscopy and multiproduct multivariate calibration are presented as an alternative to the conventional method through potentiometric titration and oxidation-reduction titration, respectively. The model was developed including different products (fruit nectar of different flavors and light fruit nectar, soybean beverages and light soybean beverages). The calibration and validation data sets were optimized through outliers identification. Partial least squares (PLS) regression was used in order to build the models. Accuracy, represented by the Root Mean Squared Errors of Calibration (RMSEC) and Prediction (RMSEP) were determined. The confidence ellipse and adjust confirms that accuracy was achieved, while residues showed random behavior, confirming linearity for the model developed. Figures of merit, such as, sensitivity, analytical sensitivity, residual prediction deviation, limits of detection and quantification were evaluated. The proposed methodology shows advantages as fastness, does not requires sample preparation and does not generate toxic waste, being an alternative to the conventional method based on potentiometry and oxidation-reduction titration.

Keywords: Vitamin C, Total acidity, Multivariate calibration, Multiproduct

LISTA DE FIGURAS

Figura 1: Vitamina C.....19
Figura 2: O ânion ascorbato (a) doa um elétron ao radical livre transformando-se em um radical ascorbila (b) que é estabilizado por ressonância (c). Posteriormente, o radical ascorbila pode reagir com outro radical livre transformando-se no ácido desidroascórbico. 20
Figura 3: Representação esquemática dos diferentes tipos de dados. 27
Figura 4: Espectros UV das amostras de néctar de frutas. 36
Figura 5: Ajuste dos modelos PLS. (A) Acidez total. (B) Vitamina C. (•) amostras de calibração. (∗) amostras de validação. 38
Figura 6: Resíduos do modelo PLS. (A) Acidez total. (B) Vitamina C. (•) amostras de calibração. (∗) amostras de validação. 39
Figura 7: Espectros UV das amostras de bebidas a base de soja. (A) Espectros sem pré-processamento. (B) Espectro após primeira derivada. 41
Figura 8: Região da elipse da confiança (99%). (A) Acidez Total. (B) Vitamina C (•) Ponto onde o intercepto é igual a zero e a inclinação igual a um.....43
Figura 9: Resíduos do modelo PLS. (A) Acidez Total. (B) Vitamina C. (•) amostras de calibração. (∗) amostras de validação. 44
Figura 10: Espectros Vis-NIR das amostras de néctar de frutas e bebidas industrializadas a base de soja. (A) Espectros brutos; (B) Espectros após correção de linha de base; (C) Espectros após primeira derivada; (D) Expansão da região entre 1900 - 2500nm do espectro Vis-NIR após a primeira derivada.....46
Figura 11: Coeficientes de regressão dos modelos PLS2. (A) Espectros após correção de linha de base; (B) Espectros após primeira derivada. (—) Região Vis-NIR; (—) Região NIR.....48
Figura 12: Amostras anômalas para o modelo PLS multicomponente. (•) amostras de calibração; (∗) amostras de validação. (A) Modelo 1; (B) Modelo 2; (C) Modelo 3; (D) Modelo 4.....50
Figura 13: Ajuste. (A) Vitamina C; (B) Acidez.....51
Figura 14: Elipse de Confiança. (A) Vitamina C; (B) Acidez.....52
Figura 15: Resíduos dos modelos PLS. (A) Vitamina C; (B) Acidez. 53

LISTA DE TABELAS

Tabela 1: Quantidade de amostras utilizadas no estudo. 23
Tabela 2: Figuras de mérito do modelo multivariado para néctar de frutas. 38
Tabela 3: Figuras de mérito do modelo multivariado para bebidas industrializadas a base de soja. 43
Tabela 4: Performance dos modelos PLS. 47
Tabela 5: Resultados para identificação de outliers. 49
Tabela 6: Figuras de mérito para o modelo multivariado multicomponente. 51

LISTA DE ABREVIATURAS

CH_3COOH – Ácido Acético

DAD – Detecção por Arranjo de Diodos

DP_{cal} – Desvio padrão da calibração

DP_{val} – Desvio padrão da validação

FIA – Análise por Injeção em Fluxo

HCl – Ácido Clorídrico

NaOH – Hidróxido de Sódio

NIR – Espectroscopia de Infravermelho Próximo (do inglês, *Near-Infrared Spectroscopy*)

PC – Componente Principal (do inglês, *Principal Component*)

PCs – Componentes Principais (do inglês, *Principal Components*)

pH – Potencial Hidrogeniônico

PLS – Minimos Quadrados Parciais (do inglês, *Partial Least Squares*)

RMSEC – Raiz Quadrada do Erro Médio Quadrático da Calibração (do inglês, *Root Mean Square Error of Calibration*)

RMSECV – Raiz Quadrada do Erro Médio de Validação Cruzada (do inglês, *Root Mean Square Error of Cross Validation*)

RMSEP – Raiz Quadrada do Erro Médio Quadrático da Previsão (do inglês, *Root Mean Square Error of Prediction*)

RPD – Relação de desempenho do desvio

UV – Ultra Violeta

VL – Variável Latente

VLs – Variáveis Latentes

VIS-NIR – Visível-Infravermelho Próximo

SUMÁRIO

1 INTRODUÇÃO 16
1.1 NÉCTARES DE FRUTAS E BEBIDAS A BASE DE SOJA 17
1.2 ACIDEZ 18
1.3 VITAMINA C 19
2 OBJETIVOS 22
2.1 OBJETIVO GERAL 22
2.2 OBJETIVOS ESPECIFICOS 22
3 MATERIAIS E MÉTODOS 23
3.1 MATERIAIS E REAGENTES 23
3.2 DETERMINAÇÃO DE ACIDEZ 24
3.3 DETERMINAÇÃO DE VITAMINA C 24
3.4 DADOS ESPECTRAIS 26
3.5 CALIBRAÇÃO MULTIVARIADA, MULTIPRODUTO E PLS 27
3.6 IDENTIFICAÇÃO DE *OUTLIERS* E FIGURAS DE MÉRITO 31
3.6.1 Exatidão 33
3.6.2 Linearidade 33
3.6.3 Ajuste 33
3.6.4 Sensibilidade e sensibilidade analítica 33
3.6.5 Limites de detecção e quantificação 34
3.6.6 RPD 34
4 RESULTADOS 36
4.1 RESULTADO DA APLICAÇÃO 1: DETERMINAÇÃO DE VITAMINA C E ACIDEZ TOTAL EM NÉCTAR DE FRUTAS INDUSTRIALIZADAS EMPREGRANDO ESPECTROSCOPIA UV. 36
4.2 RESULTADO DA APLICAÇÃO 2: DETERMINAÇÃO DE ACIDEZ TOTAL E VITAMINA C EM BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA EMPREGRANDO ESPECTROSCOPIA UV. 41
4.3 RESULTADO DA APLICAÇÃO 3: DETERMINAÇÃO DE VITAMINA C E ACIDEZ TOTAL EM NÉCTARES DE FRUTA E BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA EMPREGANDO ESPECTROSCOPIA VIS-NIR. 46
5 CONCLUSÂO 55
6 PERSPECTIVA 56
7 REFERÊNCIAS 57
ANEXOS 63

APRESENTAÇÃO

Este trabalho é composto por uma introdução abordando as definições e diferenças entre sucos de frutas, néctar de frutas e bebidas industrializadas a base de soja. A introdução trata ainda dos aspectos relevantes da determinação da acidez total e do teor de vitamina C em bebidas. Na sequência são apresentados os objetivos, subdividido em objetivos gerais e específicos. São descritos os materiais e a metodologia utilizada na execução desse trabalho bem como os resultados, subdividido em três aplicações empregando a região ultravioleta (UV) do espectro eletromagnético e a espectroscopia na região do visível e infravermelho próximo (Vis-NIR): 1) Determinação de acidez total e vitamina C em néctar de frutas empregando espectroscopia UV; 2) Determinação de acidez total e vitamina C em bebidas industrializadas a base de soja empregando espectroscopia UV; 3) Determinação de acidez total e vitamina C em néctares de frutas e bebidas industrializadas a base de soja empregando a espectroscopia Vis-NIR.

As perspectivas futuras desse trabalho continuam no sentido do desenvolvimento de um modelo não linear. Ao final são listadas as referências utilizadas na elaboração deste trabalho e o anexo lista as apresentações dos resultados do trabalho em eventos científicos, além de uma publicação no *Brazilian Journal of Analytical Chemistry*, Revista Brasileira de Pesquisa em Alimentos e *Food Chemistry* (submetido).

1 INTRODUÇÃO

Com o avanço computacional e da instrumentação utilizada nas análises químicas, uma quantidade cada vez maior de informações vem sendo obtida em diversos níveis de complexidade. Além disso, as facilidades existentes atualmente para o acoplamento entre os equipamentos de medida e os microcomputadores, associadas à facilidade na estocagem de grandes quantidades de dados, leva a uma maior rapidez nas medidas e, com isso, requer e possibilita o desenvolvimento de novos procedimentos para a análise de dados analíticos através da utilização da quimiometria (LIMA et al., 2008).

A quimiometria pode ser definida como a pesquisa e utilização de métodos matemáticos e estatísticos para o tratamento de dados químicos de forma a extrair uma maior quantidade de informações e melhores resultados analíticos. A partir dos anos 1960, com a grande quantidade de dados gerados por equipamentos capazes de realizar uma obtenção relativamente rápida e com menor esforço, observou-se a necessidade de analisar todos esses dados conjuntamente e extrair maior quantidade de informações relevantes (OTTO, 1999).

A crescente preocupação mundial com relação à questão ambiental é outro aspecto de grande relevância que tem incentivado o desenvolvimento e aperfeiçoamento das análises químicas, principalmente espectroscópicas, aliadas à quimiometria. Estas análises, além de fornecerem os resultados de maneira mais rápida, não são destrutivas e invasivas, assim como não geram subprodutos químicos tóxicos. Devido a estas vantagens, este tipo de análise tem sido aplicado no monitoramento em linha de sistemas químicos industriais, na determinação de glicose, colesterol e triglicerídeos em plasma sanguíneo, no auxílio de identificação de tumores em células, no controle de qualidade, na determinação de nitrogênio em plantas, na indústria de polímeros, no estudo da composição química de solos, na adulteração da composição química de combustíveis, óleos comestíveis e alimentos, dentre outras (COSTA FILHO; POPPI, 2002).

De acordo com as vantagens dos métodos quimiométricos associados a métodos espectroscópicos, percebe-se que eles podem ser usados como potenciais substitutos das metodologias de referência empregadas em escala industrial, como

no caso da determinação da acidez e do teor de vitamina C em néctares e em bebidas industrializadas a base de soja.

1.1 NÉCTARES DE FRUTAS E BEBIDAS A BASE DE SOJA

A instrução normativa n°12, de 4 de setembro de 2003, descreve néctares de frutas como bebidas não fermentadas, obtidos da dissolução em água potável, da parte comestível do fruto e açúcares, podendo ser adicionada de ácidos, e destinada ao consumo direto (BRASIL, 2003). Levando em conta esta regulamentação, as categorias mais comuns hoje no mercado são os néctares e as bebidas à base de soja. No entanto nenhuma delas, pela legislação, pode ser denominada suco de frutas como ocorre comumente na linguagem coloquial, por isso é necessário que o rótulo informe ao consumidor o tipo exato de bebida que está sendo ofertado (PIRILO; SABIO, 2009).

Embora não seja de amplo conhecimento do consumidor mundial, a diferença entre suco e o néctar está relacionado ao teor do suco de fruta presente na bebida envasada. No mundo todo, sucos devem conter 100% de fruta *in natura*, portanto, trata-se de um produto puro, sem conservantes ou adoçantes e sem corantes artificiais, com a possibilidade ou não de conter a polpa da própria fruta (VENÂNCIO; MARTINS, 2012).

O valor nutritivo das bebidas de fruta depende principalmente do tipo de fruta utilizado, dos métodos de processamento e do grau de diluição. O conteúdo vitamínico é inferior ao do fruto original, e as perdas dependem também do tipo de fruto. Na preparação de néctares de fruta, só se retira a parte da fibra, e o seu valor calórico é maior que o dos sucos devido à adição de açúcar. O setor dos sucos e, especificamente, o caso dos néctares de fruta tem assistido a uma evolução muito satisfatória do consumo ao longo dos últimos podendo-se constatar que o setor dos néctares está em franca expansão, sendo a sua produção uma aposta empresarial com grandes possibilidades de sucesso (GUINÉ, 2012).

O mercado de bebidas a base e soja teve início em grande escala no Brasil em 1997 e, em 2008, registrou um aumento de 6,1% em relação a 2007 (SILVA, 2010), decorrente do emprego de tecnologias adequadas que tornam a bebida mais

atraente aos consumidores e, também, devido ao marketing utilizado pelas empresas para divulgar o produto (VENTURINI FILHO, 2010).

As bebidas a base de soja representam cerca de 33% das vendas dos sucos prontos, mas o mercado para esse tipo de bebida ainda encontra-se muito restrito à parcela da população de maior poder aquisitivo, principalmente na região sul e sudeste do Brasil. Entretanto, as classes economicamente menos privilegiadas estão começando a descobrir esses tipo de produto, mostrando um grande potencial de crescimento nas vendas (VENTURINI FILHO, 2010).

Considerando o crescimento no consumo dessas bebidas pela população é relevante a quantificação de alguns constituintes e parâmetros, como a quantidade de vitamina C a acidez total; uma vez que estes fornecem informações de interesse social, como a qualidade da bebida em termos de conservação, a atividade antioxidante – que auxilia na redução de risco contra o câncer (MERCALDI, 2006), entre outras.

1.2 ACIDEZ

Acidez e vitamina C são parâmetros de qualidade responsáveis pelo aroma, sabor, características sensoriais e nutricionais, e o estado de conservação do produto alimentício (VENÂNCIO; MARTINS, 2012).

