

Ficha Catalográfica

Elaborada Pelo Serviço Técnico de Biblioteca e Documentação
Faculdade de Ciências Farmacêuticas
UNESP – Campus de Araraquara

C331u

Carvalho, Livia Cirino de

Uso da espectroscopia do infravermelho próximo (NIR) como método não destrutivo para a discriminação de aguardentes de cana-de-açúcar produzidas no Estado de São Paulo e Minas Gerais, e predição do teor de etanol / Livia Cirino de Carvalho. – Araraquara, 2015
128 f.

Dissertação (Mestrado) – **Universidade Estadual Paulista. "Júlio de Mesquita Filho"**.
Faculdade de Ciências Farmacêuticas. Programa de Pós Graduação em Alimentos e Nutrição

Orientador: Gustavo Henrique de Almeida Teixeira

1. Infravermelho próximo. 2. Quimiometria. 3. Aguardente. 4. Cana-de-Açúcar. I.
Teixeira, Gustavo Henrique de Almeida, orient. II. Título.

CAPES: 50700006

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”

Faculdade de Ciências Farmacêuticas – Campus de Araraquara

Programa de Pós-Graduação em Alimentos e Nutrição

LÍVIA CIRINO DE CARVALHO

**USO DA ESPECTROSCOPIA DO INFRAVERMELHO PRÓXIMO (NIR) COMO
MÉTODO NÃO DESTRUTIVO PARA A DISCRIMINAÇÃO DE AGUARDENTES
DE CANA-DE-AÇÚCAR PRODUZIDAS NO ESTADO DE SÃO PAULO E
MINAS GERAIS, E PREDIÇÃO DO TEOR DE ETANOL**

Araraquara

2015

LÍVIA CIRINO DE CARVALHO

**USO DA ESPECTROSCOPIA DO INFRAVERMELHO PRÓXIMO (NIR) COMO
MÉTODO NÃO DESTRUTIVO PARA A DISCRIMINAÇÃO DE AGUARDENTES
DE CANA-DE-AÇÚCAR PRODUZIDAS NO ESTADO DE SÃO PAULO E MINAS
GERAIS, E PREDIÇÃO DE ETANOL**

Dissertação apresentada ao Programa de Pós-Graduação em Alimentos e Nutrição da Faculdade de Ciências Farmacêuticas da Universidade Estadual Paulista “Júlio de Mesquita Filho” como requisito para obtenção do grau de Mestre em Alimentos e Nutrição na área de concentração de Ciências dos Alimentos.

Orientador: Prof. Dr. Gustavo Henrique de Almeida Teixeira

Co-orientador: Prof. Dr. João Bosco Faria

Araraquara

2015

” Foi o tempo que dedicastes à tua rosa que a fez tão importante”

(Antoine de Saint-Exupéry)

AGRADECIMENTOS

A realização desta dissertação marca o fim de uma importante etapa da minha vida. Gostaria de agradecer a todos aqueles que contribuíram de forma decisiva para a sua concretização

Agradeço primeiramente a Deus, pela vida e a possibilidade de trilhar esse caminho evolutivo, por proporcionar tantas oportunidades de estudos e por colocar em meu caminho pessoas amigas e preciosas.

Aos meus pais, pois de nada adiantaria minha caminhada se vocês não tivessem me ensinado a dar os primeiros passos. Durante toda essa jornada, vocês estiveram presentes, me mostrando a importância do estudo para ser uma excelente profissional. Pai, mãe, obrigada pelo apoio incondicional ao longo deste processo de dissertação e de muitos outros. Obrigada por acreditar em mim, mesmo quando eu não acreditava, amo vocês.

Ao meu orientador Professor Doutor Gustavo Henrique de Almeida Teixeira, agradeço por ter acreditado no meu potencial, mesmo sem nunca ter me visto antes, e por ter me incentivado cotidianamente o meu aperfeiçoamento acadêmico e intelectual. Obrigada pela confiança depositada em mim, pelas críticas, sugestões e conselhos.

Ao Doutor Luis Cunha Junior, cujo incentivo, conselhos e confiança, foram importantes para tornar este trabalho uma realidade.

Agradeço também aos meus queridos amigos de Araraquara, a família que construí por aqui. Aos amigos da pós-graduação que tive o prazer de fazer, obrigada pelos conselhos, conversas, enfim, pela amizade. Gostaria de

agradecer também as pessoas que cruzaram meu caminho nessa jornada, que me acolheram tão bem em uma cidade estranha, me fazendo sentir em casa, em especial gostaria de agradecer a Aline Pereira, que além de amiga, se tornou uma irmã, quero que este laço que construímos nunca se desfaça.

Aos meus familiares em especial as minhas avós, pelo incentivo, por acreditarem em mim e por sempre me apoiarem, estive ausente, mas o meu amor por vocês não diminuiu, ao contrário se fortaleceu.

Por fim, o meu profundo e sentido agradecimento a todas as pessoas que contribuíram para a concretização desta dissertação, estimulando-me intelectualmente e emocionalmente.

SUMÁRIO

LISTA DE ABREVIACÕES E SIGLAS	9
LISTA DE FIGURAS	10
ABSTRACT	17
1. INTRODUÇÃO	19
2. OBJETIVOS	21
2.1. Objetivo Geral	21
2.2. Objetivos Específicos	21
1. REVISÃO BIBLIOGRÁFICA	22
1.1. A aguardente de cana-de-açúcar	22
1.1.1. Processo produtivo da aguardente de cana-de-açúcar	24
1.2. Qualidade da aguardente de cana-de-açúcar e Indicação Geográfica	26
1.3. Espectroscopia do infravermelho próximo (NIR)	30
1.4. Quimiometria	33
1.4.1. Pré-processamentos dos dados	34
1.4.2. Análise de Componentes Principais (PCA)	36
1.4.3. Regressão por mínimos quadrados parciais (PLSR)	36
1.4.4. Análise linear discriminante (LDA)	37
1.4.5. Máquina de vetores de suporte (SVM)	38
1.4.6. Análise Discriminante por Regressão por Mínimos Quadrados Parciais	38
1.5. Espectroscopia NIR e origem dos produtos	39
1.6. Espectroscopia NIR e a determinação do teor de etanol em bebidas alcoólicas	40
2. REFERÊNCIAS BIBLIOGRÁFICAS	43
CAPÍTULO II	55

Uso da espectroscopia do infravermelho próximo (NIR) como método não destrutivo para a discriminação de aguardentes de cana-de-açúcar produzidas nos estados de São Paulo e Minas Gerais	55
RESUMO.....	55
ABSTRACT	57
1. INTRODUÇÃO	59
2. MATERIAL E MÉTODOS	62
2.1. Obtenção das amostras	62
2.2. Obtenção dos espectros NIR	63
2.3. Determinação da coloração (CIE)	63
2.4. Análises quimiométricas.....	64
2.4.1. Pré-processamento	64
2.5. Análise Univariada	68
3. RESULTADOS E DISCUSSÃO.....	68
3.1. Coloração	69
3.2. Espectros NIR	72
3.2.1. Análise por componentes principais - PCA (técnica não supervisionada)	75
3.2.2. Análises supervisionadas (PLS-DA / LDA / SVM)	78
4. CONCLUSÕES	85
5. REFERÊNCIAS BIBLIOGRÁFICAS	87
CAPÍTULO III	93
Uso da espectroscopia do infravermelho próximo (NIR) como método não destrutivo para a determinação do teor de etanol de aguardentes de cana de açúcar	93
1. INTRODUÇÃO.....	96
2. Material e métodos.....	98

2.1. Obtenção das amostras	98
2.2. Aquisição dos espectros NIR	99
2.3. Método de referência para a determinação de etanol	99
2.4. Análises quimiométricas.....	100
2.4.1. Pré-processamento	100
2.4.2. Construção dos modelos de regressão por mínimos quadrados múltiplos	101
2.5. Análise Univariada	103
3. Resultados e Discussão.....	103
3.1 Análise Univariada	103
3.2. Espectros NIR	105
3.3. Modelos PLSR para a predição do teor de etanol.....	107
4. Conclusões.....	115
5. REFERÊNCIAS BIBLIOGRÁFICAS	116
ANEXO.....	122
Anexo 1. Nome, estado produtor e teor alcoólico das aguardente de cana-de-açúcar de São Paulo e Minas Gerais.	122

LISTA DE ABREVIações E SIGLAS

INPI - Instituto Nacional de Propriedade Industrial

LDA - Análise Linear Discriminante - Linear Discriminant Analysis

MAPA - Ministério da Agricultura, Pecuária e Abastecimento

MSC - Correção Multiplicativa de Sinal - Multiple scatter correction

NIR – Infravermelho próximo – Near infrared

PCA – Análise por Componentes Principais - Principal Component Analysis

PC- - Componentes Principais - Principal Components

PLS-DA - Análise Discriminante por Regressão por Mínimos Quadrados Parciais - Discriminant Analysis by Partial least squares

PLSR - Regressão por Mínimos Quadrados Parciais - Partial Least Squares Regression

R²_c - Coeficiente de determinação do grupo de calibração

R²_{cv} - Coeficiente de determinação do grupo de validação cruzada total

R²_p - Coeficiente de determinação do grupo de predição

RMSE_c – Erro quadrado médio de Calibração - Root Mean Square Error of Calibration

RMSE_{cv} – Erro quadrado médio de Validação Cruzada - Root Mean Square Error of Cross - Validation

RMSE_p - Erro quadrado médio de Previsão - Root Mean Square Error of Prediction

RPD – Relação do Desempenho do Desvio - Residual prediction deviation

SNV - Variação Normal Padrão - Standard normal variate

SVM - Máquina de Vetores de Suporte - Support Vector Machine

VL - Variáveis Latentes

LISTA DE FIGURAS

CAPÍTULO II

- Figura 1.** Distribuição das aguardentes de cana-de-açúcar produzidas no estado de São Paulo (SP) e Minas Gerais (MG) de acordo com os componentes principais, PC1, 77,17% de variância explicada e PC2, 19,99% de variância explicada e os vetores dos parâmetros de coloração (L, a* e b*).....70
- Figura 2.** Amostras de aguardente de cana-de-açúcar e seus respectivos valores de luminosidade (L*). *número da amostra conforme o descrito no Anexo 1..... 71
- Figura 3.** Espectros NIR das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais.....72
- Figura 4.** Espectros na região do infravermelho próximo (NIR) das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais após a utilização da correção do espalhamento multiplicativo de sinal (MSC).....73
- Figura 5.** Espectros na região do infravermelho próximo (NIR) das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais após a transformação utilizando a segunda derivada de Savitsky-Golay (2ªD S-G), com janela de 4+4 pontos.....74

Figura 6. Escores do PC1 (73% de variância explicada) e PC2 (23% de variância explicada) obtidos com os espectros NIR (1000-2500 nm) sem pré-processamento das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais.....76

Figura 7. Escores do PC1 e PC2 obtidos com espectros NIR (1000-2500 nm) das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais após a utilização da correção do espalhamento multiplicativo de sinal (MSC) (A) e segunda derivada de Savitsky-Golay com janela de 4+4 pontos (B).....77

Figura 8. Valores de predição e referência de classes de aguardentes de cana-de-açúcar produzidas nos estados de São Paulo (1) e Minas Gerais (0) utilizando o modelo PLS-DA pré-processados com MSC e valor de corte de 0,5.....80

CAPÍTULO III

Figura 1. Espectros na região do infravermelho próximo de amostras puras de água, etanol, metanol e amostras aguardentes de cana-de-açúcar.....106

Figura 2. *Loadings* da análise de regressão PLS, utilizados para predizer o teor de etanol em amostras de aguardentes de cana de açúcar por espectroscopia do infravermelho próximo. *Loadings* dos espectros sem pré-processamento.....110

Figura 3. Teor de etanol (% v/v) obtidos pelo método de referência e os teores preditos pelo modelo PLS usando os espectros NIR sem pré-processamento e as faixas de comprimentos de ondas (1.393 -1.541 + 1.843 – 2.166 nm).....115