A determinação de acidez pode fornecer um dado valioso na apreciação do estado de conservação de um produto alimentício. Um processo de decomposição, seja por hidrólise, oxidação ou fermentação, altera quase sempre a concentração dos íons hidrogênio. Os métodos de determinação da acidez titulável fornecem a concentração de íons de hidrogênio livre, por meio do pH. Os métodos que avaliam a acidez titulável se resumem em titular com soluções de álcali-padrão a acidez do produto ou de soluções aquosas ou alcoólicas do produto e, em certos casos, os ácidos graxos obtidos dos lipídios. A quantidade de acidez total em bebidas pode ser expressa em volume (mL) de solução normal por cento ou em gramas do componente ácido principal (IAL, 2008).

A acidez é definida pelo número de prótons tituláveis (SKOOG et al., 2006) e, segundo a *Federation International des Producteurs de Jus de Fruits* (2005),

determinada através de potenciometria. Uma titulação potenciométrica envolve medidas do potencial de um eletrodo indicador adequado em função do volume do titulante. Nesse sentido, o eletrodo de vidro é o eletrodo indicador mais importante para os íons hidrogênio. A informação fornecida por uma titulação potenciométrica não é a mesma daquela obtida a partir de uma medida potenciométrica direta. Isso significa que a medida direta de soluções 0,100 mol/L de ácido clorídrico (HCl) e ácido acético (CH_3COOH) deveria gerar duas concentrações de íons hidrogênio substancialmente diferentes porque HCl é considerado um ácido forte e se dissocia totalmente, enquanto que CH_3COOH, um ácido fraco, se dissocia apenas parcialmente. Em contraste, a titulação potenciométrica de volumes iguais dos dois ácidos requer a mesma quantidade da base padrão porque ambos os solutos têm o mesmo número de prótons tituláveis (SKOOG et al., 2006).

A literatura reporta a determinação de acidez atravez de métodos de análise por injeção em fluxo (GRUDPAN; SRITHARATHIKHUN; JAKMUNEE, 1998) e espectroscopia no infravermelho médio (CAMPS et al., 2010). Porém, nesse trabalho a potenciometria foi empregada como método de referência.

1.3 VITAMINA C

O ácido ascórbico, mais conhecido como vitamina C (Figura 1), é um composto cristalino, que apresenta sabor ácido e pode ser encontrado em frutas cítricas; além disso, ela é bastante solúvel em água e praticamente insolúvel em solventes orgânicos. Essa vitamina é bastante importante para o organismo devido às funções que desempenha. (LAVARDA, 2011).

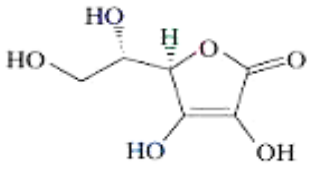

Figura 1: Vitamina C

Fonte: LIMA; FUKURO (2014)

A vitamina C foi isolada em 1928, pelo húngaro Albert Szent-Gyorgyi e sintetizada pela primeira vez em 1938, quando passou a ser reconhecido oficialmente como ácido ascórbico, por ser eficaz na prevenção e tratamento do escorbuto, doença que é responsável pela fragilidade dos vasos sanguíneos e danos ao tecido conectivo (TAVARES et al., 2000).

As principais fontes de ácido ascórbico encontram-se nas frutas e vegetais destacando-se a laranja, a tangerina, o limão, a acerola, a cereja, o melão e o abacaxi. Entre os vegetais, destacam-se o tomate, a couve, o pimentão e o brócolis (LAVARDA, 2011).

A vitamina C possui diversas funções, entre as principais pode-se citar a de antioxidante, devido a sua capacidade de ceder elétrons (Figura 2), como no caso de regeneração da forma ativa da vitamina E, é necessário para a síntese e manutenção do colágeno e também está envolvida na síntese de importantes neurotransmissores como a norepiefrina obtida a partir da dopamina e a serotonina, obtida pela conversão de triptofano em 5-hidroxitriptofano (REBOUCHE, 1991; NAGAOKA et al., 2007, APUD SANTOS, 2008).

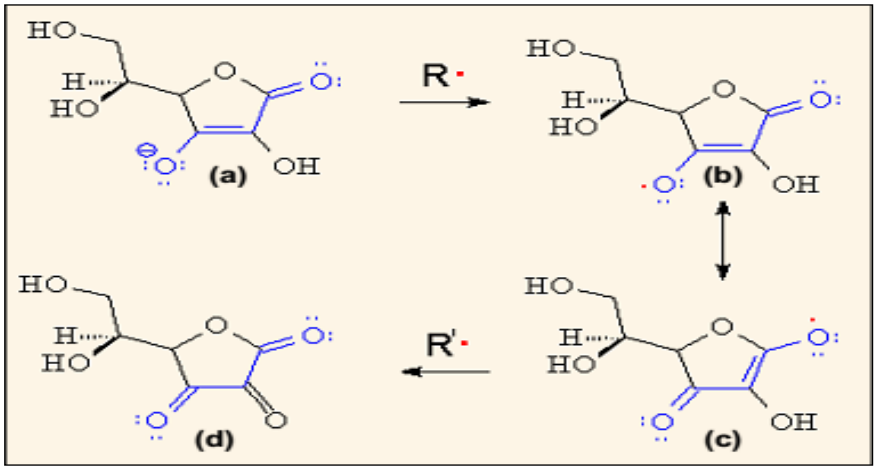

Figura 2: O ânion ascorbato (a) doa um elétron ao radical livre transformando-se em um radical ascorbila (b) que é estabilizado por ressonância (c). Posteriormente, o radical ascorbila pode reagir com outro radical livre transformando-se no ácido desidroascórbico.
Fonte: LIMA e FUKURO (2012)

Essa vitamina é importante porque atua na diminuição do risco de doenças cardiovasculares, no tratamento de hipertensão e na redução da incidência de cataratas. Estudos recentes também indicam que a vitamina C pode atuar na

prevenção do câncer, pois acredita-se que exista uma relação inversa entre frequência de tumores e consumo de alimentos ricos nesta vitamina. A vitamina C também é recomendada como suplemento alimentar, pois auxilia na absorção de minerais como ferro e zinco (CATANIA; BARROS; FERREIRA, 2009). Devido a esses benefícios proporcionados pela vitamina C, torna-se relevante identificar a quantidade desse composto nos alimentos.

Além das vantagens que o consumo de ácido ascórbico traz para o organismo, a sua quantidade nos produtos indica a qualidade nutricional e o estado de conservação dos mesmos. Por esse motivo, a vitamina C é usada como índice de qualidade nutricional de produtos derivados de frutas e vegetais, porque quando comparado a outros nutrientes, esta vitamina é mais sensível à degradação durante o processamento e subsequente estocagem (DANIELLI et. al., 2009).

Os principais fatores que podem influenciar na degradação da vitamina C em suco de fruta incluem o tipo de processamento, as condições de estocagem, o tipo de embalagem, a exposição ao oxigênio e a luz, a temperatura, a presença de catalisadores metálicos e enzimas e o pH (TEIXEIRA; MONTEIRO, 2006).

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo desse trabalho é desenvolver uma metodologia alternativa para a determinar a acidez total e o teor de vitamina C em néctares de frutas e bebidas a base de soja, utilizando a espectroscopia na região do ultravioleta, visível infravermelho próximo e calibração multivariada. Buscou-se desenvolver uma metodologia alternativa que apresentasse vantagens em relação ao custo, não utilização de reagentes químicos e, principalmente, técnicas que, além de apresentarem caráter não destrutivo da amostra, sejam suficientemente rápidas para serem implantadas na indústria de alimentos.

2.2 OBJETIVOS ESPECIFICOS

- Realizar a titulação de oxidação-redução e potenciométrica de néctares de frutas e bebidas a base de soja adquiridos em comércio local;
- Adquirir espectros da região ultravioleta, visível infravermelho próximo dos néctares de frutas e bebidas a base de soja;
- Construir um modelo de calibração multivariada para néctares de frutas e bebidas a base de soja;
- Construir um modelo de calibração multivariada multiproduto multicomponente.
- Identificar e eliminar possíveis *outliers*;
- Validar os modelos multivariados desenvolvidos através do cálculo de figuras de mérito tais como: exatidão, ajuste, linearidade, sensibilidade, sensibilidade analítica, relação de desempenho do desvio, limites de detecção e quantificação.

3 MATERIAIS E MÉTODOS

3.1 MATERIAIS E REAGENTES

Os seguintes materiais e reagentes foram empregados no desenvolvimento da parte experimental dessa dissertação: Papel de filtro qualitativo (Nalgon, 15cm, três micras de porosidade); dessecador; balança analítica; pipetas graduadas (5 e 10mL) e volumétricas; Funil de vidro; suporte para funil e bureta; Frasco erlenmeyer de 125mL; bureta de 25mL; potenciômetro de bancada, agitador magnético. Água destilada; Ácido Ascórbico (Impex); sal sódico do 2,6-diclorofenol indofenol; Indigo Carmim (Sigma-Aldrich); Ácido metafosfórico (Vetec); Ácido Acético Glacial (Vetec); Hidróxido de Sódio (Alphatec); Ácido Clorídrico (Vetec); Fenolftaleína; indicador fenolftaleína a 1%; solução de hidróxido de sódio 0,1mol/L; biftalato de potássio e um total de 123 amostras de néctares de frutas e 126 amostras de bebidas industrializadas a base de soja, cujos sabores estão descritos na tabela 1.

Tabela 1: Quantidade de amostras utilizadas no estudo.

Sabores	Número de amostras (Néctar)	Número de amostras (Soja)
Pêssego / Pêssego light	10 / 06	15 / 06
Uva / Uva light	09 / 06	15 / 09
Laranja / Laranja light	09 / 06	09 / 03
Manga / Manga light	09 / 06	03 / 00
Limão	03	03
Maçã / Maçã light	12 / 00	15 / 06
Maracujá / Maracujá light	09 / 03	06 / 00
Goiaba / Goiaba light	06 / 06	00 / 00
Caju / Caju light	06 / 03	00 / 00
Uva Branca	03	00
Abacaxi	06	21
Pera	03	00
Morango	03	06
Romã	00	03
Mamão	00	03
Tangerina	00	03
Total de Amostras	**123**	**126**

3.2 DETERMINAÇÃO DE ACIDEZ

Para determinação da acidez foram utilizadas um total de 249 amostras de bebidas industrializadas a base de soja e néctar de frutas (Tabela 1) adquiridas no comércio de Campo Mourão-Paraná.

O teor de acidez total foi realizado seguindo a metodologia proposta pela *Federation International des Producteurs de Jus de Fruit* (2005). Essa metodologia é baseada na titulação potenciométrica das amostras de suco utilizando-se NaOH na concentração de 0,1mol/L, padronizado com HCl 0,1mol/L e indicador de fenolftaleína.

Inicialmente, montou-se um sistema utilizando pHmetro, agitador magnético e uma bureta de 50mL contendo uma solução de NaOH 0,1mol/L. Em seguida retirou-se uma porção de 100mL de bebida (néctar ou soja) transferindo para um béquer de 250mL, com barra magnética (peixinho). Adicionou-se lentamente a solução de NaOH 0,1mol/L até o pH de 8,1. A medida foi realizada em triplicata para cada variedade de bebida e a acidez foi determinada de acordo com a equação 1.

$$acidez\left(\frac{g}{100mL}\right) = \frac{(6{,}4.V_{NaOH}.C_{NaOH})}{V_{amostra}} \quad (1)$$

3.3 DETERMINAÇÃO DE VITAMINA C

Para determinação de vitamina C foram utilizadas 249 amostras de bebidas industrializadas a base de soja e néctar de frutas (Tabela 1) adquiridas no comércio de Campo Mourão-Paraná.

O teor de vitamina C (ácido ascórbico) foi determinado, pelo método de Tillmans, seguindo a metodologia proposta pelo Instituto Adolfo Lutz (2008).

Inicialmente foram previamente preparadas soluções para o decorrer do experimento. Foram preparadas soluções ácidas, onde dissolveu-se 15g de ácido

metafosfórico em 40mL de ácido acético, em seguida adicionou-se 450mL de água destilada, sendo agitado e em seguida filtrado. Uma segunda solução foi preparada para o padrão de vitamina C, onde pesou-se 100mg de vitamina C, previamente dessecada, dissolvida em 100mL de solução ácida em um balão volumétrico, sendo diluída 10 vezes na mesma solução ácida. Para a terceira solução, a solução de Tillmans, foi dissolvido 42mg de bicarbonato de sódio em 50mL de água destilada, adicionou-se 50mg de 2,6-diclorofenol indofenol em agitação até a dissolução do corante. Diluiu e filtrou-se em 200mL em água destilada em balão volumétrico.

Após o preparo das soluções padronizou-se a solução de Tillmans. Em um erlenmeyer de 250mL, pipetou-se 4mL da solução diluída da vitamina C e 6mL da solução ácida, adicionou-se 50mL de água destilada e titulou com a solução de Tillmans até obter a coloração ligeiramente rosada e estável por 15 segundos, realizou-se o teste do branco, substituindo a solução de vitamina C por solução ácida com o intuito de se calcular o fator (F) da solução de Tillmans conforme a equação 2.

$$F = \frac{Massa\ (mg)\ de\ vitamina\ C\ na\ titulação}{Volume\ (mL)\ de\ solução\ de\ Tillmans\ gasto} \quad (2)$$

Filtrou-se 40mL das bebidas (néctar e soja), com quantidades iguais de solução ácida, retirando-se alíquotas de 10mL do filtrado conforme descrito na padronização da solução de Tillmans, realizando a mesma titulação para o branco constituído de 10mL da solução ácida e com volume de água igual ao da solução do corante gasto na titulação da amostra.

Em seguida foram realizados testes de interferentes de íons Fe^{2+}, Sn^{2+} e Cu^{2+}. Adicionou-se duas gotas de solução de azul de metileno 0,05% a 10mL da mistura 1:1 constituída das amostras de bebidas (néctar e soja) e do reagente ácido.