LISTA DE TABELAS

CAPÍTULO II

- Tabela 1.** Estatística descritiva dos parâmetros de coloração (L^* , a^* , b^*) das amostras de aguardente de cana-de-açúcar oriundas do estado de São Paulo e Minas Gerais.....63
- Tabela 2.** Desempenho de classificação e predição dos modelos de análise discriminante com método de regressão por mínimos quadrados parciais (PLS-DA) desenvolvidos com e sem pré-processamento no comprimento de onda de 1000 -2500 nm para aguardente de cana de açúcar em função do estado produtor, estado de São Paulo e Minas Gerais.....79
- Tabela 3.** Desempenho de classificação e predição dos modelos de análise linear discriminante (LDA) desenvolvidos com e sem pré-processamento no comprimento de onda de 1000 -2500 nm para aguardente de cana-de-açúcar em função do estado produtor, São Paulo e Minas Gerais.....82
- Tabela 4.** Desempenho de classificação e predição dos modelos desenvolvidos pela técnica de máquina de vetores de suporte (SVM), com e sem pré-processamento dos espectros no infra vermelho próximo (1000 -2500 nm) de aguardentes de cana-de-açúcar em função do estado produtor, estado de São Paulo e Minas Gerais.....84

CAPITULO III

Tabela 1. Teor de etanol (% v/v) em aguardentes de cana-de-açúcar produzidas nos estados de São Paulo e Minas Gerais nas populações obtidas após divisão pelo algoritmo de Kennard-Stone.....100

Tabela 2. Desempenho da etapa de calibração e predição de modelos de regressão dos mínimos quadrados parciais (PLS) para determinação do teor alcoólico de aguardentes de cana-de-açúcar com diferentes pré-processamentos de espectros de reflectância na região do infravermelho próximo (NIR) (1000-2500nm) de aguardentes de cana de açúcar dos estados de São Paulo e Minas Gerais.....108

Tabela 3. Resultado de calibração e predição de modelos de regressão dos mínimos quadrados parciais (PLS), em diferentes regiões do comprimento de onda, para determinação do teor alcoólico de aguardentes de cana-de-açúcar com diferentes pré-processamentos de espectros de reflectância na região do infravermelho próximo (NIR) de aguardentes de cana de açúcar dos estados de São Paulo e Minas Gerais.....112

RESUMO

A valorização e qualidade da aguardente de cana-de-açúcar é influenciada pelas condições de produção, tais como, a escala da produção, qualidade das instalações e processos produtivos e conhecimento tecnológico do processo de produção no que diz respeito às técnicas de destilação e/ou no procedimento de padronização do produto. Desta forma, o desenvolvimento de métodos que estabeleçam a autenticidade e caracterização das aguardentes de cana-de-açúcar produzidas em diferentes estados é fundamental para a garantia de qualidade do produto ofertado, possibilitando que produtos diferenciados recebam preços superiores e impeçam o uso inapropriado de nomes de regionais por produtores fora das mesmas. Desta forma, o objetivo deste trabalho foi desenvolver uma metodologia utilizando a espectroscopia NIR e ferramentas quimiométricas como método não destrutivo para atribuição de autenticidade de aguardentes de cana-de-açúcar produzidas no estado de São Paulo em comparação a outras produzidas no estado de Minas Gerais, assim como prever o teor de etanol das mesmas. Coletou-se 139 amostras de aguardente de cana-de-açúcar de Minas Gerais (n=70) e São Paulo (n=69) e estas foram divididas em dois grupos, o conjunto de calibração (n = 100) e conjunto previsão (n= 39). Os espectros na região do infravermelho próximo (NIR) foram coletados utilizando um espectrômetro FT-IR Spectrum 100N. O método de referência para determinação do teor de etanol foi o densimétrico, seguindo as recomendações da NBR 59992 da ABNT de 2008. A classificação em função do estado produtor foi feita empregando as análises de componentes principais (PCA), análise discriminante por regressão por mínimos quadrados parciais (PLS-DA), análise linear discriminante (LDA) e

máquina de vetores de suporte (SVM), com os espectros pré-processados utilizando a correção multiplicativa de sinal (MSC) e 2ª Derivada de Savitsky-Golay. Para predição do teor de etanol utilizou-se a regressão por mínimos quadrados parciais (PLSR), empregando os pré-processamentos MSC, variação normal padrão (SNV) e 2ª Derivada de Savitsky-Golay. O melhor resultado foi obtido quando se utilizou o SVM (87,2%), em espectros sem pré-processamento. Para a predição do teor de etanol, ao se utilizar os comprimentos de onda 1393-1541 nm e 1843-2166 nm foi possível obter um $RMSE_p$ de 1,98% e um RPD de 2,02. Deste modo, a espectroscopia NIR pode ser utilizada como um método não destrutivo para atribuição da autenticidade de aguardentes de cana-de-açúcar, e também para prever o teor de etanol.

ABSTRACT

The sugarcane distillate to be considered cachaça or sugarcane aguardente must comply with the parameters established by the Normative Instruction number 13, June 30th, 2005, of the Ministério da Agricultura, Pecuária e Abastecimento (MAPA), which established the Identity and Quality Standards (PIQ) for these spirits. The prestige of these spirits can be assigned to aspects well known related to the production site which can influence its characteristics. The quality of sugarcane aguardente is enhanced by the production conditions such as the scale of production, quality of facilities, production process, the alcoholic content of these spirits not in accordance with the law, regardless of site of origin of the samples are indicative of lack of technical knowledge of the production process in respect of distillation techniques and/or the product standardization procedure. Thus, the development of methods to establish the authenticity and characterization of Spirits cane sugar produced in different states is fundamental for the quality assurance of the offered product, enabling differentiated products receive higher prices and prevent inappropriate use of regional names by producers. Near-infrared spectroscopy (NIR) has been used for discrimination of various products and quantification of compounds in foods and beverages, the aim of this study was to develop a methodology using NIR spectroscopy and chemometric tools such as non-destructive method for sugarcane aguardente authenticity assignment produced in São Paulo compared to other produced in the state of Minas Gerais, as well as predict their ethanol content. It was collected 139 samples of sugarcane aguardente of Minas Gerais (n = 70) and

São Paulo (n = 69), they were divided into two groups, the calibration set (n = 100) and set prediction (n = 39). The near-infrared spectra were collected using a spectrometer FT-IR Spectrum 100N. The reference method for determining the ethanol content was the density, following the recommendations of NBR 59992 of ABNT 2008. The classification according to producer state was determined using the analysis: principal component analysis (PCA), discriminant analysis by partial least squares (PLS-DA), linear discriminant analysis (LDA) and support vector machine (SVM), and the spectra were preprocessed using the multiplicative signal correction (MSC) and 2nd Derivative of Savitskay-Golay. It was possible to obtain a classification of sugarcane aguardente of between states an accuracy of 87.2% in the three chemometric techniques. However the best model was developed by using SVM, because did not use pre-processing in the spectra. The PLSR model for the ethanol content using the wavelength 1393-1541 nm and 1843-2166 reached a RMSEP of 1.98% and a RPD of 2.02. Therefore, the NIR spectroscopy combined with SVM technique presented a satisfactory rating (87.2%) of sugarcane aguardente produced in two different states. The PLSR model for ethanol demonstrated potential to predict sugarcane aguardente offering the possibility to use as a tool to quickly determine and nondestructively the ethanol content. Thus, near infrared spectroscopy can be used as a nondestructive method for assigning the authenticity of sugarcane aguardente and also predict their ethanol content.

1. INTRODUÇÃO

A cachaça é a denominação típica e exclusiva da aguardente de cana produzida no Brasil, com graduação alcoólica entre 38 a 48% em volume a 20°C (BRASIL, 2005) e é considerada a terceira bebida mais consumida no mundo (CBRC, 2012). Esta bebida apresenta grande importância econômica no Brasil, sendo seu processo de produção constantemente melhorado visando a obtenção de uma bebida de qualidade, tanto do ponto de vista químico quanto sensorial. Em 2008, a produção de aguardente de cana-de-açúcar foi de 1,2 bilhões de litros (IBRAC, 2008) e em 2014 a aguardente de cana-de-açúcar estava sendo exportada para 66 países, gerando uma renda de US\$ 18,33 milhões, 10% a mais em relação a 2013, sendo o volume exportado de 10,18 milhões de litros (IBRAC, 2015).

Produtos de alta qualidade e com clara identidade regional estão sendo valorizados pelos consumidores, pois expressam relações de sentimento de patriotismo, preferências culinárias, qualidades organolépticas ou propriedades benéficas à saúde, associando o produto às regiões produtoras (TREGGAR et al., 1998; VAN DER LANS et al., 2001). Desta forma, a conquista de espaços nos mercados nacional e internacional está vinculada à manutenção dos padrões de qualidade da bebida, sendo estes definidos, no Brasil, pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA), BRASIL (2005).

A “cachaça”, “Brasil” e “cachaça do Brasil” são Indicações Geográficas, definidas através do Decreto número 4062, de 21 de dezembro de 2001 (BRASIL, 2001). Por este Decreto, o nome “cachaça” constitui indicação geográfica para os efeitos legais no comércio internacional.

Estudos demonstram a possibilidade de se determinar a qualidade ou a origem geográfica de produtos agrícolas por meio de técnicas analíticas (NORDON et al., 2005), como a cromatografia líquida de alta eficiência (HPLC), cromatografia gasosa (CG), cromatografia líquida (CL), espectrometria de massa, análise elementar, e isotópica análise (COZZOLINO, 2014). Dentre os métodos analíticos existentes, a espectroscopia do infravermelho próximo (NIR) apresenta um grande potencial nas certificações da origem dos produtos (NICOLAI et al., 2007), sendo considerada uma técnica rápida, segura de operação, não destrutiva, relativamente barata e proporciona automação dos processos de controle de qualidade dos produtos agroindustriais (PASQUINI, 2003).

Vários trabalhos relatam o uso da espectroscopia NIR, juntamente com técnicas quimiométricas, para a discriminação e caracterização de vinhos produzidos em diferentes regiões geográficas (LLÉTI et al., 2005; BEVIN et al., 2006; LIU et al., 2008; MARTELO-VIDAL&VAZQUEZ, 2014), para a avaliação de alguns parâmetros, entre eles o teor de etanol nas indústrias de vinho (URBANO-QUADRADO et al., 2004; URTUBIA et al.,2004; COZZOLINO et al.,2006; MARTELLO-VIDAL e VÁSQUEZ, 2014), cerveja (CASTRITIUS et al., 2010) e vodca (KOLOMIETS et al., 2010).

Em função das potencialidades da espectroscopia NIR no processo de controle de qualidade de outras bebidas alcoólicas, o presente trabalho foi concebido com os seguintes objetivos:

2. OBJETIVOS

2.1. Objetivo Geral

Desenvolver uma metodologia utilizando a espectroscopia do infravermelho próximo (NIR), como método não destrutivo utilizando técnicas quimiométricas para a atribuição da autenticidade de aguardentes de cana-de-açúcar produzidas no estado de São Paulo em comparação com as produzidas no estado de Minas Gerais.

2.2. Objetivos Específicos

Classificar as amostras de aguardentes de cana-de-açúcar de acordo com a região produtora e servir como método analítico para se avaliar fraudes e conferir autenticidade as aguardentes de cana-de-açúcar produzidas nestes estados;

Desenvolver modelos de predição para a determinação do teor de etanol nas amostras oriundas de São Paulo e Minas Gerais.

CAPÍTULO I

1. REVISÃO BIBLIOGRÁFICA

1.1. A aguardente de cana-de-açúcar

A aguardente de cana-de-açúcar é uma bebida tipicamente brasileira e, apesar de não haver um registro preciso sobre o verdadeiro local onde a primeira destilação do vinho da cana-de-açúcar tenha sido feita, pode-se afirmar que ela se deu no território brasileiro entre os anos de 1516 e 1532 (MARTINS, 2011).

No Brasil são produzidos cerca de 1,4 bilhões de litro de aguardente de cana-de-açúcar por ano, gerando uma renda de 7 bilhões de reais nesta cadeia produtiva, porém apenas 1% do total produzido é exportado, principalmente para a Alemanha, Paraguai, Portugal, Estados Unidos da América (EUA) e Bolívia, sendo considerada a terceira bebida mais consumida no mundo (CBRC, 2012).

O destilado de cana-de-açúcar para ser considerado aguardente de cana-de-açúcar, deve obedecer aos parâmetros estabelecidos pela Instrução Normativa número 13, de 30 de junho de 2005, do Ministério de Agricultura, Pecuária e Abastecimento (MAPA), que estabelece os Padrões de Identidade e Qualidade (PIQ) para aguardente de cana e para a cachaça. Desta forma, cachaça é a denominação típica e exclusiva da aguardente de cana produzida no Brasil, com graduação alcoólica entre 38 a 48% em volume a 20°C, obtida pela destilação do mosto fermentado do caldo de cana-de-açúcar,

apresentando características sensoriais peculiares, este caldo pode ser adicionada de açúcares com valor máximo de 6 g.L⁻¹. Ainda segundo esta Instrução Normativa, temos a aguardente de cana, que é definida da mesma maneira, porém com graduação alcoólica entre 38 a 54% em volume a 20°C e ainda o destilado de alcoólico simples de cana-de-açúcar, este possui graduação alcoólica superior a 54% e inferior a 70% em volume a 20°C (BRASIL, 2005).

Estas bebidas são constituídas majoritariamente por etanol e água, e também por compostos secundários, tais como álcoois superiores, ácidos, ésteres, acetais, fenóis, hidrocarbonetos, compostos nitrogenados, sulfurados e açúcares, entre outros, que as caracterizam e as qualificam de acordo com as concentrações presentes nas bebidas (ODELLO et al., 2009).

Os compostos secundários devem apresentar concentrações entre 200 mg.100 mL⁻¹ a 650 mg.100 mL⁻¹ de álcool anidro, com os seguintes limites máximos para cada componente: 150 mg.100 mL⁻¹ de álcool anidro para acidez volátil (expressa em ácido acético), 200 mg.100 mL⁻¹ de álcool anidro de ésteres, 30 mg.100 mL⁻¹ de álcool anidro de aldeídos totais, 5 mg.100 mL⁻¹ de álcool anidro da soma de furfural e hidroximetilfurfural e 360 mg.100 mL⁻¹ de álcool anidro de álcoois superiores (soma de álcoois n-propílico, isobutílico e isoamílico), (BRASIL, 2005).

A aguardente de cana-de-açúcar pode ainda ser envelhecida em tonéis de madeira, por um período não inferior a um ano, este procedimento acarreta em alterações na composição química, no aroma, no sabor e na coloração da bebida (RODRIGUES et al., 2014). Estes autores ressaltaram que as diferenças entre os tipos de aguardente de cana-de-açúcar envelhecidas estão

relacionadas ao tempo de armazenamento, a sua vida de prateleira e ao tipo de madeira utilizada no processo.

1.1.1. Processo produtivo da aguardente de cana-de-açúcar

No Brasil, aproximadamente 75% da produção total de aguardente de cana-de-açúcar é realizada em colunas de destilação e 25% em alambiques, sendo o estado de Minas Gerais o maior produtor nacional de aguardente de cana-de-açúcar de alambique (BERNARDES et al., 2014). A produção da aguardente de cana-de-açúcar de coluna e de alambique apresentam diferenças relacionadas à escala e ao sistema de produção (matéria-prima e fermentação).

Nos alambiques geralmente são produzidos de cem a mil litros de aguardente de cana-de-açúcar por dia e nas destilarias industriais são produzidos cerca de trezentos mil litros por dia (JERONIMO & SILVA, 2005). Ambos os sistemas possuem suas vantagens, sendo a aguardente de cana-de-açúcar de alambique conhecida por ser de sabor diferenciado, enquanto que a de coluna possibilita facilidades no processo de padronização do produto (SILVA et al., 2013).

O processo produtivo da aguardente de cana-de-açúcar envolve etapas como o plantio, a colheita, a moagem, a fermentação, a destilação, podendo ocorrer o envelhecimento ou não, o engarrafamento e a embalagem. Após o plantio, colheita e moagem da cana-de-açúcar, inicia-se a fermentação do seu caldo (PINHEIRO et al., 2003). Esta etapa exerce a principal influência sobre a qualidade do produto, pois os componentes menores, tais como, álcoois, ácidos orgânicos, compostos carbonílicos e ésteres, são gerados e estes são

responsáveis pelo sabor e aroma no produto final (SCANAVINI et al., 2012). Portanto, leveduras adequadas devem ser selecionadas para o sucesso da fermentação, cujas características desejáveis incluem: um início rápido e com altas taxas de fermentação, tolerância ao estresse, alto consumo e absorção do açúcar, baixa produção de ácido acético, fermentação completa do substrato e a produção eficiente de etanol (SILVA et al., 2009).

Após a fermentação, é feita a destilação, que pode ocorrer: em alambiques de cobre, onde a destilação é feita em batelada, ou em colunas de aço inoxidável, sendo a destilação contínua. Enquanto na destilação em batelada, em alambique de cobre, o destilado é dividido em cabeça, coração e cauda, na destilação feita em coluna o destilado não é dividido, pois o sistema é contínuo (BECHE & FRANCO, 2009).

A destilação influencia as características sensoriais das bebidas alcoólicas devido à concentração de compostos voláteis absolutos e relativos presentes e por reações químicas provocadas durante o aquecimento (DANTAS et al., 2007). Segundo Dantas et al. (2007), a destilação deve ser a mais lenta possível, dentro dos limites econômicos, para se obter alto rendimento e um bom produto. Podem ocorrer duas destilações sucessivas, o que é conhecido como bidestilação, podendo ser realizadas no mesmo ou em diferentes alambiques (ROTA & FARIA, 2009).

O envelhecimento da aguardente de cana-de-açúcar é realizado em tonéis de madeira, com o carvalho sendo a madeira normalmente utilizada. Os tonéis desta madeira são vendidos a um alto preço para os produtores brasileiros, por não ser uma árvore típica de climas tropicais (BERNARDES et al, 2014).

Devido à grande diversidade da flora brasileira, outros tipos de madeira são utilizados para o envelhecimento, tais como o amendoim (*Pterogune nitens*), o bálsamo (*Myroxylon peruiferum*), o jequitibá (*Cariniana estrellensis*) e a umburana (*Amburana cearenses*), (BERNARDES et al., 2014; BORBOLETTO & ALCARDE, 2013; SOUZA et al., 2014). Nesta etapa, o conteúdo alcoólico é reduzido e inúmeras reações químicas entre a bebida e a madeira ocorrem e realçando suas propriedades organolépticas, tais como cor, sabor e aroma (SOUZA et al., 2014).

1.2. Qualidade da aguardente de cana-de-açúcar e Indicação Geográfica

Apesar da tradição e importância econômica da aguardente de cana-de-açúcar no Brasil, tem-se buscado o desenvolvimento de tecnologias para aperfeiçoar o processo de controle de qualidade e padronização da bebida, visando ao atendimento dos padrões internacionais, garantindo assim a qualidade do produto para o mercado interno e externo (MIRANDA et al., 2007).

Para atender aos padrões nacionais e internacionais, o Brasil passou a adotar uma legislação para a produção de aguardente de cana-de-açúcar. Dentre as leis e decretos que fiscalizam a produção desta bebida, pode-se citar a Lei número 8918, de 15 de julho de 1994, que dispõe sobre a padronização, a classificação, o registro, a inspeção, a produção e a fiscalização de bebidas, entre elas a aguardente de cana-de-açúcar (BRASIL, 1994), sendo regulamentada pelo Decreto número 6871, de 04 de junho de 2009 (BRASIL, 2009), bem como a Instrução Normativa número 13, de 29 de junho de 2005, do MAPA que dispõem sobre uma série de requisitos como padronização,

classificação, registro, inspeção, produção e a fiscalização para a aguardente de cana-de-açúcar no Brasil (BRASIL, 2005).

A qualidade não só da aguardente de cana-de-açúcar, mas, como de todo produto alimentício, segundo a visão do consumidor, é o conjunto da qualidade nutricional, organoléptica, de uso, regulamentar e sanitária (RODRIGUES et al., 2011). Estes autores definem como qualidade nutricional a capacidade do alimento em satisfazer as necessidades fisiológicas do ser humano e se relacionam com as vitaminas, proteínas lipídios, glicídios, sais minerais. A qualidade organoléptica relaciona-se com as características sensoriais do alimento. A qualidade de uso com a facilidade e praticidade de utilização de um alimento no que diz respeito à facilidade de manipulação, conservação, preparação, etc. Já a qualidade regulamentar se refere ao respeito às normas de conservação, etiquetagem, higiene, e, por último, a qualidade sanitária se atem a ausência de contaminação química, física e biológica. Sendo a qualidade organoléptica a que possui mais influência na aguardente de cana-de-açúcar.

O Brasil apresenta uma grande diversidade de produtos agropecuários com qualidade diferenciada, podendo ter forte influência de acordo com sua origem geográfica devido a fatores étnicos, culturais, geográficos e climáticos. Reconhecer esta qualidade em relação à sua origem é uma prática antiga (BARROS & SANTOS, 2011), que envolvem sentimentos de patriotismo, preferências culinárias, qualidades organolépticas ou propriedades benéficas a saúde (TREGGAR et al., 1998; VAN DER LANS et al., 2001), sendo consideradas um fator relevante para o consumidor na tomada decisão de compra por um produto em detrimento de outro.

No Brasil, o Instituto Nacional de Propriedade Industrial (INPI) é o órgão responsável por fiscalizar a Indicação Geográfica (IG) dos produtos brasileiros, entre eles a aguardente de cana-de-açúcar. A IG é definida como um registro conferido a produtos ou serviços que são característicos do seu local de origem e também garante credibilidade do produto e potencializa seu valor comercial. Isso porque a IG permite que se mantenha o padrão de qualidade, relacionadas ao local de produção, e inibe que pessoas utilizem o nome da região em seus produtos indevidamente (SEBRAE, 2014).

No Brasil, a Lei número 9279, de 14 de maio de 1996 (BRASIL, 1996) define duas formas de indicação geográfica, a Indicação de Procedência (IP) e a Denominação de Origem (DO). O artigo 177 da referida Lei, define a IP como sendo o nome geográfico de país, cidade, região ou localidade de seu território que se tenha tornado conhecido como centro de extração, produção ou fabricação de determinado produto ou de prestação de determinado serviço. Enquanto que o artigo 187 considera a DO como sendo o nome geográfico de país, cidade, região ou localidade de seu território, que designe produto ou serviço cujas qualidades ou características se devam exclusiva ou essencialmente ao meio geográfico, incluídos fatores naturais e humanos.

Na DO, as características geográficas (solo, subsolo, vegetação), meteorológicas (mesoclima) e humanas (cultivo, tratamento, tradição, cultura) que participam de todo o processo de produção são determinantes na caracterização e na diferenciação do produto. Por outro lado, na IP basta que o bem produzido nessa área possua reputação e notoriedade junto ao mercado consumidor (VALENTE et al., 2012).

De maneira geral, a obtenção da IG assegura aos produtores o reconhecimento das características qualitativas de seus produtos (notoriedade) o que pode ser frequentemente associado à diferenciação de preços (COZZOLINO, 2012), agregando valor ao produto e conferindo-lhe um diferencial de mercado em função das características de seu local de origem (VALENTE et al., 2008). Esta certificação tem se tornado cada vez mais popular (VAN DER LANS et a., 2001), devido ao crescente interesse por parte dos produtores de alimentos em assegurar aos consumidores a proveniência dos seus produtos (COZZOLINO, 2014).

A “cachaça”, “Brasil” e “cachaça do Brasil” são Indicações Geográficas, definidas através do Decreto número 4062, de 21 de dezembro de 2001 (BRASIL, 2001). Por este Decreto, o nome “cachaça” constitui indicação geográfica para os efeitos, no comércio internacional.

No Brasil, a primeira Indicação de Procedência (IP) registrada para a cachaça foi a de Paraty – RJ, em 2007. Atualmente, há mais duas IPs para a cachaça, a da Região de Salinas – MG, cujo registro foi concedido em 2009, e a da Micro Região da Abaíra – BA, concedido em 2014 (INPI, 2015).