O teor de vitamina C foi determinado através do cálculo, equação 3.

$$\frac{V \times F \times 100}{A} = Ácido\ ascórbico\ mg/100mL \quad (3)$$

Em que V é o volume da solução de Tillmans gasto na titulação; F é o fator da solução de Tillmans; A é quantidade em volume (mL) da amostra utilizada.

3.4 DADOS ESPECTRAIS

A realização das medidas espectrais registradas na região Ultravioleta (UV, 200 – 400nm, intervalo de 1nm) foi realizada em cubeta de quartzo de 1mm de caminho óptico. Esta cubeta foi utilizada por evitar a etapa de preparo de amostra por diluição, já que aumenta a sensibilidade espectral, melhorando a visualização do perfil espectral por diminuir a quantidade de espécies absorventes. Mesmo assim, para as medidas a base de soja foi necessário uma etapa prévia de diluição (300µL:10mL). As medidas foram realizadas com um equipamento da marca Ocean Optics, modelo USB-650-UV-VIS.

As medidas espectrais na região Visível-Infravermelho Próximo (Vis-NIR, 400 – 2500nm, intervalo de 0,0005nm) foram realizadas em um equipamento da marca Methrom Pensalab, modelo DS2500. Para realização das medidas foi empregado um conjunto acessório de quartzo e ouro com caminho óptico de 1mm. Nesse caso, não foi necessário a etapa de diluição das amostras de bebidas industrializadas a base de soja.

O software empregado nas análises foi o Matlab R2007b empregando-se o PLS Toolbox 5.2 (Eingevector Research Inc.).

Após a determinação do teor de acidez total e vitamina C através dos métodos de referência, foi empregado o método quimiométrico de calibração multivariada para dados de primeira ordem, denominado mínimos quadrados parciais (PLS). Para tanto, os espectros UV ou Vis-NIR foram organizados no formato de uma matriz, em que cada linha da matriz **X** corresponde a um espectro e cada coluna corresponde à absorbância nos diferentes comprimentos de onda. A matriz **X** foi então correlacionada a um vetor **y** que continha os valores de acidez ou vitamina C determinados por titulação e esta correlação entre **X** e **y** foi realizada através do método de calibração multivariada PLS.

3.5 CALIBRAÇÃO MULTIVARIADA, MULTIPRODUTO E PLS

Atualmente os equipamentos disponíveis permitem a geração de dados cada vez mais complexos. A Figura 3 mostra uma representação esquemática dos tipos de dados que são atualmente obtidos com muita facilidade nos laboratórios de análises.

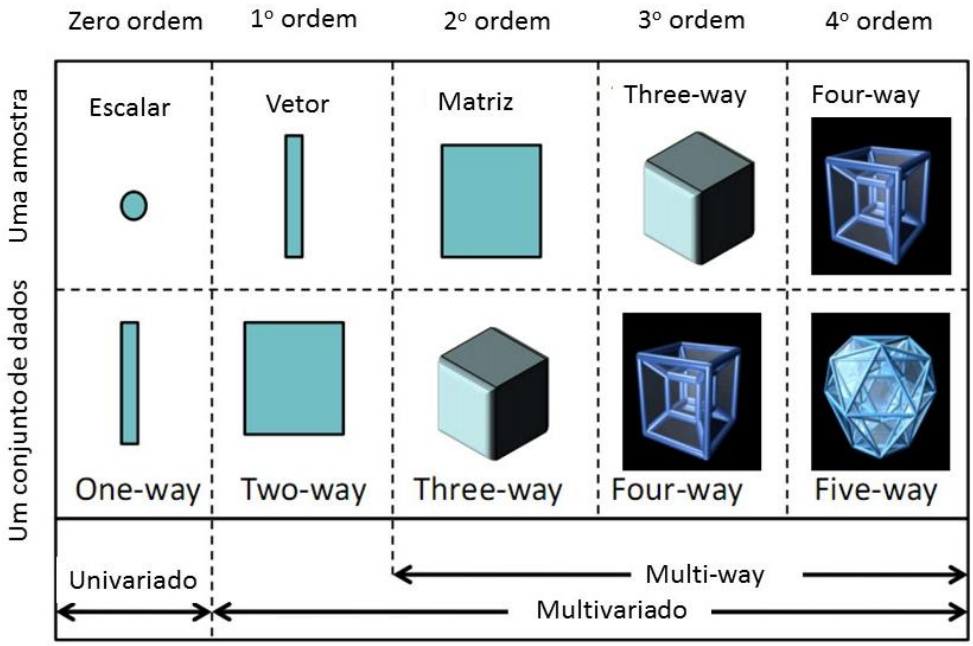

Figura 3: Representação esquemática dos diferentes tipos de dados.
Fonte: adaptado de ESCANDAR *et al.* (2014)

Em modelos de calibração de ordem zero é obtido um único valor escalar por amostra, sendo classificada como uma calibração univariada. Um exemplo típico de instrumentos que geram esse tipo de dados inclui os fotômetros e espectrofotômetros com monitoramento de uma única variável. Esse tipo de calibração ainda é o mais aplicado em análises de rotina, dispondo de procedimentos de validação bem estabelecidos e descritos na literatura em diversos trabalhos (MILLER; MILLER, 2000; RIBANI et al., 2004; VALDERRAMA; BRAGA; POPPI, 2009). Entretanto, a aplicação de calibrações univariadas requer que a grandeza que é medida diretamente no sistema, por exemplo, a absorbância de um composto, seja livre de interferentes que possam provocar desvios entre sua relação com a propriedade de interesse.

Métodos de calibração de primeira ordem fazem uso de um vetor de medidas instrumentais para cada amostra. Exemplos deste tipo de dados incluem as medidas espectroscópicas na região do ultravioleta e visível, infravermelho médio e próximo, os cromatogramas, etc. Esses métodos possibilitam análises mesmo na presença de interferentes desde que eles estejam presentes nas amostras de calibração, determinação simultânea e análises sem resolução (VALDERRAMA; BRAGA; POPPI, 2009).

Ainda de acordo com a Figura 3, verifica-se a possibilidade de obtenção de uma matriz de dados por amostra, de modo que a calibração envolvendo esse tipo de dados é chamada de calibração de segunda ordem. Esse tipo de dados pode ser gerado através de diversas técnicas como, por exemplo, a cromatografia líquida ou gasosa com detecção por espectrômetro de massas ou arranjo de diodos, respectivamente (HPLC-MS e CG-MS, HPLC-DAD), fluorescência molecular de excitação e emissão, análise por injeção em fluxo com gradiente de pH e detecção por arranjo de diodos (VALDERRAMA; BRAGA; POPPI, 2009). Esses métodos têm a grande vantagem de permitirem a determinação de espécies de interesse na presença de interferentes, mesmo que estes interferentes não tenham sido incluídos nas amostras de calibração. Essa característica é a chamada vantagem de segunda ordem. Além disso, o perfil de cada composto linearmente independente, presente na amostra, pode ser estimado com esse tipo de dados (BOOKSH; KOWALSKI, 1994).

Os dados de terceira e quarta ordem podem ser gerados a partir de instrumentos hifenados tais como: cromatografia gasosa bidimensional abrangente com detector de espectrometria de massas ou ainda através da fluorescência molecular de excitação e emissão monitorada em diferentes valores de pH e em diferentes tempos de reação (MAGGIO; MUÑOZ DE LA PEÑA; OLIVIERI, 2011).

A utilização de um modelo de calibração multivariada para cada tipo de produto apresenta desvantagens, como o número de etapas computacionais requeridas em cada previsão de um produto em análises de rotina. Ainda, um modelo de calibração específico para cada produto deve ser desenvolvido para uma pequena população de amostras específicas e cada nova amostra deve ser claramente identificada de modo a ter seus componentes previstos pelo modelo multivariado correto (BERZAGHI; SHENK; WESTERHAUS, 2000) e, a manutenção dos modelos de calibração multivariada pode ser algo trabalhoso. Assim, é preciso

pensar em alternativas que prolonguem o uso de um determinado modelo, economizando tempo em atualizações. Para contornar essas desvantagens, uma alternativa ainda pouco explorada consiste na construção de modelos de calibração multivariada multiproduto (MICKLANDER et al., 2006).

A calibração multivariada multiproduto utiliza diferentes tipos de amostras em um mesmo modelo de calibração multivariada visando à determinação de um mesmo analito de interesse (BERZAGHI; SHENK; WESTERHAUS, 2000). Este tipo de calibração é recente podendo apresentar como vantagens modelos mais robustos, fáceis de atualizar, além da praticidade.

Os primeiros trabalhos reportados na literatura acerca da calibração multivariada multiproduto datam de 1992 (NAES; ISAKSSON, 1992) e 1994 (WANG; ISAKSSON; KOWALSKI, 1994) e foram desenvolvidos em conjuntos de dados com relativa homogeneidade, visando avaliar a performance de novos algoritmos.

Berzaghi e colaboradores (2000) avaliaram um conjunto de dados não homogêneo na construção de um modelo multiproduto. O objetivo do trabalho desses autores era avaliar a performance de um método de regressão denominado de LOCAL. Micklander e colaboradores (2006) testaram cinco diferentes métodos de regressão (linear e não-linear) no desenvolvimento de modelos de calibração multivariada multiproduto. Os autores concluíram que os métodos de regressão não lineares apresentavam menores erros médios quadráticos de previsão (RMSEP) que o método PLS.

Atualmente, a indústria alimentícia ainda é um pouco reticente na implantação de espectroscopia combinada com quimiometria em suas rotinas de controle de qualidade. Além disso, ainda muitas indústrias adquirem os modelos de calibração multivariada previamente desenvolvidos utilizando amostras que não fazem parte de suas realidades industriais. Com isso, os resultados obtidos empregando esses modelos não são satisfatórios (SANTOS; MARÇO; VALDERRAMA, 2013). Isso porque, modelos PLS podem ser construídos na presença de interferentes, desde que esses interferentes estejam presentes na etapa de calibração (BRERETON, 2000). Se com o PLS, que é um método linear, todos esses problemas já são vivenciados, a utilização de métodos de regressão não lineares ou matematicamente mais complexos pode comprometer ainda mais negativamente a implantação desses métodos em aplicações práticas. Visando reduzir o número de etapas computacionais e facilitar o trabalho de rotina industrial

pretende-se avaliar a potencialidade do método PLS na construção de modelos de calibração multivariada multiproduto.

PLS (do inglês, *Partial Least Squares*) (GELADI; KOWALSKI, 1986) é baseado no método de reconhecimento de padrões não supervisionado conhecido como Análise de Componentes Principais (PCA, do inglês *Principal Component Analysis*) (WOLD, 1987). Ambos são métodos quimiométricos aplicáveis a dados de primeira ordem, ou seja, quando se tem um vetor de respostas instrumentais para cada amostra (VALDERRAMA; BRAGA; POPPI, 2009), como, por exemplo, um espectro na região do UV ou Vis-NIR. Na utilização desses métodos, os espectros UV ou Vis-NIR são organizados no formato de uma matriz **X**, em que cada linha da matriz corresponde a um espectro e cada coluna corresponde à absorbância nos diferentes comprimentos de onda.

Considerando um caso geral para a determinação de mais de uma espécie de interesse, tem-se uma outra matriz, denominada **Y**, que contém as informações dessas propriedades. No PLS, as matrizes **X** e **Y** são decompostas em um produto de duas matrizes, uma matriz de *scores* (**T** e **U**) e uma matriz de *loadings* (**P** e **Q**), conforme equações (4) e (5) (OTTO, 1999).

$$\mathbf{X}=\mathbf{T}\mathbf{P}^{\mathbf{T}} \tag{4}$$

$$\mathbf{Y}=\mathbf{U}\mathbf{Q}^{\mathbf{T}} \tag{5}$$

O conjunto de *scores* e *loadings* é denominado de Componentes Principais (PCs) que, no PLS sofre pequenas modificações para buscar a máxima covariância entre **X** e **Y**, assim, os PCs que são obrigatoriamente ortogonais, no PLS recebem a terminologia de Variáveis Latentes (VLs) por perderem sua ortogonalidade. A decomposição das matrizes em *scores* e *loadings* é realizada por algoritmos, como por exemplo, o SIMPLS (JONG, 1993), e após a decomposição os coeficientes de regressão para o modelo são estimados por mínimos quadrados entre os *scores* e *loadings* de **X** e os *scores* de **Y** (OTTO, 1999). Nesse caso será obtido um vetor de coeficientes de regressão, ou seja, um coeficiente de regressão para cada comprimento de onda utilizado na matriz **X** (GELADI; KOWALSKI, 1986).

Na escolha adequada do número de variáveis latentes para composição do modelo, um procedimento de validação cruzada (*Cross Validation*) normalmente é empregado (BRERETON, 2000). Nesse trabalho, a validação cruzada foi realizada

através de blocos contínuos de 9 amostras. Neste caso, as primeiras 9 amostras são deixadas de fora no processo de construção do modelo e a seguir essas amostras são previstas pelo modelo construído. Esse processo se repete até que todas as amostras tenham sido deixadas de fora e previstas. O número de variáveis latentes pode ser então escolhido com base no valor da raiz quadrada do erro médio quadrático de validação cruzada (RMSECV, *Root Mean Square Error of Cross Validation*), conforme a equação 6.

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (6)$$

Em que, y_i é o valor de referência para a propriedade de interesse da amostra "i", $\hat{y}_i$ é o valor da propriedade de interesse da amostra "i" previsto pelo modelo e n é o número de amostras previstas no processo de validação cruzada.