O uso inapropriado de indicações geográficas por partes não autorizadas é prejudicial aos consumidores e aos legítimos produtores, sendo no Brasil considerado crime, conforme descrito na Lei número 9279, de 14 de maio de 1996, nos artigos 192 a 193 (BRASIL, 1996).

Desta forma, o desenvolvimento de novas e sofisticadas técnicas para a determinação da origem geográfica de produtos agrícolas é altamente interessante para os consumidores, produtores rurais, revendedores e autoridades administrativas (COZZOLINO, 2012).

As diferenças da qualidade da aguardente de cana-de-açúcar são intensificadas pelas condições e perspectivas dos produtores em termos de escala da produção, qualidade das instalações e processo produtivo (DIESEL et al., 2005). Bogusz et al. (2006) observaram variações no teor alcoólico desta bebida, algumas não estando dentro do que estabelece a legislação, que podem ser causadas por diferenças no tipo de fermento empregado, na levedura utilizada, no sistema e na temperatura de aquecimento bem como pelo tipo de aparelho de destilação.

Teores alcoólicos para bebidas destiladas em não conformidade com a legislação, independentemente do local de origem das amostras, são um indicativo de falta de conhecimento tecnológico do processo de produção no que diz respeito a técnicas de destilação e/ou no procedimento de padronização do produto (BOGUSZ et al., 2006).

Diante disto, torna-se de extrema relevância fazer a determinação do teor alcoólico da aguardente de cana-de-açúcar, com intuito de padronizá-la de acordo com os requisitos legais e evitar fraudes, assegurando a qualidade da bebida.

1.3. Espectroscopia do infravermelho próximo (NIR)

Os métodos indicados para determinação da origem geográfica em alimentos são: cromatografia líquida de alta eficiência (HPLC), cromatografia gasosa (CG), cromatografia líquida (CL), espectrometria de massa, análise elementar, e isotópica análise (COZZOLINO, 2014). Entretanto, estas técnicas são destrutivas, dispendiosas e geradoras de resíduos químicos.

Estudos tem demonstrado a possibilidade de se determinar a qualidade e/ou origem geográfica de produtos agrícolas de forma não destrutiva, utilizando métodos como a espectroscopia Raman, apontada para determinação do teor alcoólico em bebidas (NORDON et al., 2005), a técnica de ultrassom, para avaliar a qualidade da madeira (MINÁ et al., 2004), a aplicação da tomografia de ressonância magnética nuclear (RMN) que fornece informações sobre a qualidade das frutas (MATTIUZ et al., 2002).

Dentre as técnicas não destrutivas, a espectroscopia do infravermelho próximo (NIR) apresenta um grande potencial de uso nas certificações de origem produtos (NICOLAÏ et al., 2007), considerada uma técnica rápida, segura de operação, não destrutiva, relativamente barata e proporciona automação dos processos de controle de qualidade nos produtos da agroindústria (PASQUINI, 2003). Neste contexto, a espectroscopia do infravermelho próximo (NIR) pode ser utilizada para conferir autenticidade aos produtos oriundos de diferentes regiões produtoras (COZZOLINO, 2012).

A radiação na região do infravermelho próximo, em inglês, *Near Infrared* (NIR), foi descoberta pelo astrônomo alemão, Frederick William Herschel em 1800 (HERSCHEL, 1800). Embora Herschel tenha descoberto a radiação no NIR em 1800, muitos espectroscopistas na primeira metade do último século ignoravam esta região para algum interesse analítico. As aplicações mais recentes da espectroscopia NIR foram relatadas por volta dos anos de 1950 e tornaram-se ainda mais evidentes quando o grupo dirigido por Karl Norris, em 1965, aplicou-a na área da agricultura. O desenvolvimento da instrumentação, da eletrônica, dos componentes ópticos, o advento dos computadores capazes

de processamento e armazenamento de dados foram fatores que facilitaram a expansão desta técnica em diversas áreas (BLANCO & VILLARROYA, 2002).

Segundo Pasquini (2003), a espectroscopia NIR é um tipo de espectroscopia vibracional que emprega a energia dos fótons ($h\nu$) na faixa de $2,65 \times 10^{-19}$ a $7,96 \times 10^{-20}$ J, que corresponde a faixa de comprimento de onda de 750 a 2.500 nm (13.000 a 4.000 cm^{-1}). A radiação NIR excita as moléculas levando-as a apresentarem movimentos harmônicos e combinações de vibrações moleculares de níveis energéticos mais elevados, permitindo que a menor absorbância da radiação incidente, infravermelho próximo, penetre em amostras opacas, provocando vibrações moleculares em elétrons externos, sem necessidade de preparação prévia, modificação da matriz ou diluição do material (CIENFUEGOS & VAITSMAN, 2000).

Esta técnica analítica é classificada como versátil e universal de análise estrutural para uma grande variedade de amostras orgânicas, inorgânicas, como complexos metálicos, e biológicas, podendo ser utilizada para determinações quali e quantitativas. O espectro de infravermelho é considerado a impressão digital (fingerprint) de um composto químico e o processo é simples, como descrito por Cienfuegos e Vaitsman (2000), onde a amostra é analisada no equipamento e um valor de absorbância é determinado para cada comprimento de onda da região do infravermelho próximo. Os resultados são armazenados no computador, o que significa que o equipamento tem dois resultados, o do laboratório convencional e o obtido através do espectro. Uma relação é calculada e o resultado dessa relação é a curva de calibração, criando um banco de dados com os resultados obtidos.

Dentre as características da espectroscopia NIR, as mais significantes são: a sua velocidade (um minuto ou menos por amostra), aplicação universal (qualquer molécula contendo C-H; N-H; S-H ou bandas de O-H), necessita de o mínimo de preparação da amostra, é não destrutivo e não invasivo (PASQUINI, 2003). Segundo Pasquini (2003), o fato desta técnica ser não destrutiva e não invasiva se deve ao uso da refletância difusa, comprovado por um trabalho realizado pelo pesquisador Karl Norris(1965) na determinação de umidade em produtos agrícolas.

A utilização desta técnica na avaliação dos parâmetros físico-químicos oferece inúmeras vantagens como a possibilidade de realizar medidas em um grande número de amostras, assim como, repetir a análise em tempo real sobre a mesma amostra acompanhando as suas mudanças (NICOLAI et al., 2007).

Deste modo, a espectroscopia NIR pode ser utilizada para identificar um composto ou investigar a composição de uma amostra (SIMÕES, 2008).

1.4. Quimiometria

Quimiometria é o campo da química que utiliza ferramentas estatísticas e matemáticas para o planejamento e otimização das condições experimentais, e para a extração de informação química relevante de dados químicos (RIBEIRO et al., 2007). No Brasil podem ser agrupada em três áreas principais: planejamento e otimização de experimentos, reconhecimento de padrões abrangendo métodos de análise exploratória e classificação, e calibração multivariada (NETO et al., 2006).

O reconhecimento de padrões desenvolve ferramentas computacionais que permitem explorar os resultados obtidos por meio de análises químicas, com intuito de verificar similaridades entre as amostras que possuem correlação com as semelhanças na composição química (CORREIA & FERREIRA, 2007)

A calibração multivariada é considerada uma das subáreas mais importantes da quimiometria, pois permite estimar com precisão o valor da propriedade desejada em matrizes de amostras complexas, com forte interferência química ou física (COSTA FILHO & POPPI, 2002).

Com o crescimento da quimiometria foram desenvolvidas novas ferramentas para tratamento de dados, encontrando aplicações distintas conforme o objetivo do estudo (SOUZA & POPPI, 2012). As técnicas quimiométricas tem sido amplamente utilizada na espectrometria NIR, pois é considerada uma ferramenta valiosa para retirar as informações contidas nos espectros NIR, uma vez que o espectro NIR possui bandas largas e sobrepostas, a sua aplicação permite a identificação e a quantificação de diversos parâmetros em diferentes matrizes (SOUZA et al., 2013).

1.4.1. Pré-processamentos dos dados

Muitas vezes é necessário realizar uma etapa de pré-processamento dos espectros NIR para remover informações físicas e/ou instrumentais resultantes de pequenas vibrações, granulometria, correntes de ar, movimentos aleatórios das moléculas, umidade presente na amostra, que não estão relacionadas à informação química de interesse da amostra (COZZOLINO, 2011; CHEN et al., 2013; SOUZA et al., 2013).

Alguns dos pré-processamentos mais frequentemente utilizados aos espectros NIR incluem os que visam minimizar a interferência relativa ao espalhamento de luz. Como exemplo, pode-se citar a correção do espalhamento multiplicativo de sinal (MSC - Multiplicative Scatter Correction) que corrige os espectros NIR com base no espalhamento médio de todos os espectros empregados na calibração multivariada (MARTENS & NAES, 1996). Similarmente, é utilizado a variação normal padrão (SNV - Standard normal variate) que promove a remoção das interferências relativas ao espalhamento de luz e os efeitos do espalhamento multiplicativo entre diferentes amostras (BARNES et al., 1989; AZZOUZ et al., 2003).

Além destes, a primeira e segunda derivada de Savitsky-Golay (McCLURE, 1994), são empregadas com o objetivo de corrigir os deslocamentos da linha de base e acentuar picos que possam conter alguma informação importante, possibilitando maior detalhamento e desenvolvimento de estudos quali e quantitativos dos espectros NIR (SUSI e BYLER, 1983).

Depois de realizar os pré-processamentos adequados, pode-se aplicar os tratamentos dos dados espectrais, sendo impossível definir uma regra universal para qualquer estudo de calibração envolvendo dados de espectrofotometria NIR. Dentre as técnicas mais conhecidas destacam-se *Multiple Linear Regression* (MLR), *Principal Component Analysis* (PCA), *Principal Component Regression* (PCR), *Classic Least Squares* (CLS), *Partial Least Squares* (PLS). Dentre as técnicas existentes a mais utilizada para predição é a PLS (GY et al., 1998; FERREIRA et al., 1999; REBOUÇAS & BARROS NETO, 2001; PASQUINI, 2003; SANTOS et al., 2005, NICOLAÏ et al., 2007).

1.4.2. Análise de Componentes Principais (PCA)

O PCA é uma análise exploratória de dados utilizada para reduzir a dimensionalidade do conjunto de dados, preservando ao mesmo tempo o máximo da informação (NETO et al, 2006). O PCA emprega um procedimento matemático que transforma um conjunto de variáveis respostas possivelmente correlacionado a um novo conjunto de variáveis, chamados componentes principais (COZZOLINO, 2011).

Os cálculos utilizados nas construções das componentes podem ser processados por uma série de algoritmos, sendo que o mais utilizado é o *Nonlinear Interactive Partial Least Squares* (NIPALS) (DASZYKOWSKI et al., 2007).

1.4.3. Regressão por mínimos quadrados parciais (PLSR)

A regressão por mínimos quadrados parciais (PLSR) é uma técnica de análise multivariada de dados que combina as características de análise de componentes principais e regressão linear múltipla (SHETTY & GISLUM, 2011).

O seu objetivo é encontrar uma relação entre a matriz (**X**) contendo os espectros das amostras e o vetor que armazena as respectivas concentrações (**Y**). O resultado é uma equação semelhante à:

$$y = Xb + e$$

Onde **b** é o vetor de regressão e **e** é o vetor que representa os erro do modelo.

Por meio de combinações lineares dos dados espectroscópicos (matriz X) e dos dados de referência (matriz Y), se obtém o número de variáveis latentes necessárias para correlacionar os espectros e as concentrações. O número de variáveis latentes utilizados construção do modelo de calibração deve ser o que proporcione o menor erro possível de previsão ($RMSE_c$), ou seja, que as diferenças entre os valores de referência e os valores previstos sejam os menores (MORGANO et al., 2005; SOUZA et al., 2013).

Determinado o número de variáveis latentes, o modelo de regressão pode ser construído e deve ser validado empregando amostras que não participaram da calibração (validação externa), deve-se conhecer os valores da propriedade de interesse, para avaliar a real capacidade de previsão do modelo. Calcula-se nesta etapa o erro de validação ($RMSE_{cv}$) (SOUZA et al., 2013).

Este método não requer um conhecimento exato de todos os componentes presentes nas amostras podendo realizar a previsão de amostras mesmo na presença de interferentes, desde que estes também estejam presentes por ocasião da construção do modelo (BRERETON, 2000).

1.4.4. Análise linear discriminante (LDA)

O LDA é uma técnica de reconhecimento de padrão supervisionado (NAES et al., 2002) e é utilizado para extrair informação e detectar tendências nos dados, baseados nas medidas multivariadas das amostras. Nas técnicas supervisionadas é necessário que exista alguma informação inicial sobre a identidade das amostras para a formação das classes e o objetivo é

desenvolver um modelo baseado nas informações contidas nas amostras (BRERETON, 2007).

O objetivo da LDA é determinar os melhores parâmetros de ajuste para a classificação das amostras por um modelo desenvolvido. O modelo pode, em seguida, ser utilizado para classificar amostras desconhecidas (NAES et al., 2002).

1.4.5. Máquina de vetores de suporte (SVM)

O SVM é uma ferramenta de classificação binária que é baseada na teoria da aprendizagem estatística e concebida para resolver um problema de classificação, sendo considerada uma metodologia poderosa para realizar a classificação não linear, estimativa de função ou regressão multivariada não linear (XIE et al., 2009)

A técnica SVM tem sido empregada para discriminação de produtos, como demonstrado por Zhao et al. (2006) que utilizaram a espectroscopia NIR combinado com SVM para identificar chá verde, preto e chá Oolong .Desta forma o uso de SVM auxilia na identificação ou classificação de produtos, sendo utilizado para solucionar problemas de calibração não linear em função do mapeamento do espaço original dos dados em um espaço de alta dimensão, denominado de espaço de características (CORTES & VAPNIK, 1995).

1.4.6. Análise Discriminante por Regressão por Mínimos Quadrados Parciais

A regressão por Mínimos Quadrados Parciais (PLS) é um método de projeção multivariada para modelar uma relação entre variáveis dependentes (Y) e variáveis independentes (X), seu objetivo é encontrar as variáveis e direções no espaço multivariado que discriminam as classes estabelecidas no conjunto de calibração. O número ótimo de variáveis latentes pode ser estimado por meio de validação cruzada ou conjunto de teste externo (BERRUETA et al., 2007).

Em PLS-DA, uma matriz Y "simulada" é construída com zeros e uns. A matriz consiste em quantas colunas existem classes e uma observação tem o valor 1 para a classe a que pertence e 0 para o resto. A matriz X é composta dos dados originais (pré-processados). As matrizes X e Y são decompostas num produto de duas outras matrizes de escores e pesos. Assim, os pesos de X são calculados a partir dos escores de Y , enquanto que os pesos de Y são determinados a partir dos escores de X . As decomposições não são independentes, mas relacionados com os escores de X e Y . Um modelo é desenvolvido para cada classe. Quanto mais próximo um elemento de uma determinada coluna em Y está de 1 e os elementos das outras colunas está de 0, maior a probabilidade de um objeto ser um membro de uma classe particular (BERRUETA et al., 2007).

1.5. Espectroscopia NIR e origem dos produtos

A espectroscopia NIR vem sendo satisfatoriamente aplicada em bebidas alcoólicas para se determinar a origem geográfica destes produtos, como o trabalho descrito por LIU et al. (2006) que classificaram vinhos "Tempranillo" provenientes da Austrália e Espanha. A origem de vinhos de diferentes castas

oriundas da Austrália, Nova Zelândia, França e Alemanha também foi atribuída utilizando a espectroscopia NIR (LIU et al., 2008). Há aplicações também em outros produtos, como em azeite de oliva (GALTIER et al., 2007), pistaches (VITALE et al., 2013) e produtos cárneos (PRIETO et al., 2009).

É possível também encontrar trabalhos utilizando a espectroscopia NIR para prever ou quantificar componentes químicos em alimentos, podendo prever teor de proteína, matéria-seca e fósforo em grãos de milho (NETO et al., 2009), determinar proteína em café cru (MORGANO et al., 2005), ambos aplicando o método de regressão por quadrados mínimos parciais (PLS) para a análise dos dados, pode-se também caracterizar bebidas alcoólicas, como o vinho, utilizando esta técnica (SIMÕES, 2008).

Desta forma, a espectroscopia NIR pode ser utilizada visando à discriminação de amostras de aguardente de cana produzidas em diferentes regiões produtoras com o intuito de classificá-las de acordo com a região de origem e servir como método analítico para se avaliar fraudes e conferir autenticidade as bebidas, assim como prever componentes químicos de interesse qualitativo.

1.6. Espectroscopia NIR e a determinação do teor de etanol em bebidas alcoólicas

A determinação de etanol em bebidas alcoólicas é um problema importante tanto por razões sociais quanto econômicas, particularmente em relação aos impostos incidentes sobre bebidas alcólicas em diferentes países (GALLIGNANI et al., 1993), pois quanto maior o teor de etanol maior é o imposto. Assim, um método de análise quantitativo simples e preciso é

fundamental para a determinação do teor de etanol em diferentes níveis (0 a 50% v/v) para o controle de qualidade das bebidas e como forma de fiscalização por parte do governo (WANG et al., 2003).

Muitos métodos têm sido utilizados para a determinação da graduação ou teor alcoólico em bebidas. Entre os descritos pela “Association of Official Analytical Chemists” (AOAC) constam: os métodos refratométricos e métodos de densidade, que utilizam picnômetros ou densímetros para medir o teor alcoólico de bebidas destiladas (BARBOSA & POPPI, 2003). Porém, devido aos procedimentos analíticos, por exemplo, destilação de amostras, pesagem precisa, e grande volume de amostra necessário, esses métodos tornam-se inviáveis para amostras em pequena quantidade e, além disso, os reagentes utilizados são altamente tóxicos (WANG et al, 2003).

Por outro lado, outros métodos podem ser encontrados na literatura, tais como os cromatográficos, métodos eletroquímicos e a espectroscopia NIR (BARBOZA & POPPI, 2003; MENDES et al., 2003).

A técnica do infravermelho próximo com análises quimiométricas, como o PLSR, tem sido muito utilizada em bebidas alcoólicas, possibilitando a sua caracterização e a avaliação de alguns parâmetros, entre eles o teor de etanol, como na indústria de vinhos (URBANO-QUADRADO et al., 2004; URTUBIA et al., 2004; COZZOLINO et al., 2006; MARTELLO-VIDAL & VÁSQUEZ, 2014), cervejas (CASTRITIUS et al., 2010) ou em vodcas (KOLOMIETS et al., 2010). Isto é possível por que o etanol, como descrito por Cozzolino et al (2006), tem um forte sinal de absorbância na região do infravermelho próximo em bebidas alcoólicas, com uma intensidade menor apenas que a água, porém a precisão

e robustez da calibração pode ser limitada pela variação das amostras e variações na concentração de açúcar.

Desta maneira, a espectroscopia do infravermelho próximo associada à técnicas quimiométricas demonstra um grande potencial de uso quando empregada com a finalidade de caracterização e quantificação de compostos em bebidas e alimentos.

2. REFERÊNCIAS BIBLIOGRÁFICAS

- AZZOUZ, T.; PUIGDOMÉNECH, A.; TAULER, R. Comparison between different data pré-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method. **Analytica Chimica Acta**, v. 484, p. 121-134, 2003.
- BARBOSA, F.D.; POPPI, R.J. Determination of alcohol content in beverages using short-wave near-infrared spectroscopy and temperature correction by transfer calibration procedures. **Analytical and Bioanalytical Chemistry**.v.377, pp.695-701, 2003.
- BARROS, R.C.; SANTOS, T.S.; Indicação geográfica no brasil: um instrumento para o desenvolvimento rural sustentável (uma revisão). **Revista Geográfica de América**, p. 1-21, 2011.
- BECHE, R. V.; FRANCO, D. W.. Distinção entre cachaças destiladas em alambiques e em colunas usando quimiometria. **Química Nova**, São Paulo , v. 32, n. 2, p. 332-336, 2009.
- BERNARDES, C.D.; FIGUEIREDO, M.C.P. e BARBEIRA, P. J.S. Developing a PLS model for determination of total phenolic content in aged cachaças. **Microchemical Journal**, v. 116, p. 173-177, 2014.
- BERRUETA LA, ALONSO-SALCES RM, HÉBERGER K .Supervised pattern recognition in food analysis. **Journal of Chromatography**, v. 1158, p. 196-214, 2007.

- BEVIN, C. J.; DAMBERGS, R. G.; FERGUSON, A.J.; COZZOLINO, D. Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. **Analytica chimica acta**, v.621, p.19–23, 2006.
- BLANCO, M.; VILLARROYA, I. NIR spectroscopy: a rapide-response analytical tool. *Trends in analytical chemistry*, v.21, 2002.
- BOGUSZ JUNIOR, S.; MERTINS, K.D.C.; RAQUEL, G, BACARIN, A. Composição química da cachaça produzida na região noroeste do Rio Grande do Sul, Brasil. **Ciênc. Tecnol. Aliment.**, Campinas , v. 26, n. 4, p. 793-798, Dec. 2006
- BORTOLETTO, A.M. e ALCARDE, A. R. Congeners in sugar cane spirits aged in casks of different woods. **Food Chemistry**, v. 139, p. 695-701, 2013.
- BRASIL. Decreto número 4062, de 21 de dezembro de 2001. Define as expressões “cachaça”, “Brasil” e “cachaça do Brasil” como indicações geográficas e dá outras providências. **Diário Oficial da República federativa do Brasil**. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.
- BRASIL. Decreto número 6871, de 4 de junho de 2009. Regulamenta a lei número 8.918 de julho de 1994, sobre a padronização, a classificação, o registro, a inspeção, a produção e a fiscalização de bebidas. **Diário Oficial da República federativa do Brasil**. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.
- BRASIL. Instrução Normativa número 13, de 29 de junho de 2005. Aprova o regulamento técnico para fixação dos padrões de identidade e qualidade para aguardente de cana e para cachaça. **Diário Oficial da República**

federativa do Brasil. Disponível em: <<http://www.agricultura.gov.br/>>.

Acesso em 09 de maio de 2015.

BRASIL. Lei número 8918, de 15 de julho de 1994. Dispõe sobre a padronização, a classificação, o registro, a inspeção, a produção e a fiscalização de bebidas. **Diário Oficial da República federativa do Brasil.** Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

BRASIL. Lei número 9279, de 14 de maio de 1996. Dispõe sobre direitos e obrigações relativos à propriedade industrial. **Diário Oficial da República Federativa do Brasil.** Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

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

CASTRITIUS, S.; KRON, A.; SCHAFFER, T.; RADLE, M.; HARMS, D. Determination of Alcohol and Extract Concentration in Beer Samples Using a Combined Method of Near-Infrared (NIR) Spectroscopy and Refractometry. **Journal Agricultural Food Chemistry**, v.58, pp.12634-12641, 2010

CBRC- Centro Brasileiro de Referência da Cachaça. Números da cachaça. Expo cachaça, 2012. Disponível em: <http://www.expocachaca.com.br/bh/numeros-da-cachaca.shtml>. Acesso em 08 de agosto de 2014.

CIENFUEGOS, F.; VAITSMAN, D. **Análise Instrumental.** São Paulo: Interciência, 2000, 606 p.

- 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, p.46-52, 2002.
- CORREIA, P.R.M.; FERREIRA, M.M.C. Reconhecimento de padrões por métodos não supervisionados: explorando procedimentos quimiométricos para tratamento de dados analíticos. *Química nova*, v.30, p.481-487, 2007.
- COZZOLINO, D.; DAMBERGS, R.G.; JANIK, L.; CYNKAR W.U.; GISHEN, M. Analysis of grapes and wine by near infrared spectroscopy. *Journal Near Infrared Spectrosc.* v. 14, p. 279-289, 2006.
- COZZOLINO, D.; CYNKAR, W.; SHAH, N.; SMITH, P.A. Can spectroscopy geographically classify Sauvignon blanc wines from Australia and New Zealand? **Food Chemistry**, n. 126, p.673-678, 2011.
- COZZOLINO, D. Recent Trends on the Use of Infrared Spectroscopy to Trace and Authenticate Natural and Agricultural Food Products. **Applied Spectroscopy Reviews**, v.47, n.7, p.518-530, 2012.
- COZZOLINO, D. Use of infrared spectroscopy for in field measurement and phenotyping of plant properties: instrumentation, data analysis and examples. **Applied Spectroscopy Reviews**, 2014.
- DANTAS, H.J.; VILAR, F.A.; SILVA, F. L. H.SILVA, A.S. Avaliação da influência da velocidade de destilação na análise físico-química de aguardente de cana-de-açúcar. **Revista Brasileira de Produtos Agroindustriais**, Campina Grande, v.9, n.2, p.101-109, 2007.