3.6 IDENTIFICAÇÃO DE *OUTLIERS* E FIGURAS DE MÉRITO

Outliers é o termo utilizado para designar amostras anômalas que podem estar presentes nos conjuntos de calibração e validação. Normalmente esse tipo de amostra possui um comportamento diferente das demais amostras do conjunto de dados e a sua presença no conjunto de calibração pode conduzir a modelos com baixa capacidade de previsão, produzindo elevados valores de erro. Quando presentes no conjunto de validação, podem influenciar, geralmente, levando a resultados que indicam que o modelo não é adequado ou que a sua capacidade é inferior à que poderia ser apresentada na ausência das amostras anômalas. Nesse sentido, a identificação de *outliers* é um passo importante para a otimização dos conjuntos de calibração e validação, sendo que, a exclusão destes permite a construção de modelos mais eficientes e precisos e com melhor capacidade de previsão (VALDERRAMA; BRAGA; POPPI, 2007b; MARTENS; NAES, 1996).

A presença de *outliers* pode ocorrer por diferentes razões, tais como, erros nas medidas experimentais no laboratório, erros instrumentais, amostras pertencentes a outra população (MARTENS; NAES, 1996). Nas aplicações 1 e 2

dessa dissertação, os *outliers* foram identificados com base no *leverage* (ASTM, 2005), resíduos espectrais não modelados (MARTENS; NAES, 1996; ASTM, 2005) e resíduos não modelados na variável dependente (MARTENS; NAES, 1996). O *leverage*, os resíduos espectrais não modelados e os resíduos não modelados na variável dependente foram avaliados para o conjunto de calibração, enquanto que na validação os *outliers* foram avaliados com base no *leverage* e nos resíduos espectrais não modelados.

Não é incomum, quando os *outliers* são eliminados em um primeiro modelo e este é reconstruído, novos *outliers* serem detectados no modelo reconstruído (segundo modelo) (VALDERRAMA; BRAGA; POPPI, 2007b). Nestes casos, o teste pode ser “relaxado”: 1) O primeiro modelo foi construído. 2) Os *outliers*, identificados com base no *leverage*, nos resíduos espectrais não modelados e nos resíduos na variável dependente foram removidos. 3) Um segundo modelo foi construído, com o mesmo número de variáveis latentes. 4) Os *outliers*, identificados com base no *leverage*, nos resíduos espectrais não modelados e nos resíduos na variável dependente foram removidos. 5) Um terceiro modelo foi construído, com o mesmo número de variáveis latentes. 6) Os *outliers* foram avaliados e removidos do conjunto de validação.

Para a terceira aplicação dessa dissertação, os *outliers* foram avaliados com base no gráfico do *leverage* versus Q resíduos (referente aos resíduos espectrais) (SILVA et al., 2014). Nesta aplicação, os seguintes passos foram seguidos para a otimização dos conjuntos de calibração e validação: 1) O primeiro modelo foi construído. 2) Os *outliers*, identificados com base no gráfico do *leverage* versus Q resíduos foram removidos. 3) Um segundo modelo foi construído. 4) Os *outliers*, identificados com base no gráfico do *leverage* versus Q resíduos foram removidos. 5) Um terceiro modelo foi construído. 6) Os *outliers*, identificados com base no gráfico do *leverage* versus Q resíduos foram removidos. 7) Um quarto modelo foi construído. 8) Os *outliers* foram avaliados e removidos do conjunto de validação.

A validação dos modelos de calibração multivariada foi realizada através da determinação das figuras de mérito: exatidão, linearidade, ajuste, sensibilidade, sensibilidade analítica, limites de detecção e quantificação. O cálculo foi realizado com base em trabalhos descritos na literatura (VALDERRAMA; BRAGA; POPPI, 2007; VALDERRAMA; BRAGA; POPPI, 2009). A relação de desempenho do desvio

(RPD, *Residual Prediction Deviation*) foi avaliada de acordo com (BOTELHO; MENDES; SENA, 2013).

3.6.1 Exatidão

A exatidão expressa o grau de concordância entre o valor estimado pelo modelo multivariado e o valor tido como referência. Em calibração de primeira ordem a exatidão pode ser avaliada com base nos valores de RMSEP e RMSEC, além da avaliação do ajuste e em alguns casos, da elipse de confiança.

3.6.2 Linearidade

A avaliação da linearidade é problemática em calibração multivariada empregando o método PLS. Isso porque as variáveis são decompostas pelos componentes principais. Assim, a linearidade é avaliada qualitativamente a partir do gráfico dos resíduos de previsão para as amostras de calibração e validação. Quando este gráfico apresenta um comportamento aleatório, isto é um indicativo que os dados seguem um comportamento linear.

3.6.3 Ajuste

O ajuste foi avaliado a partir da correlação entre os valores tidos como referência e os valores estimados pelo modelo de calibração multivariada, podendo ser considerado também um indicativo de exatidão do modelo.

3.6.4 Sensibilidade e sensibilidade analítica

Em modelos de calibração multivariada baseados no método PLS a sensibilidade pode ser avaliada a partir do vetor de coeficientes de regressão (**b**):

$$\boldsymbol{sensibilidade} = \frac{\mathbf{1}}{\|\mathbf{b}\|} \tag{7}$$

A sensibilidade analítica, apresenta a sensibilidade do método em termos de unidades de concentração e é definida como a razão entre a sensibilidade e o desvio padrão do sinal de referência (δx):

$$\boldsymbol{sensibilidade\ analítica} = \frac{\boldsymbol{sensibilidade}}{\delta \mathbf{x}} \tag{8}$$

O inverso da sensibilidade analítica (sensibilidade analítica^{-1}) permite estabelecer a menor diferença de concentração entre amostras, que pode ser distinguida pelo método.

3.6.5 Limites de detecção e quantificação

O limite de detecção e o limite de quantificação de um procedimento analítico, expressam as menores quantidades da espécie de interesse que podem ser detectadas e determinadas quantitativamente, respectivamente. Esses limites são calculados a partir da sensibilidade e do desvio padrão do sinal de referência (δx):

$$\boldsymbol{Limite\ de\ detecção} = \mathbf{3,3}\ \delta\mathbf{x}\ \frac{\mathbf{1}}{\boldsymbol{sensibilidade}} \tag{9}$$

$$\boldsymbol{Limite\ de\ quantificação} = \mathbf{10}\ \delta\mathbf{x}\ \frac{\mathbf{1}}{\boldsymbol{sensibilidade}} \tag{10}$$

3.6.6 RPD

A relação de desempenho do desvio (do inglês, *Residual Prediction Deviation*) é uma figura de mérito mais recente, utilizada para estimar a capacidade preditiva dos métodos. Essa figura de mérito é mais adequada para comparações em termos absolutos e foi estimada para os conjuntos de calibração e validação de acordo com as Equações 11 e 12, onde RMSECV é o erro quadrático médio de

validação cruzada e DP_{cal} e DP_{val} são os desvios padrão dos valores de referência dos conjuntos de calibração e validação, respectivamente.

$$\boldsymbol{RPD_{cal}} = \frac{\boldsymbol{DP_{cal}}}{\boldsymbol{RMSECV}} \tag{11}$$

$$\boldsymbol{RPD_{val}} = \frac{\boldsymbol{DP_{val}}}{\boldsymbol{RMSEP}} \tag{12}$$

4 RESULTADOS

4.1 RESULTADO DA APLICAÇÃO 1: DETERMINAÇÃO DE VITAMINA C E ACIDEZ TOTAL EM NÉCTAR DE FRUTAS INDUSTRIALIZADAS EMPREGRANDO ESPECTROSCOPIA UV.

Na determinação do teor de vitamina C e acidez total por espectroscopia na região do UV obtém-se um vetor de respostas instrumentais para cada amostra analisada e, portanto modelos de calibração multivariada de primeira ordem através de mínimos quadrados parciais (PLS) foram empregados. A Figura 4 apresenta os espectros na região do UV obtidos para todas as amostras analisadas.

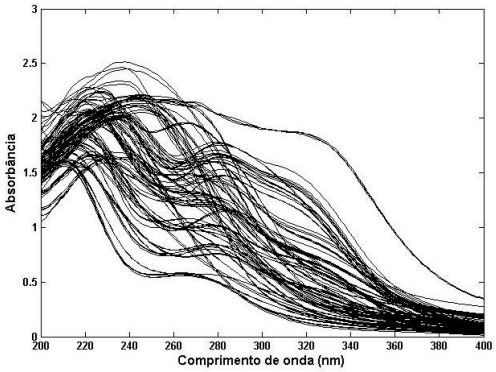

Figura 4: Espectros UV das amostras de néctar de frutas.
Fonte: Acervo do autor (2014)

Na construção dos modelos de calibração as amostras foram separadas entre conjunto de calibração e validação através do algoritmo de Kennard-Stone (KENNARD; STONE, 1969). Este algoritmo seleciona as amostras com base em suas distâncias e é de fácil aplicação. A primeira amostra selecionada é a que apresenta a maior distância em relação à amostra média. A segunda amostra a ser selecionada será a que apresentar maior distância em relação à primeira amostra selecionada. A próxima amostra a ser selecionada apresentará maior distância em relação à última amostra selecionada, e assim sucessivamente até atingir o número de amostras desejadas. Esse algoritmo é aplicado para realizar a seleção das amostras que irão compor o conjunto de calibração, uma vez que procede a seleção

das amostras de maior variabilidade, ou seja, as amostras mais “externas” do conjunto total. Assim, o conjunto de calibração foi inicialmente composto por 75% das amostras e o conjunto de validação foi composto por 25% das amostras.

O próximo passo consistiu na análise para identificação de amostras anômalas, onde no conjunto de calibração, as amostras consideradas anômalas com base nos testes de *leverage*, resíduos espectrais não modelados e resíduos na variável dependente eram eliminadas e o modelo de calibração novamente reconstruído. Após a otimização dos modelos de calibração, os testes para identificação de amostras anômalas foram aplicados a cada conjunto de validação e os *outliers* para esses conjuntos também foram eliminados.

A otimização dos conjuntos de calibração e validação pela eliminação das amostras anômalas resultaram em 81 e 79 amostras de calibração e 16 e 25 amostras de validação para os modelos PLS na determinação de acidez total e vitamina C, respectivamente.

Os modelos foram centrados na média, construídos com 9 variáveis latentes na determinação de acidez total e 3 variáveis latentes na determinação de vitamina C. O número de variáveis latentes foi determinado através dos resultados da raiz quadrada do erro médio quadrático de validação cruzada (RMSECV) para as amostras de calibração, obtido por validação cruzada em blocos contínuos de 9 amostras. A percentagem de variância acumulada explicada no bloco **y** também foi levada em consideração na escolha do número de variáveis latentes e apresentou valores de 71,19% e 88,90% para o modelo otimizado na determinação da acidez total e vitamina C, respectivamente.

Um dos parâmetros avaliados na exatidão dos modelos desenvolvidos foram os valores do erro médio quadrático de calibração (RMSEC) e previsão (RMSEP), apresentados na Tabela 2, para os modelos após otimização.

Tabela 2: Figuras de mérito do modelo multivariado para néctar de frutas.

Figuras de mérito		Vitamina C	Acidez
Exatidão	RMSEC[a]	4,2504	0,0641
	RMSEP[a]	3,0945	0,0534
Coeficiente de correlação		0,8890	0,7119
RPD_{cal}		2,8	1,6
RPD_{val}		2,4	1,5
Sensibilidade analítica^{-1} [a]		0,3315	0,0057
Limite de detecção[a]		1,0941	0,0188
Limite de quantificação[a]		3,3153	0,0570

[a] (mg/100mL)

Fonte: Acervo do autor (2014)

Os resultados de RMSEC e RMSEP indicam que o número de variáveis latentes utilizados em cada um dos modelos estão corretas, não havendo sobreajuste ou subajuste. Entretanto, os valores de RMSEP e RMSEC são parâmetros globais que incorporam tanto erros sistemáticos quanto aleatórios. Sendo assim, um outro indicador de exatidão é a regressão entre os valores de referência e os valores estimados pelo modelo, e os valores da inclinação e do intercepto, apresentados na Figura 5.

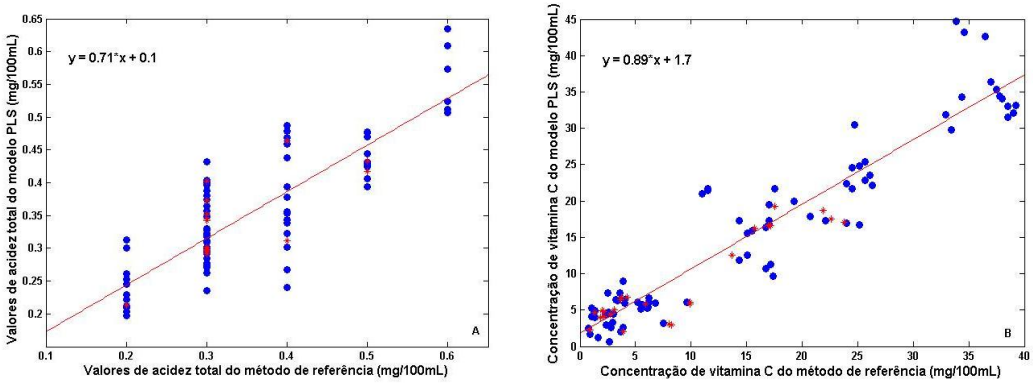

Figura 5: Ajuste dos modelos PLS. (A) Acidez total. (B) Vitamina C. (•) amostras de calibração. (*) amostras de validação.

Fonte: Acervo do autor (2014)

Para o ajuste apresentado na Figura 5, os coeficientes de correlação foram 0,7119 e 0,8890 para acidez total e vitamina C, respectivamente. Apesar de parecerem valores baixos, resultados de coeficiente de correlação dessa ordem já foram reportados antes na literatura quando o método de referência, na calibração multivariada, é a titulação (VALDERRAMA; BRAGA; POPPI 2007, VALDERRAMA; BRAGA; POPPI 2007b, FERREIRA; PALLONE; POPPI, 2013).