DASZYKOWSKI, M., KACZMAREK, K., HEYDEN, Y.V., WALCZAK, B. Robust statistics in data analysis - A review: basic concepts. **Chemometrics And Intelligent Laboratory Systems**, v. 85, n.2, p.203-219, 2007.

DIESEL, V. et al. Trabalho apresentado no **I Congresso Internacional de Desenvolvimento Rural e Agroindústria Familiar**.

FERREIRA, A.L.A., ANTUNUS, A.M., MELGO, M.S., VOLPE, P.L.O. Quimiometria I: calibração multivariada, um tutorial, **Química Nova**, v. 22, n.5, p. 724-731. 1999.

GALTIER O., DUPUY N., LE DRÉAU Y., OLLIVER D., PINATEL C., KISTER J., ARTAUD J.. **Geographic origins and compositions of virgin olive oils determined by chemometric analysis of NIR spectra**. *Analytica Chimica Acta*, v.595, p.136–144, 2007.

GALLIGNANI, M.; GARRIGUES, S.; LA GUARDIA, M. Direct determination of ethanol in all types of alcoholic beverages by near-infrared derivative spectrometry. **Analytica Chimica Acta**, v. 118, p. 1167-1173, 1993.

GY, B., OVÁDI, Z., SALGÓ, A. Prediction of Gasoline Properties with Near Infrared Spectroscopy. **Journal Near Infrared Spectroscopy**, v. 6, p.341–348, 1998.

HERSCHEL, W. Experiments on the Refrangibility of the invisible Rays of the Sun. **Philosophical Transactions of the Royal Society Of London**, v. 90, p. 284-292, 1800.

IBRAC. Mercado interno. 2008. Disponível em: < <http://www.ibrac.net/index.php/servicos/estatisticas/mercado-interno> > Acesso em: 11 de maio de 2015.

IBRAC. Mercado externo. 2015. Disponível em:< <http://www.ibrac.net/index.php/servicos/estatisticas/mercado-externo> > Acesso em: 11 de maio de 2015.

INPI. Registros - indicação geográfica. 2015. Disponível em: http://www.inpi.gov.br/portal/artigo/registros_indicacao_geografica. Acesso em: 11 de maio de 2015.

JERÔNIMO, E., M., SILVA, M., A.. Cachaça: uma bebida brasileira. **Pesquisa & Tecnologia**, vol. 2, n.2, 2005.

KOLOMIETS, O.A.; LACHENMEIER D.W.; HOFFMANN U.; SIESLER W. Quantitative determination of quality parameters and authentication of vodka using near infrared spectroscopy. **Journal of Near Infrared Spectroscopy**, v. **18**, p.59–67, 2010.

LIU, L., COZZOLINO, D., CYNKAR,W.U., DAMBERGS, R.G., JANIK, L., O'NEILL, B.K., COLBY, C.B., AND GISHEN, M. Geographic classification of Spanish and Australian Tempranillo red wines by visible and near infrared spectroscopy combined with multivariate analysis. **Food Chem.**, v. 106:781, 2006.

LIU, L.; COZZOLINO, D.; CYNKAR W.U.; DAMBERGS R.G.; JANIK L.; O'NEILL B.K.; COLBY C.B.; GISHEN M.. Preliminary study on the application of visible–near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. **Food Chemistry**, v.106, p. 781-786, 2008.

LLÉTI, R.; MELÉNDEZ E.; ORTIZ M.C.; SARABIA L.A.; SÁNCHEZ, M.S. Outliers in partial least squares regression.Application to calibration of wine

- grade with mean infrared data. **Analytica Chimica Acta**, v. 544, p. 60–70, 2005.
- MARTELO-VIDAL, M.J.; VÁZQUEZ, M. Evaluation of Ultraviolet, Visible, and Near Infrared Spectroscopy for the Analysis of Wine Compounds. **Czech Journal of Food Sciences**. v.32, n.1, pp.37-47, 2014.
- MARTENS, H., NAES, T. **Multivariate Calibration**. New York: John Wiley & Sons, 1996.
- MARTINS, J. **Um pouco da história**. IBRAC, 2011. Disponível em: <http://www.ibrac.net/index.php/ig-cachaca/historia-da-cachaca>. Acesso em: 30 de janeiro de 2015.
- MATTIUZ, B.; BISCEGLI, C. I.; DURIGAN, J. F.. Aplicações da tomografia de ressonância magnética nuclear como método não-destrutivo para avaliar os efeitos de injúrias mecânicas em goiabas 'Paluma' e 'Pedro Sato'. **Rev. Bras. Frutic.**, Jaboticabal , v. 24, n. 3, 2002
- MENDES, L.S.; OLIVEIRA, F.C.C.; SUAREZ, P.A.Z., RUBIM, J.C. Determination of ethanol in fuel ethanol and beverages by Fourier transform (FT)-near infrared and FT-Raman spectrometries. **Analytica Chimica Acta** v.493, p.219–231, 2003.
- MINÁ, A.J.S.; OLIVEIRA, F.G.R.; CALIL Jr., C.; DIAS, A.A.; SALES, A. Avaliação não destrutiva de postes de madeira por meio de ultra-som. **Scientia Forestalis**, Piracicaba, v.1, n.65, p.188-189, 2004.
- MIRANDA, M.B.; MARTINS, N.G.; BELLUCO, A.E.; HORII, J.; ALCARD, A.R. Qualidade química de cachaças e de aguardentes brasileiras. **Ciênc. Tecnol. Aliment.**, Campinas , v. 27, n. 4, 2007 .

- MORGANO, M.A.; FARIA, C. G.; FERRÃO, M.F.; BRAGAGNOLO, N.; FERREIRA, M.C.F. Determinação de proteína em café cru por espectroscopia NIR e regressão PLS. **Ciência e Tecnologia de Alimentos**, v. 25, n.1. p. 25-31, 2005.
- NICOLAI, B. M.; BEULLENS, K.; BOBELYN, E.; PEIRS, A.; SAEYS, W.; THERON, K. I. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. **Postharvest Biology and Technology**, v. 46, p.99-118, 2007.
- NETO, B.B.; SCARMINIO, I.S.; BRUNS, R.I. 25 anos de quimiometria no Brasil. *Química nova*, v.29, n6, p. 1401-1406, 2006.
- NORDON, A.; MILLS, A.; BURN, R.T.; CUSICK, F.M.; LITTLEJOHN, D..Comparison of non-invasive NIR and Raman spectrometries for determination of alcohol content of spirits. **Analytica Chimica Acta**, v.548, p. 148–158, 2005.
- ODELLO, L.; BRACESCHI, G.P.; SEIXAS, F.R.F.; SILVA, A.A.; GALINARO, C, A.; FRANCO, D.W. Avaliação sensorial de cachaça. **Química Nova**, v. 32, n. 7, p.1839-1844, 2009 .
- PASQUINI, C. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. **Journal of Brazilian Chemistry Society**, v.14, n.2, p.198-219, 2003.
- PINHEIRO, P.C.; LEAL M.C.; ARAÚJO, D.A.; Origem, produção e composição química da cachaça. **Química nova da escola**, n. 18, 2003.
- PRIETO, N., ROEHE, R., LAV'IN, P., BATTEN, G., AND ANDR'ES, S. Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review. **Meat. Sci.**, v.83, p.175–186, 2009.

- REBOUÇAS, M.V., BARROS NETO, B. Near infrared spectroscopic prediction of physical properties of aromatic-rich hydrocarbon mixtures. **Journal Near Infrared Spectroscopy**, v.9, p.263-273, 2001.
- RIBEIRO, F.A.L.; BARBOZA, F.D.; BREITKREITZ, M.C.; MARTINS, J.A. Quimiometria: inovação no desenvolvimento e validação de métodos analíticos para controle de qualidade na indústria química. 2007. Disponível em: http://www.crq4.org.br/informativomat_235. Acesso em: 11 de maio de 2015.
- RODRIGUES, I.C.; RODRIGUES, A. M.; ALVES, F.J.C. Produtos típicos e a diferenciação de produtos: a iniciativa da “Cachaça de Minas”. **XXI Encontro Nacional de Engenharia de Produção**, 2001.
- RODRIGUES, B.U.; COSTA, R.M.; SALVINI, R.L.; SOARES, A.S.; SILVA, F.A.; CALIAR, M.; CARDOSO, K.C.R.; RIBEIRO, T.I.M. Cachaça Classification Using Chemical Features and Computer Vision. **Procedia Computer Science** v.29, p. 2024–2033, 2014.
- ROTA, M.B.; FARIA, J.B. Efeito do processo de bidestilação na qualidade sensorial da cachaça. **Alimentos e Nutrição**, v.20, n.1, p. 121-127, 2009.
- SANTOS, A.F., SILVA, F.M., LENZI, M.K., PINTO, J.C. Monitoring And Control of Polymerization Reactors Using NIRS, **Polymer-Plastics Technology And Engineering**, v.44, p.1–61, 2005.
- SAVITSKY, A., GOLAY, M.J.E. Smoothing and differentiation by simplified least squares procedures. **Analytical Chemistry**, v.36, p.1627-1632, 1964.

- SCANAVINI, H. F. A.; CERIANI, R.; MEIRELLES, A. J. A.. Cachaça distillation investigated on the basis of model systems. **Braz. J. Chem. Eng.**, São Paulo, v. 29, n. 2, 2012.
- SEBRAE. **Potencial da cachaça**. Boletim, 2014.
- SHETTY, N.; GISLUM, R. Quantification of fructan concentration in grasses using NIR spectroscopy and PLSR. **Field Crops Research**. V. 120, p. 31-37, 2011.
- SILVA, M.C.V.G.; MACIEL, D.S.C.; FREITAS, L.A. Avaliação do processo produtivo de cachaça de alambique do engenho Pé da Serra à luz da produção mais limpa. XVI SemeAD, Seminários da Administração, Campina Grande, 2013.
- SILVA, C. L.C.; VIANNA, C. R.; CADETE, R. M.; SANTOS, R.O.; GOMES, F.C.O.; OLIVEIRA, E.S. e ROSA, C.A. Selection, growth, and chemo-sensory evaluation of flocculent starter culture strains of *Saccharomyces cerevisiae* in the large-scale production of traditional Brazilian cachaça. **International Journal of Food Microbiology**, v. 131, pp. 203-210, 2009.
- SIMÕES, SILVA F.S., Potencialidades da Espectroscopia de infravermelho com Transformadas de Fourier na caracterização de aguardentes vínicas. **Dissertação de mestrado**. Universidade de Aveiro, 2008.
- SORATTO, A.N.; VARVAKIS, G.; HORII, J.. A certificação agregando valor à cachaça do Brasil. **Ciênc. Tecnol. Aliment.**, v. 27, n. 4, p. 681-687, 2007
- SOUZA, A.M.; POPPI, R.J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial, parte I. **Química Nova**, v. 35, n. 1, pp. 223-229, 2012

- SOUZA, A.M.; BREITKREITZ, M.C.; FILGUEIRAS, P.R.; ROHWEDDER, J.J.R.; POPPI, R.J. Experimento didático de quimiometria para calibração multivariada na determinação de paracetamol em comprimidos comerciais utilizando espectroscopia no infravermelho próximo: um tutorial, parte II. **Quim. Nova**, v. 36, n. 7, pp. 1057-1065, 2013
- SOUZA, P. P.; RESENDE, A.M.M.; AUGUSTI, D.V.; BADOTTI, F.; CASSIA, F.; GOMES, R.R.C.; EBERLIN, M.N.; AUGUSTI, R. Artificially-aged cachaça samples characterised by direct infusion electrospray ionisation mass spectrometry. **Food Chemistry**, v. 143, pp.77–81, 2014.
- SUSI, H. e BYLER, D. M. Biochem. Biophys. **Res. Commun.** 115, 391-397, 1983.
- NAES, T.; ISAKSSON, T.; FEARN T.; DAVIES, T. *A User-friendly Guide to Multivariate Calibration and Classification*, **NIR Publications**, Chichester, UK, 2002.
- TREGEAR, A.; KUZNESO, S.; MOXEY, A.. Policy initiatives for regional foods: some insights from consumer research. **Food Policy**, v. 23, n. 5, p. 383–394, 1998.
- URBANO-QUADRADO, M.; LUQUE de CASTRO, M.D.; PEREZ, P.M.; GÓMEZ NIETO, M.A. Near infrared reflectance spectroscopy and multivariate analysis in enology Determination or screening of fifteen parameters in different types of wines. **Analytica Chimica Acta**, v.527, pp.81-88, 2004.
- URTUBIA, A.; CORREA, J.R.P.; MEURENS, M.; AGOSIN, E.; Monitoring large scale wine fermentations with infrared spectroscopy. *Talanta* , v. 64, p. 778–784, 2004.

- VALENTE, M.E.; PEREZ, R.; RAMOS, A.M.; CHAVES, J.B. Indicação geográfica de alimentos e bebidas no Brasil e na União Europeia. **Ciência Rural**, v.42, n.3, 2012.
- VAN DER LANS, I.; ITTERSUM, K.V.; CICCIO, A.; LOSEBY, M. The role of the region of origin and EU certificates of origin in consumer evaluation of food products. **European Review of Agricultural Economics**, v.28, n.4, p. 451-477, 2001.
- VITALE R., BEVILACQUA M., BUCCI R., MAGRI A.D., MAGRI A.L., MARINI, F. A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 121, p.90–99, 2013.
- XIE, L.; YING, Y.; YING, T. Classification of tomatoes with different genotypes by visible and short-wave near-infrared spectroscopy with least-squares support vector machines and other chemometrics. **Journal of Food Engineering**, v.94, p. 34-39, 2009.
- WANG, M.; CHOONG, Y.; SU, N; LEE, M. A rapid method for determination of ethanol in alcoholic beverages using capillary gas chromatography. **Journal of Food and Drug Analysis**, v. 11, n. 2, p.133-140, 2003.
- J.W. ZHAO, J.W.; CHEN, Q.S.; HUANG, X.Y.; FANG, C.H. Qualitative identification of tea categories by near infrared spectroscopy and support vector machine. **Journal of the Pharmaceutical and Biomedical Analysis** v. 41 p. 1198–1204, 2006.

CAPÍTULO II

Uso da espectroscopia do infravermelho próximo (NIR) como método não destrutivo para a discriminação de aguardentes de cana-de-açúcar produzidas nos estados de São Paulo e Minas Gerais

RESUMO

A aguardente de cana-de-açúcar é uma bebida alcoólica produzida a partir da destilação do vinho obtido da fermentação do mosto de cana-de-açúcar, sendo considerada uma bebida tipicamente brasileira. Como muitos produtos, a valorização da aguardente de cana-de-açúcar pode ser atribuída a aspectos da notoriedade que ela possui, principalmente aqueles relacionados com o local de produção. Como a espectroscopia do infravermelho próximo (NIR) vem sendo utilizada para discriminar bebidas em função da sua região produtora, o objetivo deste trabalho foi desenvolver um método utilizando a espectroscopia NIR como método não destrutivo para atribuir autenticidade a aguardente de cana-de-açúcar produzidas nos estados de São Paulo em comparação com as produzidas no estado de Minas Gerais. Foram utilizados 139 amostras de aguardentes de cana-de-açúcar produzidas nos estados de Minas Gerais e São Paulo. Estas foram divididas em dois grupos, o conjunto de calibração ($n = 100$) e conjunto de previsão ($n = 39$) usando o algoritmo clássico Kernnard-Stone. De cada amostra coletou-se 5 sub-amostras de aproximadamente 5 mL. Os espectros NIR das sub-amostras foram obtidos na região do NIR (4.000 a 10.000 cm^{-1}) como $\log 1/R$ com 64 scans a uma resolução espectral de 16 cm^{-1} , com intervalo de 8 cm^{-1} , utilizando um

espectrômetro FT-IR Spectrum 100N. A classificação das aguardentes de cana-de-açúcar em função dos estados foi feita empregando a análise de componentes principais (PCA), análise discriminante por regressão de mínimos quadrados parciais (PLS-DA), análise linear discriminante (LDA) e máquina de vetores de suporte (SVM). Os resultados da PCA com espectros brutos e pré-processados não foram eficazes para classificar as amostras entre os estados. Por outro lado, quando foram utilizados PLS-DA e LDA, foi possível obter uma predição correta de 87,2% com espectros pré-processados com a correção do espalhamento multiplicativo de sinal (MSC). O modelo SVM apresentou 87,2% de acerto na predição, porém os espectros não foram pré-processados. Deste modo, o emprego da espectroscopia NIR e análises quimiométricas podem ser usados como método não destrutivo para atribuir autenticidade de aguardentes de cana-de-açúcar.

ABSTRACT

Sugarcane aguardente is an alcoholic beverage produced by the distillation of the wine obtained from the fermentation of sugarcane most and it is considered a typically Brazilian spirit. Like many products, the appreciation of sugarcane aguardente can be attributed to aspects of the notoriety, especially those related to the production origin. As near-infrared spectroscopy (NIR) has been used to discriminate alcoholic beverages according to their production region, the aim of this study was to develop a non-destructive method using NIR spectroscopy for authenticity sugarcane aguardente produced in São Paulo, compared with those produced in Minas Gerais. A total of 139 samples were used and these samples were divided into two groups, the calibration set ($n = 100$) and prediction set ($n = 39$) using the classical algorithm Kernard-Stone. From each sample it was collected 5 sub-samples with approximately 5 mL. The NIR spectra of the sub-samples was obtained in the NIR region ($4000-10000\text{ cm}^{-1}$) using an FT-IR spectrometer Spectrum 100N reflectance as $\log 1/R$ with 64 scans at a spectral resolution of 16 cm^{-1} with interval of 8 cm^{-1} , totaling five spectra of each sample. The classification of sugarcane aguardente in function of the state was done using the principal component analysis (PCA), discriminant regression analysis by partial least squares (PLS-DA), linear discriminant analysis (LDA) and support vector machine (SVM). The results of PCA with pre-processed and raw NIR spectra were not effective to classify sugarcane aguardente between the states. On the other hand, were used PLS-DA and LDA was possible to obtain a correct prediction of 87.2% using pre-processed spectra with signal multiplicative scattering (MSC). The SVM model also showed 87.2% accuracy in predicting, but the spectra were not

preprocessed. Therefore, the use of NIR spectroscopy and chemometric analysis can be used as a non-destructive method to attribute authenticity to sugarcane aguardentes.

1. INTRODUÇÃO

A bebida alcoólica chamada de cachaça é uma denominação típica e exclusiva da aguardente de cana-de-açúcar brasileira, com graduação alcoólica variando de 38 a 48% em volume, a 20°C, obtida pela destilação do mosto fermentado do caldo de cana-de-açúcar com características sensoriais peculiares, podendo ser adicionada ou não de açúcares em até 6 g.L⁻¹ (BRASIL, 2009). O destilado de cana-de-açúcar para ser considerado cachaça, deve obedecer aos parâmetros estabelecidos pela Instrução Normativa número 13, de 30 de junho de 2005 do Ministério de Agricultura, Pecuária e Abastecimento (MAPA), que estabelece os Padrões de Identidade e Qualidade (PIQ) para aguardente de cana e para a cachaça (BRASIL, 2005).

Deste modo, a cachaça deve ser produzida exclusivamente no Brasil, constituindo assim uma Indicação Geográfica (IG) nacional, (BRASIL, 2001) e o seu uso inapropriado por partes não autorizadas é prejudicial aos consumidores e aos legítimos produtores, sendo, no Brasil, considerado crime, conforme o descrito na Lei número 9279, de 14 de maio de 1996 (BRASIL, 1996).

Os principais estados produtores de aguardente de cana-de-açúcar são: São Paulo, Pernambuco, Ceará, Minas Gerais e Paraíba e os principais estados consumidores são: São Paulo, Pernambuco, Rio de Janeiro, Ceará, Bahia e Minas Gerais (IBRAC, 2008).

O estado de São Paulo destaca-se na produção de aguardente de cana-de-açúcar em coluna, enquanto o estado de Minas Gerais é o principal produtor de aguardente de cana-de-açúcar de alambique (VERDI, 2006), representando

25% deste mercado (BERNARDES et al., 2014). Na produção em alambique, a aguardente de cana-de-açúcar é dividida em três frações, ou seja: cabeça, coração e cauda, produzindo um resíduo da destilação conhecido como vinhaça, enquanto na destilação realizada em coluna não há separação do destilado em frações, isto explica em grande parte as diferenças na composição química da bebida produzida por estes dois processos (BECHE & FRANCO, 2009). Essas diferenças também podem correlacionar-se com a origem geográfica, a fatores climáticos, matéria prima, etapas do processo produtivo (INPI, 2014; DANTAS et al., 2007), e de tradições e hábitos durante o processo de produção.

Desta forma, técnicas para a determinação da origem geográfica de produtos agrícolas é altamente interessante para os consumidores, produtores rurais, revendedores e autoridades administrativas (COZZOLINO, 2012). As aguardentes de cana-de-açúcar, ou produtos, que tem indicação geográfica possuem notoriedade, valor intrínseco e identidade própria, além de distingui-los em relação aos seus similares disponíveis no mercado, apresentando uma qualidade única em função de recursos naturais como solo, vegetação, clima e processo de produção (MAPA, 2015).

Técnicas que permitem a determinação da origem geográfica de maneira rápida e com mínimo de preparo da amostra são interessantes para os órgãos fiscalizadores, produtores e consumidores. Assim a espectroscopia do infravermelho próximo (NIR) apresenta um grande potencial nas determinações de origem dos produtos (NICOLAÏ et al, 2007), sendo considerada uma técnica rápida, segura de operação, não destrutiva, relativamente barata e proporciona

automação dos processos de controle de qualidade nos produtos (PASQUINI, 2003).

Em bebidas alcoólicas, a espectroscopia NIR, juntamente com técnicas quimiométricas, como a Análise de Componentes Principais (PCA) ou análises discriminantes, vem sendo aplicada para a discriminação de vinhos produzidos em diferentes regiões geográficas (LIU et al., 2008; LLÉTI et al., 2005; BEVIN et al., 2006). Cozzolino et al. (2011) desenvolveram modelo de análise discriminante por mínimos quadrados parciais (PLS-DA), com espectros na região do infravermelho (400 – 2500 nm) para classificar vinhos da Austrália e Nova Zelândia, obtendo uma acurácia de 76% para a classificação geográfica desta bebida. Liu et al. (2006) correlacionaram as diferenças químicas do vinho com a sua região produtora, buscando comprovar a sua autenticidade e origem. Não há trabalhos relatando o uso da espectroscopia NIR para autenticação de aguardentes de cana-de-açúcar.

Assim, este trabalho deve por objetivo desenvolver uma metodologia utilizando a espectroscopia NIR e ferramentas quimiométricas de classificação, como as técnicas de PCA, Análise Linear Discriminante (LDA), PLS-DA e Máquina de Vetores de Suporte (SVM), como método não destrutivo para a atribuição da autenticidade de aguardentes de cana-de-açúcar produzidas no estado de São Paulo em comparação com as produzidas no estado de Minas Gerais.