A linearidade do modelo de calibração multivariada de primeira ordem foi avaliada através do gráfico dos resíduos da calibração e validação, mostrados na Figura 6. Qualitativamente, este tipo de gráfico pode indicar se os dados seguem ou não um comportamento linear, se a distribuição dos resíduos for aleatória.

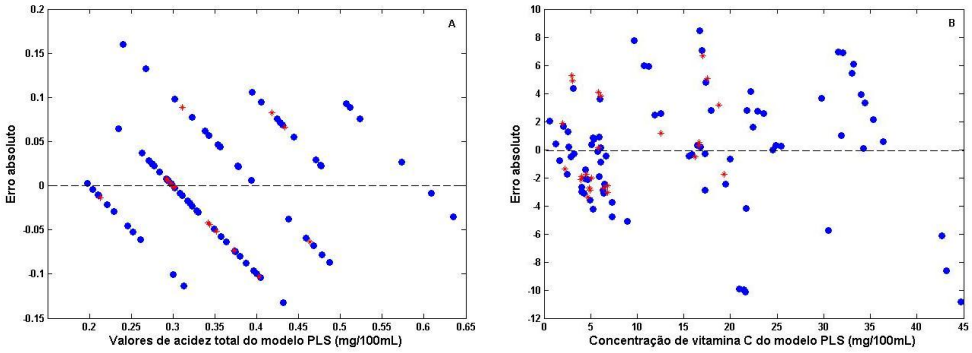

Figura 6: Resíduos do modelo PLS. (A) Acidez total. (B) Vitamina C. (•) amostras de calibração. (*) amostras de validação.

Fonte: Acervo do autor (2014)

Através dos resíduos observa-se o comportamento aleatório para os modelos construídos para determinação de acidez total e vitamina C sugerindo que os dados se adequam a um modelo linear.

A sensibilidade e sensibilidade analítica apresentaram resultados adequados na determinação de vitamina C e acidez total, considerando-se a faixa de concentração utilizada no trabalho, 0,74 – 39,25 e 0,20 – 0,60 mg/100mL, para vitamina C e acidez, respectivamente. O inverso da sensibilidade analítica permite estabelecer a menor diferença de concentração entre as amostras que pode ser distinguida pelo método. Sendo assim, é possível fazer a distinção de amostras com diferença de concentração de vitamina C da ordem de 0,33 mg/100mL e amostras com diferença de acidez da ordem de $5{,}7 \times 10^{-3}$ mg/100mL.

Os resultados de RPD podem ser considerados satisfatórios, ou seja, entre 2,4 e 1,5 (BOTELHO; MENDES; SENA, 2013). Os limites de detecção e quantificação do modelo multivariado para determinação de acidez e vitamina C apresentaram resultados coerentes com as quantidades medidas desses parâmetros.

Os resultados mostraram a possibilidade de utilização de um mesmo modelo PLS na avaliação do teor de acidez total e vitamina C em diferentes produtos (néctar de diferentes sabores) simultaneamente. Assim, a espectroscopia UV aliada ao método de calibração multivariada PLS permite a construção de modelos de calibração multiproduto. O método proposto é rápido, não emprega reagentes/solventes e não gera resíduos tóxicos, sendo uma alternativa aos métodos convencionais baseados em titulação.

4.2 RESULTADO DA APLICAÇÃO 2: DETERMINAÇÃO DE ACIDEZ TOTAL E VITAMINA C EM BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA EMPREGRANDO ESPECTROSCOPIA UV.

Para determinar o teor de vitamina C e acidez total por espectroscopia na região do UV os espectros, vetores de respostas instrumentais para cada amostra, foram organizados no formato de uma matriz. O método de mínimos quadrados parciais (PLS) foi empregado no desenvolvimento dos modelos de calibração multivariada multiproduto. A Figura 7 apresenta os espectros na região do UV obtidos para todas as amostras analisadas. Nessa aplicação foi necessário a aplicação da primeira derivada para o pré-processamento dos espectros. Isso pode ter ocorrido devido a cor opaca da amostra de sucos industrializados a base de soja, mesmo após ter passado por processo de diluição.

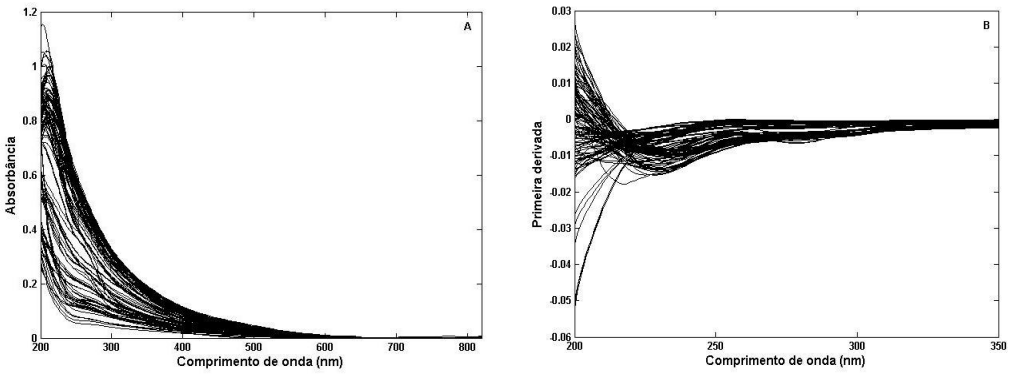

Figura 7: Espectros UV das amostras de bebidas a base de soja. (A) Espectros sem pré-processamento. (B) Espectro após primeira derivada.
Fonte: Acervo do autor (2014)

As amostras foram separadas em um conjunto de calibração e validação através do algoritmo de Kennard-Stone, resultando em 75% das amostras no conjunto de calibração e 25 das amostras no conjunto de validação (KENNARD; STONE, 1969).

O próximo passo consistiu na análise para identificação de amostras anômalas que, no conjunto de calibração, as amostras consideradas anômalas com

base nos testes de *leverage*, resíduos espectrais não modelados e resíduos na variável dependente foram eliminadas e o modelo de calibração novamente reconstruído. Após a otimização dos modelos de calibração, os testes para identificação de amostras anômalas foram aplicados a cada conjunto de validação e os *outliers* para esses conjuntos também foram eliminados.

A otimização dos conjuntos de calibração e validação pela eliminação das amostras anômalas resultaram em 80 e 76 amostras de calibração e 21 e 23 amostras de validação para os modelos PLS na determinação de acidez total e vitamina C, respectivamente.

Os modelos foram centrados na média, construídos com 10 variáveis latentes na determinação de ambos os parâmetros, acidez total e vitamina C. O número de variáveis latentes foi determinado através dos resultados da raiz quadrada do erro médio quadrático de validação cruzada (RMSECV) para as amostras de calibração, obtido por validação cruzada em blocos contínuos de nove amostras. A percentagem de variância acumulada no bloco **y** também foi levada em consideração para a escolha do número de variáveis latentes. Nesse caso, em ambos os casos, a percentagem de variância acumulada mínima para os modelos otimizados foi de 70%.

A exatidão dos modelos desenvolvidos foi avaliada de acordo com os valores do erro médio quadrático de calibração (RMSEC) e previsão (RMSEP), apresentados na Tabela 3, para os modelos após otimização.

Tabela 3: Figuras de mérito do modelo multivariado para bebidas industrializadas a base de soja.

Figuras de mérito		Vitamina C	Acidez
Exatidão	RMSEC[a]	2,5332	0,0218
	RMSEP[a]	1,6973	0,0199
Coeficiente de correlação		0,7188	0,7435
RPD_{cal}		1,3	1,5
RPD_{val}		2,6	1,7
Sensibilidade analítica^{-1} [a]		0,7356	0,0061
Limite de detecção[a]		2,4275	0,0200
Limite de quantificação[a]		7,3561	0,0606

[a] (mg/100mL)

Fonte: Acervo do autor (2014)

Os resultados de RMSEC e RMSEP indicam que o número de variáveis latentes utilizados em cada um dos modelos foram escolhidos corretamente, não havendo sobre ajuste ou falta de ajuste. Entretanto, os valores de RMSEP e RMSEC são parâmetros globais que incorporam erros sistemáticos e aleatórios. Sendo assim, outros indicadores de exatidão são a regressão entre os valores de referência e os valores estimados pelo modelo (ajuste), através dos valores da inclinação, do intercepto e coeficiente de correlação, incluindo a região da elipse de confiança (RIU; RIUS, 1996; VALDERRAMA; BRAGA; POPPI, 2009) apresentada na Figura 8.

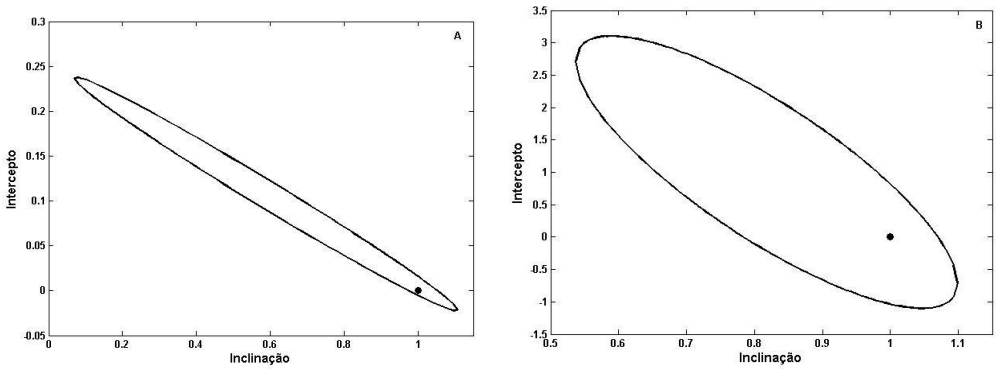

Figura 8: Região da elipse da confiança (99%). (A) Acidez Total. (B) Vitamina C (•) Ponto onde o intercepto é igual a zero e a inclinação igual a um.

Fonte: Acervo do autor (2014)

Através da Figura 8 observa-se a região da elipse de confiança para o intercepto e a inclinação da regressão entre os valores de acidez total e vitamina C previstos pelo modelo PLS e os valores de acidez total e vitamina C determinados por titulação potenciométrica e titulação de oxidação-redução, respectivamente. Pode-se verificar que, a região da elipse de confiança contém o ponto ideal (1,0) para a inclinação e o intercepto, respectivamente. Isto mostra que os valores de acidez total e os valores de vitamina C determinados por titulação e os valores de acidez total e vitamina C determinados pelo modelo PLS não apresentam diferença significativa com 99% de confiança.

Para o ajuste entre os valores de acidez total obtidos por potenciometria e os valores de acidez total obtidos pelo modelo PLS, o coeficiente de correlação foi 0,7435. Para a vitamina C, o ajuste apresentou coeficiente de correlação 0,7188. Apesar dos resultados para o coeficiente de correlação estarem abaixo de 0,9 resultados da ordem de 0,7 já foram reportados antes na literatura quando o método de referência, na calibração multivariada, é a titulação (VALDERRAMA; BRAGA; POPPI 2007, VALDERRAMA; BRAGA; POPPI 2007b, FERREIRA; PALLONE; POPPI, 2013).

A linearidade do modelo de calibração multivariada de primeira ordem foi avaliada através do gráfico dos resíduos da calibração e validação, mostrados na Figura 9. Através dos resíduos observa-se o comportamento aleatório para os modelos construídos para determinação de acidez total e vitamina C indicando que o modelo PLS segue um comportamento linear.

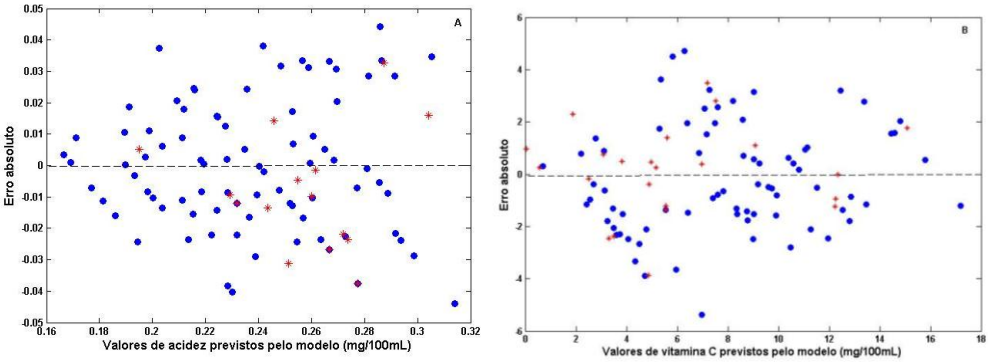

Figura 9: Resíduos do modelo PLS. (A) Acidez Total. (B) Vitamina C. (●) amostras de calibração. (*) amostras de validação.

Fonte: Acervo do autor (2014)

O valor de RPD do modelo de calibração para determinação de vitamina C mostrou resultado muito próximo do que é considerado satisfatório, podendo ser considerado adequado em relação ao valor de RPD para a validação desse parâmetro. No modelo para determinar acidez o valor de RPD pode ser considerado satisfatório para ambos os casos, calibração e validação. De acordo com a literatura (BOTELHO; MENDES; SENA, 2013) são considerados bons modelos de calibração os modelos que apresentam valores para RPD acima de 2,4. Os modelos com valores de RPD entre 2,4 e 1,5 são considerados satisfatórios.

A sensibilidade e sensibilidade analítica apresentaram resultados adequados na determinação de vitamina C e acidez, considerando-se a faixa de concentração utilizada no trabalho, 0,83 – 16,83 e 0,17 – 0,34 mg/100mL, para vitamina C e acidez, respectivamente. O inverso da sensibilidade analítica permite estabelecer a menor diferença de concentração entre as amostras que pode ser distinguida pelo método. Sendo assim, é possível fazer a distinção de amostras com diferença de concentração de vitamina C da ordem de 0,73 mg/100mL e amostras com diferença de acidez da ordem de $6{,}1 \times 10^{-3}$ mg/100mL.

Os limites de detecção e quantificação do modelo multivariado para determinação de acidez apresentaram resultados coerentes com as quantidades medidas da acidez. No modelo para determinação de vitamina C os resultados alcançados foram menos satisfatórios indicando que o modelo proposto não consegue detectar e quantificar amostras com concentração de vitamina C abaixo de 2,43 e 7,36 mg/100mL, respectivamente.

Os resultados mostram a possibilidade de utilização de um modelo PLS na avaliação do teor de acidez total e vitamina C em diferentes produtos (bebidas a base de soja de diferentes sabores, bebidas *light* a base de soja) simultaneamente. Assim, a espectroscopia UV aliada ao método de calibração multivariada PLS permite a construção de modelos de calibração multiproduto. O método proposto é rápido, não emprega reagentes/solventes e não gera resíduos tóxicos, sendo uma alternativa aos métodos convencionais baseados em titulação.

4.3 RESULTADO DA APLICAÇÃO 3: DETERMINAÇÃO DE VITAMINA C E ACIDEZ TOTAL EM NÉCTARES DE FRUTA E BEBIDAS INDUSTRIALIZADAS A BASE DE SOJA EMPREGANDO ESPECTROSCOPIA VIS-NIR.

Nesta aplicação, buscou-se desenvolver um modelo de calibração multivariada multiproduto para determinar vitamina C e acidez total em néctar de frutas e bebidas industrializadas a base de soja, concomitantemente, a partir da espectroscopia Vis-NIR e calibração multivariada de primeira ordem através de mínimos quadrados parciais multicomponente (PLS2). Acredita-se que o teor de acidez total apresenta alguma relação com a concentração de vitamina C, o que justificaria o emprego do método PLS2 nessa aplicação. A Figura 10 mostra os espectros obtidos na região Vis-NIR para as amostras de néctar de frutas e bebidas a base de soja. Para a construção do modelo, testou-se algumas estratégias de pré-processamento, tais como correção de linha de base e aplicação de primeira derivada (janela de 15 variáveis e polinômio de segunda ordem).

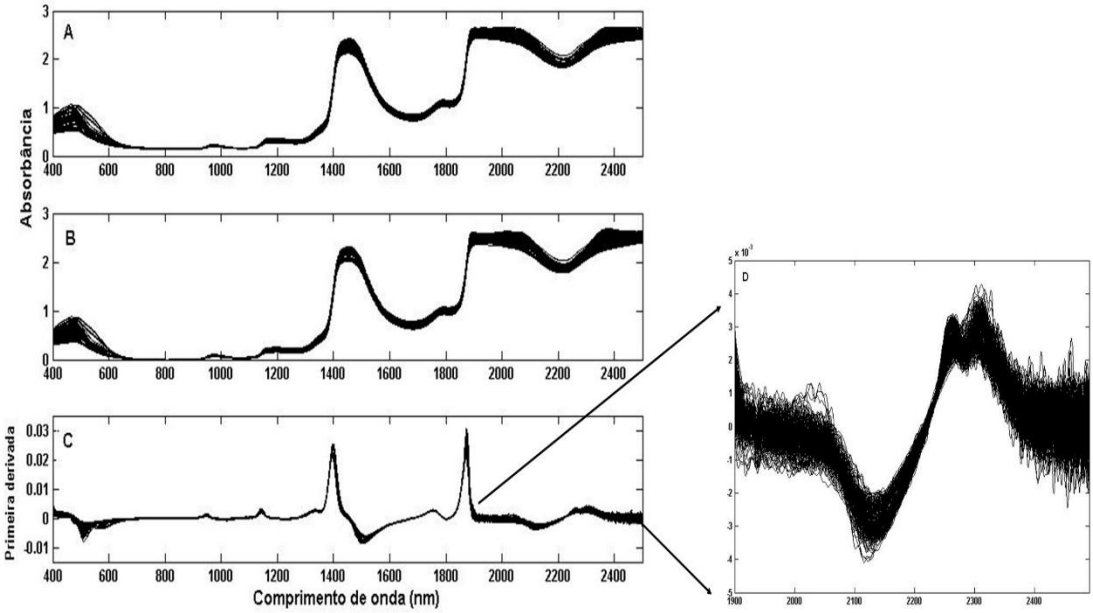

Figura 10: Espectros Vis-NIR das amostras de néctar de frutas e bebidas industrializadas a base de soja. (A) Espectros brutos; (B) Espectros após correção de linha de base; (C) Espectros após primeira derivada; (D) Expansão da região entre 1900 - 2500nm do espectro Vis-NIR após a primeira derivada.

Fonte: Acervo do autor (2015)

Um total de 240 amostras foram utilizadas nessa aplicação, sendo 123 amostras de néctar de frutas e 117 amostras de bebidas industrializadas a base de soja. A separação das amostras entre os conjuntos de calibração e validação, realizado pelo algoritmo de Kennard-Stone (KENNARD; STONE, 1969), resultou em um total de 180 amostras para o conjunto de calibração e 60 amostras para o conjunto de validação.

A performance dos modelos desenvolvidos foi testada a partir dos valores da raiz quadrada do erro médio quadrático de calibração (RMSEC) e previsão (RMSEP), e dos valores do coeficiente de correlação do ajuste entre os valores de referência e os valores previstos pelo modelo. Estes resultados são apresentados na Tabela 4. Modelos de calibração, a partir do método PLS2, foram construídos para as amostras após aplicação dos pré-processamentos baseline e primeira derivada. Testou-se também a possibilidade de construir modelos apenas com a região espectral do infravermelho próximo (780 – 2500nm). Entretanto, verifica-se que com aplicação de primeira derivada ocorre a intensificação de ruídos na região entre 1900 - 2500nm (Figura 10 – D), assim, quando os espectros passaram pelo pré-processamento de primeira derivada foram construídos modelos com as regiões Vis-NIR entre 400 – 1900 nm e NIR entre 780 - 1900 nm.

Tabela 4: Performance dos modelos PLS.

Modelo	RMSEC Vitamina C/Acidez (mg/100mL)	RMSEP Vitamina C/Acidez (mg/100mL)	R Vitamina C/Acidez
$Vis\text{-}NIR_{baseline}$	8,8464 / 0,0783	7,8456 / 0,0751	0,68 / 0,62
$NIR_{baseline}$	9,2360 / 0,0882	10,5557 / 0,0712	0,65 / 0,61
$Vis\text{-}NIR_{1derivada}$	7,7514 / 0,0795	8,8305 / 0,0785	0,75 / 0,57
$NIR_{1derivada}$	10,0090 / 0,0872	8,0705 / 0,0804	0,72 / 0,49

Fonte: Acervo do autor (2014)

A partir dos resultados da Tabela 4, verifica-se que o modelo construído a partir da região Vis-NIR, após correção de linha de base, alcançou performance mais satisfatória. Analisando o gráfico dos coeficientes de regressão dos modelos (Figura

11), verifica-se que realmente a região do visível apresenta importância na regressão. Esta importância pode estar relacionada com a vitamina C que, segundo Choi et al. (2002), apresenta influência na coloração de sucos de frutas, que por sua vez, absorvem na região do visível.

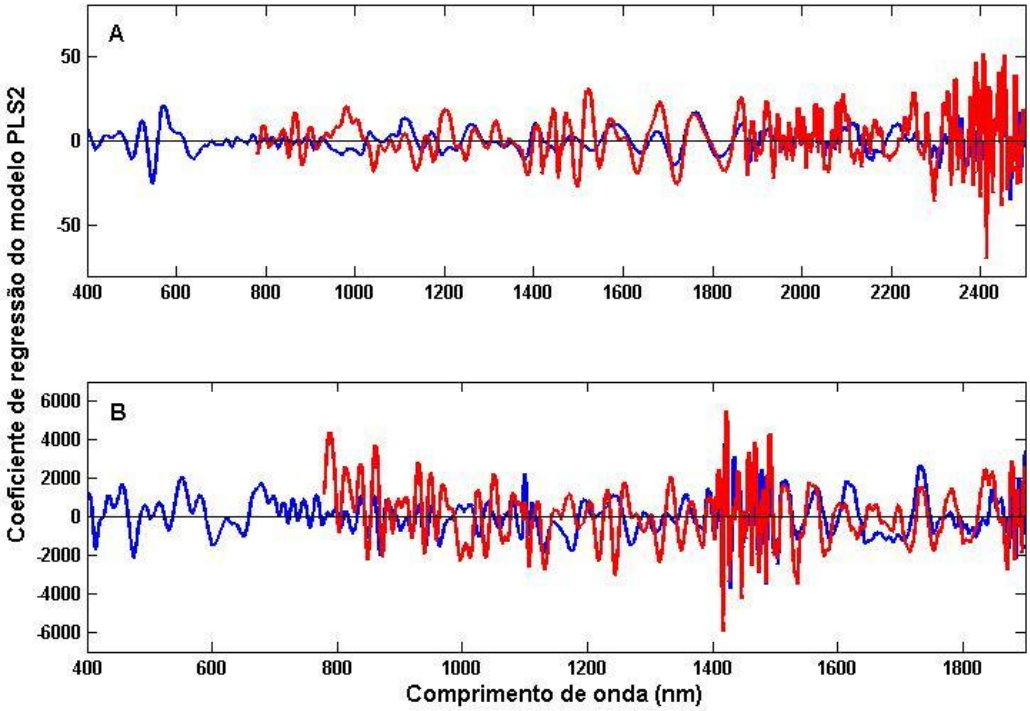

Figura 11: Coeficientes de regressão dos modelos PLS2. (A) Espectros após correção de linha de base; (B) Espectros após primeira derivada. (—) Região Vis-NIR; (—) Região NIR.

Fonte: Acervo do autor (2015)

O próximo passo consistiu na análise para identificação de amostras anômalas, no modelo construído com a região espectral Vis-NIR com correção de linha de base. Para tanto, avaliou-se o gráfico do *leverage* versus resíduos Q do modelo Vis-NIR, após correção de linha de base. Esses *outliers* foram removidos e um segundo modelo foi construído. Os *outliers* do segundo modelo foram avaliados da mesma forma e eliminados e um terceiro modelo foi construído. Esse terceiro modelo também teve seus *outliers* avaliados e eliminados de acordo com o gráfico do *leverage* versus resíduos e um quarto modelo foi construído. Assim, foram realizadas três etapas de eliminação de *outliers* na fase de calibração, resultando no

quarto modelo que foi considerado otimizado. Quando a remoção de *outliers* é realizada para um dado modelo e este, após sua reconstrução, continua apresentando amostras anômalas significa que o espaço das variáveis do modelo pode apresentar uma distribuição de "*clusterings*" (agrupamentos) não usual. Assim, quando isso acontece, o teste para identificação de *outliers* pode ser relaxado (ASTM E1655-05), ou seja, em modelos de calibração de primeira ordem os *outliers* podem ser removidos e o modelo reconstruído mais de uma vez. Alternativamente, um método não linear, como as redes neurais, poderia ser empregado e será uma perspectiva futura desse trabalho.

Após a otimização dos modelos de calibração, a identificação de amostras anômalas foi realizada, da mesma forma, para o conjunto de validação e os *outliers* foram eliminados em uma única etapa. A Tabela 5 mostra os resultados para a avaliação de *outliers* bem como, os valores de RMSEC e RMSEP para os modelos à medida que os *outliers* foram eliminados (Mod1= primeiro modelo; Mod2= segundo modelo; Mod3= terceiro modelo; Mod4= quarto modelo). Foram avaliados os *outliers* até o terceiro modelo de calibração e as amostras anômalas do conjunto de validação foram avaliadas com base no quarto modelo de calibração, tido como otimizado. O resultado da avaliação de *outliers* pode ser visualizado também a partir da Figura 12.

Tabela 5: Resultados para identificação de outliers.

Modelos	% variância acumulada explicada em y	Nº amostras	Nº outliers baseado no leverage	Nº outliers baseado no resíduo	Nº total de outliers descartados	RMSEC (mg/100mL) Vit. C Acidez	RMSEP (mg/100mL) Vit. C Acidez
Mod1 (20VL)	68,15	180	14	23	30	8,8464 0,0783	7,8456 0,0751
Mod2 (19VL)	71,77	150	6	12	17	7,5486 0,0775	10,2095 0,0742
Mod3 (18VL)	76,04	133	4	10	13	6,8990 0,0782	12,3076 0,0850
Mod4$_{otimizado}$ (18VL)	81,17	120	0	0	0	9,8771 0,0826	6,3230 0,0769
Validação	-	60	6	13	41	5,0909 0,0758	6,3230 0,0769

Fonte: Acervo do autor (2015)

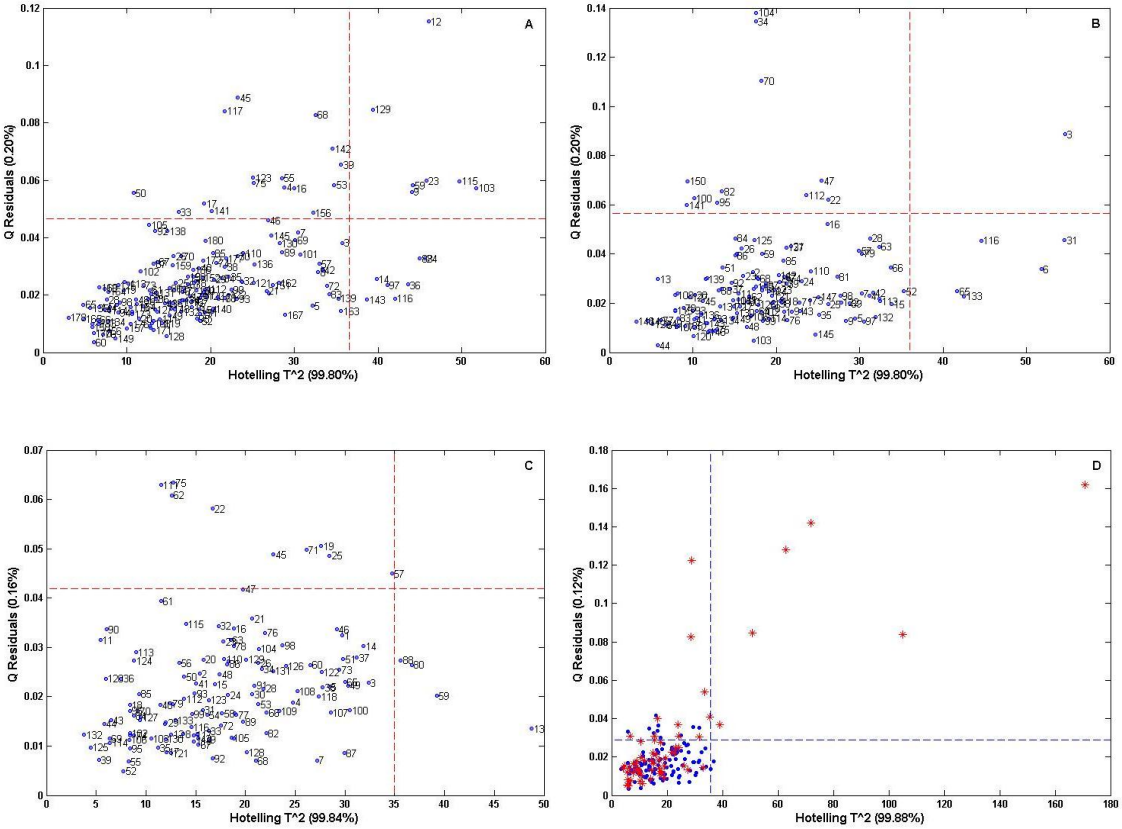

Figura 12: Amostras anômalas para o modelo PLS multicomponente. (●) amostras de calibração; (*) amostras de validação. (A) Modelo 1; (B) Modelo 2; (C) Modelo 3; (D) Modelo 4. Fonte: Acervo do autor (2015)

Os modelos foram centrados na média e o número de variáveis latentes (VL) foi determinado através dos resultados da raiz quadrada do erro quadrático médio de validação cruzada (RMSECV) para as amostras de calibração, obtido por validação cruzada em blocos contínuos de 9 amostras, e levando-se em consideração também a percentagem de variância explicada na matriz **Y**.

Os resultados das figuras de mérito são mostrados na Tabela 6. Os valores de exatidão são representados pelos valores de RMSEC e RMSEP para os modelos após otimização. Entretanto, como o RMSEC e RMSEP são considerados parâmetros globais que incorporam erros sistemáticos e aleatórios, outros indicadores de exatidão são o ajuste (Figura 13) dos valores de referência versus os

valores previstos pelo modelo, representados pelo coeficiente de correlação, e a consideração da elipse de confiança (Figura 14).

Tabela 6: Figuras de mérito para o modelo multivariado multicomponente.

Figuras de mérito		Vitamina C	Acidez
Exatidão	RMSEC[a]	6,3230	0,0769
	RMSEP[a]	5,0909	0,0758
Coeficiente de correlação		0,8117	0,7103
RPD_{cal}		1,1	1,4
RPD_{val}		1,8	1,7
Sensibilidade analítica^{-1} [a]		0,9469	0,0021
Limite de detecção[a]		3,1249	0,0069
Limite de quantificação[a]		9,4692	0,0209

[a] (mg/100mL)

Fonte: Acervo do autor (2015)

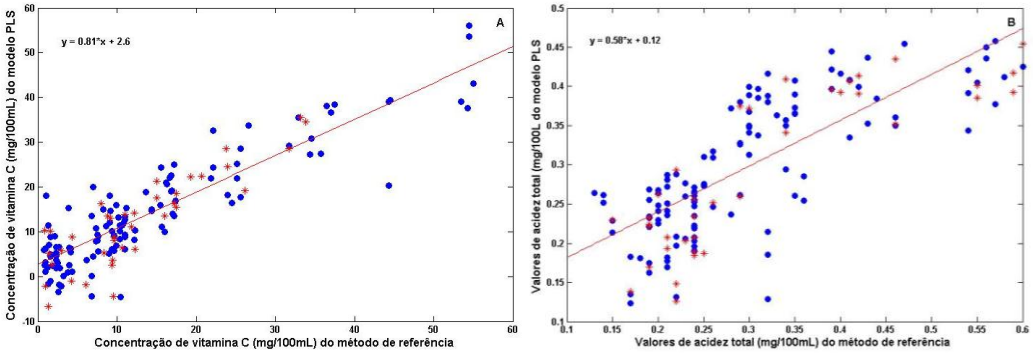

Figura 13: Ajuste. (A) Vitamina C; (B) Acidez.

(●) amostras de calibração; (*) amostras de validação.

Fonte: Acervo do autor (2015)

Para o ajuste apresentado na Figura 13, os coeficientes de correlação foram 0,8117 e 0,7103 para vitamina C e acidez total, respectivamente. Esses resultados encontram-se de acordo com os resultados de coeficiente de correlação para o ajuste de modelos de calibração multivariada, cujo método de referência tenha sido

a titulação (VALDERRAMA; BRAGA; POPPI, 2007; VALDERRAMA; BRAGA; POPPI, 2007b; FERREIRA; PALLONE; POPPI, 2013).

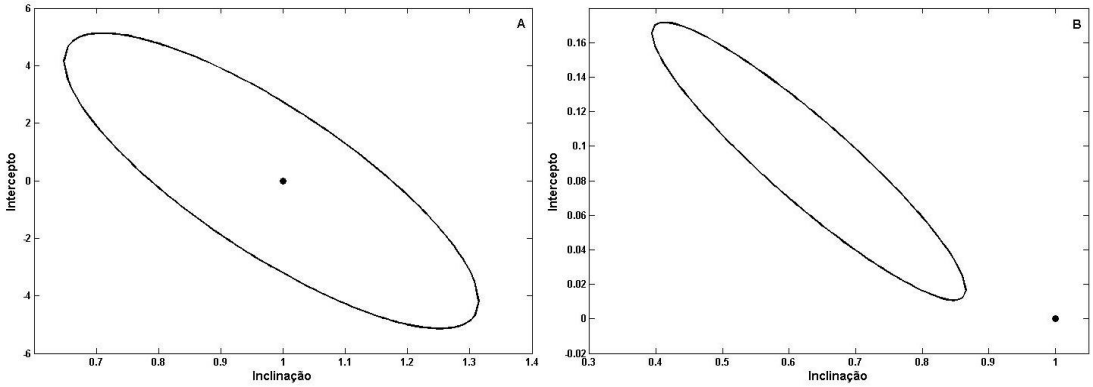

Figura 14: Elipse de Confiança. (A) Vitamina C; (B) Acidez.
Fonte: Acervo do autor (2015)

Como se pode observar na Figura 14, a região da elipse de confiança contém o ponto ideal (1,0), para a inclinação e o intercepto, respectivamente, no modelo construído para determinação da vitamina C. Este resultado mostra que a concentração de vitamina C do método de referência e os valores previstos pelo modelo não apresentam diferença significativa com 99% de confiança. O mesmo, não pode ser observado para o modelo construído para determinação da acidez. Entretanto, o valor de RPD para o modelo de calibração para determinação de acidez mostrou resultado muito próximo do que é considerado satisfatório, podendo ser considerado satisfatório em relação ao valor de RPD para a validação desse parâmetro. No modelo para determinar vitamina C o valor de RPD somente pode ser considerado satisfatório para a validação, isso porque de acordo com a literatura (BOTELHO; MENDES; SENA, 2013) valores para RPD entre 2,4 e 1,5 são considerados satisfatórios.

Para avaliar os resultados da previsão do modelo otimizado, um teste t pareado no nível de confiança de 95% (MILLER; MILLER, 2000) foi realizado. O valor para t calculado na determinação de vitamina C foi de 0,644 enquanto que para determinação do teor de acidez total foi de 1,483 contra um valor para t tabelado de 2,042. Como os valores de t calculado foram menores que o valor de t

tabelado, isso indica que não existem diferenças significativas entre os métodos de referência para a determinação de vitamina C e acidez total e a calibração multiproduto multicomponente.

A sensibilidade e sensibilidade analítica apresentaram resultados adequados na determinação de vitamina C e acidez, considerando-se a faixa de concentração utilizada no trabalho, 0,74 – 55,00 e 0,13 – 0,60 mg/100mL, para vitamina C e acidez, respectivamente. O inverso da sensibilidade analítica permite estabelecer a menor diferença de concentração entre as amostras que pode ser distinguida pelo método. Sendo assim, é possível fazer a distinção de amostras com diferença de concentração de vitamina C da ordem de 0,95 mg/100mL e amostras com diferença de acidez da ordem de $2{,}1 \times 10^{-3}$ mg/100mL.

Os limites de detecção e quantificação do modelo multivariado para determinação de acidez apresentaram resultados coerentes com as quantidades medidas da acidez. No modelo para determinação de vitamina C os resultados alcançados foram menos satisfatórios e o modelo proposto não consegue detectar e quantificar amostras com concentração de vitamina C abaixo de 3,12 e 9,47 mg/100mL, respectivamente.

A linearidade dos modelos construídos foi avaliada através do gráfico dos resíduos de calibração e validação, mostrados na Figura 15. Qualitativamente, pode-se sugerir que os dados se ajustam a um modelo linear, pois a distribuição dos resíduos apresentou comportamento aleatório.

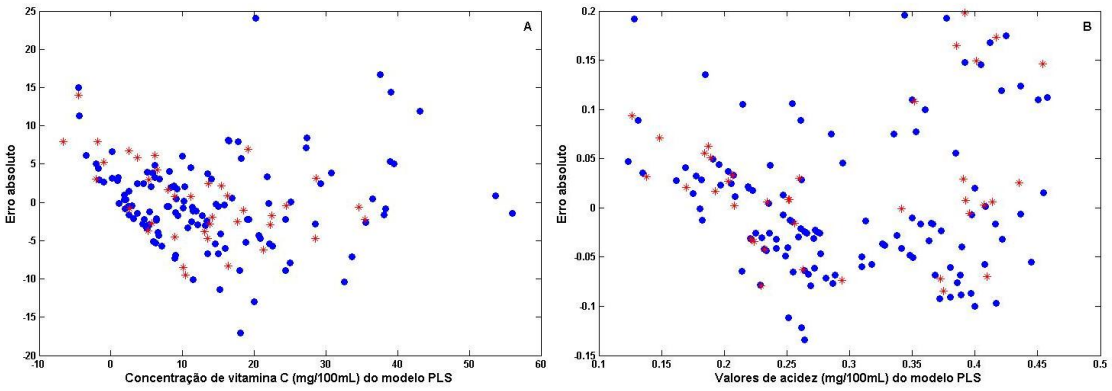

Figura 15: Resíduos dos modelos PLS. (A) Vitamina C; (B) Acidez.

(•) amostras de calibração. (*) amostras de validação.

Fonte: Acervo do autor (2015)

Os resultados sugerem a possibilidade da utilização de um mesmo modelo PLS na avaliação simultânea do teor de acidez total e vitamina C em diferentes produtos (néctar de diferentes sabores, néctar *light*, bebidas industrializadas a base de soja e bebidas industrializadas a base de soja *light*) concomitantemente. O modelo construído a partir da região Vis-NIR mostrou maior capacidade preditiva em relação ao modelo construído a partir da região NIR. Assim, a espectroscopia Vis-NIR aliada ao método de calibração multivariada PLS permite a construção de modelos de calibração multiproduto. O método proposto é rápido, não emprega reagentes/solventes e não gera resíduos tóxicos, sendo uma alternativa aos métodos convencionais baseados em titulação.

5 CONCLUSÂO

Modelos de calibração multivariada multiproduto podem ser construídos com base no método linear PLS. Os resultados mostraram a possibilidade de utilização de um mesmo modelo PLS na avaliação do teor de acidez total e de vitamina C em diferentes produtos (néctar de diferentes sabores juntamente com néctar *light*, bebidas industrializadas a base de soja em conjunto com bebidas a base de soja *light*). Assim, a espectroscopia UV aliada ao método de calibração multivariada PLS permite a construção de modelos de calibração multiproduto.

A espectroscopia Vis-NIR e método quimiométrico PLS permitem a construção de modelos de calibração multivariada multiproduto multicomponente, ou seja, empregando todo o espectro Vis-NIR para avaliar simultaneamente a acidez total e a vitamina C em néctar de frutas e bebidas industrializadas a base de soja de diferentes sabores, além de néctar de frutas e bebidas industrializadas a base de soja *light* de diferentes sabores. A espectroscopia Vis-NIR mostrou-se vantajosa em relação à espectroscopia UV pois permitiu análises simultâneas e determinações de diferentes parâmetros ao mesmo tempo. Entretanto, os resultados dos modelos construídos a partir da espectroscopia UV mostraram resultados superiores para a exatidão.

As metodologias propostas são rápidas, não empregam reagentes/solventes e não geram resíduos tóxicos, sendo uma alternativa aos métodos convencionais baseado em titulação potenciométrica e titulação por oxidação-redução.

Visto que o modelo de calibração multivariada multiproduto, validado a partir de parâmetros como figuras de mérito, apresenta muitos pontos positivos quando comparado à metodologia de referência, além de apresentar um grande potencial em ser um possível substituto dos métodos de titulação para a determinação de propriedades de interesse, no que concerne a acidez total e teor de vitamina C, também tem a capacidade de poder ser empregado futuramente em escala industrial.

6 PERSPECTIVA

Como perspectivas futuras, pretende-se:

- Desenvolver modelos de calibração multivariada não linear para determinar o teor de acidez total e vitamina C em néctares de frutas e bebidas industrializadas a base de soja.

- Avaliar o potencial de imagens digitais na construção de modelos de calibração multivariada visando a determinação do teor de acidez total e vitamina C em néctares de frutas e bebidas industrializadas a base de soja.

7 REFERÊNCIAS

ASTM, **Annual Book of ASTM Standards**. Standards practices for infrared, multivariate, quantitative analysis, E1655-05. Pennsylvania-USA: West Conshohocken, 2005.

BERZAGHI, P.; SHENK, J.S.; WESTERHAUS, M.O. LOCAL prediction with nearinfrared multi-product databases. **Journal of Near Infrared Spectroscopy**, v.8, p. 1-9, 2000.

BOOKSH, K. S.; KOWALSKI, B. R. Theory in analytical chemistry. **Analytical Chemistry**, v.66, n.15, p. 782A-791A, 1994.

BOTELHO, B. G.; MENDES, B. A. P.; SENA, M. M. Implementação de um método robusto para o controle fiscal de umidade em queijo minas artesanal. Abordagem metrológica multivariada. **Química Nova,** v.36, n.9, p. 1416-1422, 2013.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento – Instrução normativa n. 12 de 04 de setembro de 2003. **Diário Oficial da República Federativa do Brasil**, Brasília-DF, seção 1, p. 2, 09 de set. 2003.

BRERETON, R.G. Introduction to multivariate calibration in analytical chemistry. **Analyst**, v.125, p. 2125-2154, 2000.

CAMPS, C.; ROBIC, R.; LAURENS, F. Rapid determination of soluble solids content and acidity of Black currant (*Ribes nigrum L.*) juice by mid-infrared spectroscopy performed in series. **Food Science and Technology**. v.43, n.7, p. 1164-1167, 2010

CATANIA, A. S.; BARROS, C. R.; FERREIRA, S. R. G.. Vitaminas e minerais com propriedades antioxidantes e risco cardiometabólico: Controvérsias e perspectivas. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v.53, n.5, p. 550-559, 2009.

CHOI, M. H.; KIM, G. H.; LEE, H. S. Effects of ascorbic acid retention on juice color and pigment stability in blood orange (*Citrus sinesis*) juice during refrigerated storage. **Food Research International**, v.35, n.8, p. 753-759, 2002.

COSTA FILHO, P.A.; POPPI, R.J.; Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio. Determinação simultânea de glicose, maltose e frutose. **Química Nova**, v.25, n.1, p. 46-52, 2002.

DANIELLI, F.; COSTA, G. L. L. R.; SILVA. L. C.; HARA, S. A. S.; SILVA, A. A.; Determinação de vitamina C em amostras de suco de laranja in natura e amostras comerciais de suco de laranja pasteurizado e envasado em embalagem Tetra Park. **Revista do Instituto de Ciências da Saúde**, v.27, n.4, p. 361-365, 2009.

ESCANDAR, G. M.; GOICOECHEA, H. C.; MUÑOZ DE LA PEÑA, A.; OLIVERI, A. C. Second-end higher-order data generation and calibration: A tutorial. **Analytica Chimica Acta**, v. 806, p. 8-26, 2014.

FEDERATION INTERNATIONAL DES PRODUCTEURS DE JUS DE FRUITS. **Determination of titrable acidity** – IFUMA03 – França, 2005.

FERREIRA, D.S.; PALLONE, J.A.L.; POPPI, R.J. Fourier transform near-infrared spectroscopy (FT-NIRS) application to estimate Brazilian soybean [*Glycine max* (L.) Merril] composition. **Food Research International**, v.51, n.1, p. 53-58, 2013.

GELADI, P.; KOWALSKI, B. R. Partial Least-Squares Regression - A Tutorial. **Analytica Chimica Acta**, v.185, p. 1-17, 1986.

GRUDPAN, K.; SRITHARATHIKHUN, P.; JAKMUNEE, J. Flow injection conductimetric or spectrophotometric analysis for acidity in fruit juice. **Analytica Chimica Acta**, v.363, n.2-3, p. 199-202, 1998.

GUINÈ, R.; Projeto de uma indústria de processamentos de Néctar de Maça e de Pera. **Millenium**, v.42, p. 175-189, 2012.

INSTITUTO ADOLFO LUTZ. **Métodos físico-químicos para análise de alimentos**. São Paulo: Instituto Adolfo Lutz, 2008.

JONG, S. SIMPLS: An alternative approach to partial least squares regression. **Chemometrics and Intelligent Laboratory Systems,** v.18, n.3, p. 251-263, 1993.

KENNARD, R.W.; STONE, L.A. Computer aided desing of experiments. **Technometrics**, v.11, n.1, p. 137-148, 1969.

LAVARDA, L. **Determinação da cinética de degradação térmica da vitamina C em polpa de acerola via aquecimento ôhmico**. 2011. 44 p. Trabalho de Conclusão de Curso – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.

LIMA, K.M.G.; TREVISAN, M.G.; POPPI, R.J.; DE ANDRADE, J.C.; 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, K. P.; FUKURO, I. Y. Compostos polifenólicos presentes em alimentos e sua utilização no combate de radicais livres. **Jornal Eletrônico de Química**. v.1, p.4-5, III edição. Campo Mourão – PR, 2012.

MAGGIO, R. M.; MUÑOZ DE LA PEÑA, A.; OLIVERI, A. C. Unfolded partial least-squares whith residual quadrilenearization: A new multivariate algotithm for processing five-way data achieving the second-order advantage. Apllication to fourth-order excitacion-emission-kinetic-pH fluorescence analytical data. **Chemometrics and Intelligent Laboratory Systems**, v.109, p. 178-185, 2011.

MARTENS, H.; NAES, T.; Multivariate Calibration. New York: Wiley, 1996.

MERCALDI, J. C.; **Desenvolvimento de bebidas a base de “leite” de soja acrescedida de suco de graviola.** Dissertação de Mestrado. USP: Araraquara-SP, 2006.

MICKLANDER, E.; KJELDAHL, K.; EGEBO, M.; N∅RGAARD, L. Multi-product calibration models of near infrared spectra of foods. **Journal of Near Infrared Spectroscopy**, v.14, p. 395-402, 2006.

MILLER, J.N., MILLER, J. C. **Statistic and chemometrics for analytical chemistry**. 4ª ed, Pentrice Hall: United Kingdom, 2000.

NAES, T.; ISAKSSON, T. Locally weighted regression in diffuse near-infrared transmittance spectroscopy. **Applied Spectroscopy**, v.46, n.1, p. 34-43, 1992.

NAGAOKA, S.; KAKIUCHI, T.; OHARA, K.; MUKAI, K. Kinetics of the reaction by which natural vitamin E is regenerated by vitamin C. **Chemistry and Physics of Lipids**, v.146, n.1, p. 26-32, 2007.

OTTO, M. **Chemometrics, Statistics and Computer Application in Analytical Chemistry**, Wiley-VCH: Weinheim, 1999.

PIRILLO, C. P.; SABIO, R. P. 100% Suco: Nem tudo é suco nas bebidas de frutas. **Revista de Hortifruti Brasil**, v.81, p. 6-12, 2009.

REBOUCHE, C. J. Ascorbic acid and carnitine biosynthesis. **The American Journal of Clinical Nutrition**, v.54, n.6, p.1147S-1152S, 1991.

RIBANI, M.; BOTTOLI, C. B. G.; COLLINS, C. H.; JARDIM, I. C. S. F.; MELO, L. F. C. Validação em métodos cromatográficos e eletroforéticos. **Química Nova**, v.27, n.5, p. 771-780, 2004.

RIU, J.; RIUS, F.X. Assessing the accuracy of analytical methods using linear regression with errors in both axes. **Analytical Chemistry**, v.68, p. 1851-1857, 1996.

SANTOS, P. H. da S. **Estudo da cinética de degradação do ácido ascórbico na secagem de abacaxi em atmosfera modificada**. Dissertação (Mestrado). UNICAMP: Campinas - SP, 2008.

SANTOS, D.A.; MARÇO, P.H.; VALDERRAMA, P. Multi-product calibration: preliminary studies to determine quality parameters in industrialized juices based on ultraviolet spectroscopy. **Brazilian Journal of Analytical Chemistry**, v.12, p. 495-498, 2013.

SILVA, T. A. A.; SAMPAIO, C. S.; FURTADO, J. E. A. S.; ABILIO, G. M. F.; XAVIER, A. F. C.; CAVALCANTI, A. L. Avaliação do potencial Erosivo de bebidas à base de soja. **Revista Brasileira de Ciência e Saúde**, v.14, p. 109-114, 2010.

SILVA, V. A. G.; TALHAVINI, M.; PEIXOTO, I. C. F.; ZACCA, J. J.; MALDANER, A. O.; BRAGA, J. W. B. Non-destructive indentification of diferente types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis. **Microchemical Journal**, v.116, p. 235-243, 2014.

SKOOG, D.A.; WEST, D.M.; HOLLER, F.J.; CROUCH, S.R. **Fundamentos de química analítica**. São Paulo: Pioneira Thompson Learning, cap. 21, p. 552-556, 2006.

TAVARES, J.T.Q. ; SILVA, C.L.; CARVALHO, L.A.; SILVA, M.A.; SANTOS, C.M.G. Estabilidade do ácido ascórbico em suco de laranja submetido a diferentes tratamentos. **Magistra**, v.12, n.1/2, 2000.

TEIXEIRA, M.; MONTEIRO, M. Degradação da vitamina C em sucos de fruta. **Alimentos e Nutrição**. v.17, n.2, p. 219-227, 2006.

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.

VALDERRAMA, P.V.; BRAGA, J.W.B.; POPPI, R.J. Validation of multivariate calibration models in the determination of sugar cane quality parameters by near infrared spectroscopy. **Journal of the Brazilian Chemical Society**, v.18, n.2, p. 259-266, 2007.

VALDERRAMA, P.; BRAGA, J.W.B.; POPPI, R.J. Variable selection, outlier detection, and figures of merit estimation in a Partial Least-Squares regression multivariate calibration model. A case Study for the determination of quality parameters in the alcohol industry by near-infrared spectroscopy. **Journal of Agricultural and Food Chemistry**, v.55, p. 8331-8338, 2007b.

VENÂNCIO, A. A.; MARTINS, O. A. Análise química de diferentes marcas de néctares e suco de laranja comercializada na cidade de Cerqueira César – São Paulo. **Revista Eletrônica de Educação e Ciência - REEC**, v.2, p.1-6, 2012.

VENTURINI FILHO, W. G. Bebidas não alcoólicas: Ciência e Tecnologia. Editora: Blucher. v.2, São Paulo, 2010.

WANG, Z.; ISAKSSON, T.; KOWALSKI, B.R. New approach for distance measurement in locally weighted regression. **Analytical Chemistry**, v. 66, n. 2, p. 249-260, 1994.

WOLD, S. Principal Component Analysis: Tutorial. **Chemometrics and Intelligent Laboratory Systems**, v.2, p. 37-52, 1987.

ANEXOS

Apresentação em eventos científicos

SANTOS, D.A.; LIMA, K.P.; Março, P.H.; VALDERRAMA, P. Utilização de calibração multivariada Multiproduto na determinação de acidez total em bebidas industrializadas a base de soja. In: V Workshop de Quimiometria, 2014, Arembepe. Anais do V Workshop de quimiometria, 2014.

SANTOS, D.A.; LIMA, K.P.; MARÇO, P.H.; VALDERRAMA, P. Calibração multivariada Multiproduto e espectroscopia ultravioleta na determinação da acidez total em bebidas industrializadas a base de soja e néctar de frutas. In: XX – COBEQ – XX Congresso Brasileiro de Engenharia, 2014, Florianópolis-SC. Anais do XX COBEQ, 2014.

SANTOS, D.A.; LIMA, K.P.; MARÇO, P.H.; VALDERRAMA, P. Modelo de calibração Multiproduto em sucos industrializados a base de soja com seu teor de acidez, baseado em espectroscopia de ultravioleta. In: VI SIMTEA – Simpósio de Tecnologia e Engenharia de Alimentos, 2014, Campo Mourão-PR. Anais do VI SIMTEA, 2014.

SANTOS, D.A.; LIMA, K.P.; MARÇO, P.H.; VALDERRAMA, P. Determinação de vitamina C em sucos industrializados, utilizando calibração multivariada Multiproduto. In: IV SIMALI – Simpósio de Alimentos, 2014, Medianeira-PR. Anais do IV SIMALI, 2014.

LIMA, K.P.; SANTOS, D.A.; VALDERRAMA, P. Determinação da acidez total em bebidas industrializadas a base de néctar a partir da calibração multivariada e da espectroscopia UV. II Encontro Regional de Química da UTFPR-CM, 2014, Campo Mourão-PR.

SANTOS, D.A.; MARÇO, P.H.; VALDERRAMA, P. Resultados preliminares de calibração multivariada multiproduto para quantificação de carboidratos em sucos industrializados. In: 10 Simpósio Latino Americano de Ciência de Alimentos, 2013, Campinas. Caderno de resumos 10 SLACA, 2013.

SANTOS, D.A.; MARÇO, P.H.; VALDERRAMA, P. Calibração multivariada multiproduto para a quantificação de vitamina C em sucos industrializados – Resultados preliminares. 17º Encontro Nacional de Química Analítica, 2013, Belo Horizonte-MG.

Artigos

SANTOS, D. A.; LIMA, K.P.; MARÇO, P. H.; VALDERRAMA, P. Vitamin C determination by ultraviolet spectroscopy and multiproduct calibration. **Food Chemistry** (submetido).

SANTOS, D. A.; LIMA, K.P.; MARÇO, P. H.; VALDERRAMA, P. Determinação da acidez total em bebidas industrializadas a base de néctar a partir da calibração multivariada e da espectroscopia UV. **REBRAPA**, v.6, n.1, p. 1-8, 2015.

SANTOS, D. A.; MARÇO, P. H.; VALDERRAMA, P. Multi-product calibration: preliminary studies to determine quality parameters in industrialized juices based on ultraviolet spectroscopy. **Brazilian Journal of Analytical Chemistry**, v. 12, p. 495-498, 2013.