2. MATERIAL E MÉTODOS

2.1. Obtenção das amostras

Segundo Pasquini (2003), para se produzir um modelo robusto, o mesmo deve incluir a variabilidade natural da propriedade de interesse e também um número mínimo de amostras empregadas na calibração (50 a 100, dependendo da complexidade e variabilidade do material). Desta forma, foram coletadas 139 amostras de aguardentes de cana-de-açúcar dos estados de São Paulo e Minas Gerais (Anexo 1). As amostras foram cedidas pelo Centro de Pesquisa e Desenvolvimento Qualidade da Cachaça da Universidade Estadual Paulista, Faculdade de Ciências Farmacêuticas (UNESP-FCFAR, n = 83), pelo laboratório de Análises Físico-Químicas de Aguardente (LAFQA) da Universidade Federal de Lavras (UFLA) (n = 49); e algumas foram compradas no comércio de Jaboticabal - SP (n = 7). Este conjunto de amostras foi dividido em dois grupos, ou seja, o conjunto de calibração (n = 100) e o conjunto de predição (n = 39), utilizando o algoritmo de seleção Kennard-Stone clássico (KENNARD & STONE, 1969).

As amostras de aguardente de cana-de-açúcar foram deixadas no laboratório visando à estabilização da temperatura (~25 °C). Após estabilizada a temperatura, cada amostra foi dividida em 5 sub-amostras de 5 mL para a obtenção dos espectros NIR. Em seguida, foi determinada a coloração das mesmas (Tabela 1)

Tabela 1. Estatística descritiva dos parâmetros de coloração (L*, a*, b*) das amostras de aguardente de cana-de-açúcar oriundas do estado de São Paulo e Minas Gerais.

Pop	Estado	N	Nota	Coloração				
				L*	a*	b*	C	Ângulo hue
Pop1	São Paulo (i)	69	1	47,41a	-0,36a	4,09a	4,11a	107,54a
Pop2	Minas Gerais (ii)	70	0	47,14a	-0,44a	4,60a	4,70a	102,20a
Test F				NS	NS	NS	NS	NS
*Cal	*i+ii	100		47,21b	-0,37b	4,48b	4,59b	102,65b
Pred.	*i+ii	39		47,42b	-0,48b	3,85b	3,94b	110,48c
Test F				NS	NS	NS	NS	*

Dividido por algoritmo de seleção Kernard Stone clássico. N - número de amostra. Nota classe utilizada no PLS-DA em função da região. Parâmetros da coloração (L, a*, b*, cromaticidade e ângulo hue). Pop: população. Test F: teste de significância. Cal: calibração. Pred: predição. Médias seguidas de pelo menos uma letra comum nas colunas, para cada elemento, não diferem entre si pelo teste de Tukey (P<0,05). ** = significativo a 1% de probabilidade; * = significativo a 5% de probabilidade pelo teste F. NS = não significativo.

2.2. Obtenção dos espectros NIR

De cada sub-amostra (aproximadamente 5 mL) foram obtidos espectros na região NIR utilizando um espectrômetro FT-IR Spectrum 100N (PerkinElmer, Shelton, EUA), na geometria óptica de reflectância como $\log 1/R$, na faixa de número de ondas de 4.000 a 10.000 cm^{-1} (1.000-2.500 nm) com 64 scans a uma resolução espectral de 16 cm^{-1} , com intervalo de 8 cm^{-1} , totalizando cinco espectros de cada amostra.

2.3. Determinação da coloração (CIE)

De cada amostra foram determinados parâmetros de coloração utilizando um reflectômetro Minolta CR 400 (Minolta Corp., Ozaka, Japão), que expressa este parâmetro segundo o sistema proposto pela “Commission Internationale de L’Eclairage” (CIE) em L^* , a^* , b^* . Os valores de L indicam a luminosidade e varia de 100 (branco) a 0 (preto), o parâmetro a^* é a direção do verde ($-a^*$) ao vermelho ($+a^*$) e b^* do azul ($+b^*$) ao amarelo ($-b^*$), McGuire (1992).

Além destes valores, foram calculados os parâmetros ângulo hue $\{^\circ h = [\arctan(b^*/a^*)]\}$ e cromaticidade $\{C^* = \sqrt{[(a^*)^2 + (b^*)^2]} \times 0,5\}$, segundo a metodologia descrita por McGuire (1992).

2.4. Análises quimiométricas

Os cinco espectros coletados de cada amostra foram reduzidos à sua média e as informações presentes nos mesmos foram extraídas utilizando o programa Unscrambler®, versão X.3 (CAMO, Oslo, Noruega), assim como os pré-processamentos e a construção dos modelos de classificação quanto ao estado de origem.

2.4.1. Pré-processamento

O uso de pré-processamento nos espectros foi utilizado para remover qualquer informação irrelevante que não poderia ser tratada por meio das técnicas de regressão (COZZOLINO, 2011). Empregaram-se os seguintes pré-processamentos nos espectros:

- Multiplicative Scatter Correction ou correção do espalhamento multiplicativo de sinal – MSC (ISAKSSON & TORMOD, 1988). O MSC foi

utilizado para compensar o espalhamento de luz, que normalmente é provocado pela falta de homogeneidade ótica das amostras. O MSC visa minimizar a interferência relativa ao espalhamento de luz com base no espalhamento médio de todos os espectros empregados na calibração multivariada (MARTENS & NAES, 1996).

- Segunda Derivada de Savitsky-Golay - 2ª D SG (SAVITSKY & GOLAY, 1964). A segunda derivada proporciona uma forma direta para identificar picos que possam conter alguma informação importante, assim, permite um estudo mais detalhado, qualitativo e, eventualmente, estudos quantitativos dos espectros (SUSI & BYLER, 1983).

Os pré-tratamentos foram avaliados pela capacidade de predizer as amostras de Minas Gerais e São Paulo.

2.4.2. Análise multivariada: classificação

A calibração multivariada tem como ideia básica relacionar duas matrizes de dados (**X** e **Y**), em que **X** (n, m) é o conjunto dos dados experimentais contendo as variáveis independentes, representado em suas linhas pelas amostras do conjunto e em suas colunas pelas variáveis medidas e o outro bloco (**Y**), formado pela(s) variável(eis) dependente(s). Dessa forma, na calibração multivariada, ao contrário do que ocorre na univariada, utiliza-se de múltiplas respostas para correlacionar a(s) propriedade(s) medida(s) das amostras.

2.4.2.1 Análise de Componentes Principais (PCA)

A análise de componentes principais é amplamente utilizada na quimiometria, pois é a base para diversos métodos de reconhecimento de padrões, classificação e calibração multivariada, usualmente a PCA é utilizada

com o objetivo de visualizar a estrutura dos dados, encontrar similaridades entre as amostras e detectar amostras anômalas (outliers) (SOUZA & POPPI, 2012).

Os espectros NIR brutos, sem pré-processamento, e os pré-processados foram submetidos à análise de componentes principais (PCA). Nos dados de coloração (L^* , a^* e b^*) foi feita obtendo a média e a padronização do conjunto de dados, de modo que cada variável mantivesse a média nula e variância unitária (HARTINGAN, 1975).

2.4.2.1. Técnicas Supervisionadas de Classificação

Ao contrário das técnicas quimiométricas não supervisionadas, como o PCA, nas técnicas supervisionadas é necessário que exista alguma informação inicial sobre a identidade das amostras para a formação das classes e o objetivo é desenvolver um modelo baseado nas informações contidas nas amostras (BRERETON, 2007). Como técnica supervisionada utilizou-se: análise discriminante por mínimos quadrados parciais (PLS-DA), análise linear discriminante (LDA) e máquinas de vetores de suporte (SVM).

A análise PLS-DA maximiza a separação entre classes pré-definidas, São Paulo e Minas Gerais, em vez de explicar as variações dentro de um conjunto geral de dados (WONG et al., 2013). Os modelos PLS-DA foram desenvolvidos com base nos algoritmos de mínimos quadrados parciais (PLS), onde as variáveis da matriz X (dados espectrais) foram relacionadas com as classes determinadas, São Paulo e Minas Gerais, no vetor Y (MAZIVILA et al., 2015), sendo atribuídos aos modelos desenvolvidos um valor de referência para cada amostra.

Esta técnica de análise discriminante é recomendada em dados de alta dimensionalidade (NAES et al., 2002). No entanto, como o PLS-DA explica as diferenças entre propriedades gerais das classes, a interpretação pode se tornar complicada quando o número de classes aumenta (BYLESJO et al., 2006). Neste trabalho, foram utilizadas duas classes, onde para cada amostra foi atribuído um valor de 0 ou 1 (Tabela 1), indicando se a amostra pertence ao estado de São Paulo (1) ou ao estado de Minas Gerais (0). Adotou-se como critério de separação o valor de 0,5 (ponto de corte) como indicado por Cunha Júnior et al (2015), onde amostras classificadas com valores menores que 0,5 foram consideradas como sendo de Minas Gerais (classe 0) e as com valores maiores que 0,5 como sendo originárias de São Paulo (classe 1).

O LDA foi utilizado para classificar as amostras de aguardente de cana-de-açúcar proveniente do estado de Minas Gerais e de São Paulo. A discriminação dos grupos consiste no cálculo da distância entre as amostras, para tal, foi utilizada a medida de Mahalanobis, sendo que quanto maior a distância Mahalanobis entre duas amostras, maior as diferenças espectrais entre os mesmos (MIRALBÉS, 2008). Os modelos LDA foram construídos utilizando PCA, com os 10 primeiros componentes principais (PCs) resultantes da PCA, este número de PCs foi escolhido, por contemplarem mais que 98 % da variância total dos dados.

O SVM é aplicado aos espectros NIR para identificar ou classificar produtos, para solucionar problemas de calibração não linear em função do mapeamento do espaço original dos dados em um espaço de alta dimensão, denominado de espaço de características (CORTES & VAPNIK, 1995). Assim como o PLS-DA, o SVM foi utilizado para classificar as amostras do presente

estudo, em função de sua origem, utilizou-se o SVM tipo 1 ou C-SVC (classificação de vetores de suporte), função kernel linear e este tipo de SVM tem um parâmetro de entrada chamado C, que é um fator de capacidade ou de grande penalidade, uma medida da solidez do modelo, como o C deve ser maior do que 0 (HSU et al., 2009), adotou-se um C com valor 5.

Os modelos construídos para a classificação foram avaliados quanto a porcentagem de acertos na classificação, conforme descrito por Cunha Júnior (2015) :

$$\text{Classificação correta (\%)} = 100 - \frac{\{\sum y^1 + \sum y^2\}}{N} \times 100$$

Onde y^1 e y^2 representam o número de amostras classificadas incorretamente para cada grupo (SP e MG), respectivamente, e N é o número de amostras total. Os modelos PLS-DA, LDA e SVM foram desenvolvidos usando validação cruzada total (full cross validation).

2.5. Análise Univariada

Os parâmetros de coloração (L^* , a^* , b^* , cromaticidade e ângulo hue) foram submetidos à análise de variância (ANOVA). A média das variáveis de coloração para cada estado (Minas Gerais e São Paulo) e também para os conjuntos de calibração e validação foram analisados individualmente utilizando-se o teste de Tukey, a 5% de significância. A análise estatística foi realizada utilizando o programa Statistica, versão 7.0 (STATSOFT, 2004).

3. RESULTADOS E DISCUSSÃO

3.1. Coloração

Pode ser observado na Tabela 1 que não houve diferença significativa ($p < 0,05$) para os parâmetros de coloração L^* , a^* , b^* , cromaticidade e ângulo hue entre amostras de aguardentes de cana-de-açúcar produzidas nos estados de São Paulo (Pop 1) e Minas Gerais (Pop 2), ou seja, as amostras não apresentaram diferenças significativas quanto a sua coloração. Por outro lado, entre o grupo de calibração e predição houve diferença significativa ($p < 0,05$) para o ângulo hue, com o conjunto de calibração apresentando valores menores (102,65) em relação ao de predição (110,48), Tabela 1.

A PCA foi utilizada como análise exploratória para visualizar a existência ou não de agrupamentos de amostras de aguardente em função da coloração. Na Figura 1 pode-se observar os resultados do PCA para os dados de coloração (L^* , a^* e b^*). O primeiro componente principal (PC1) correspondeu a 77,17% da variância explicada e o segundo componente principal (PC2) 19,99%. A soma destes dois PCs gerou uma explicação de 97,16 % do total da variância do banco de dados.

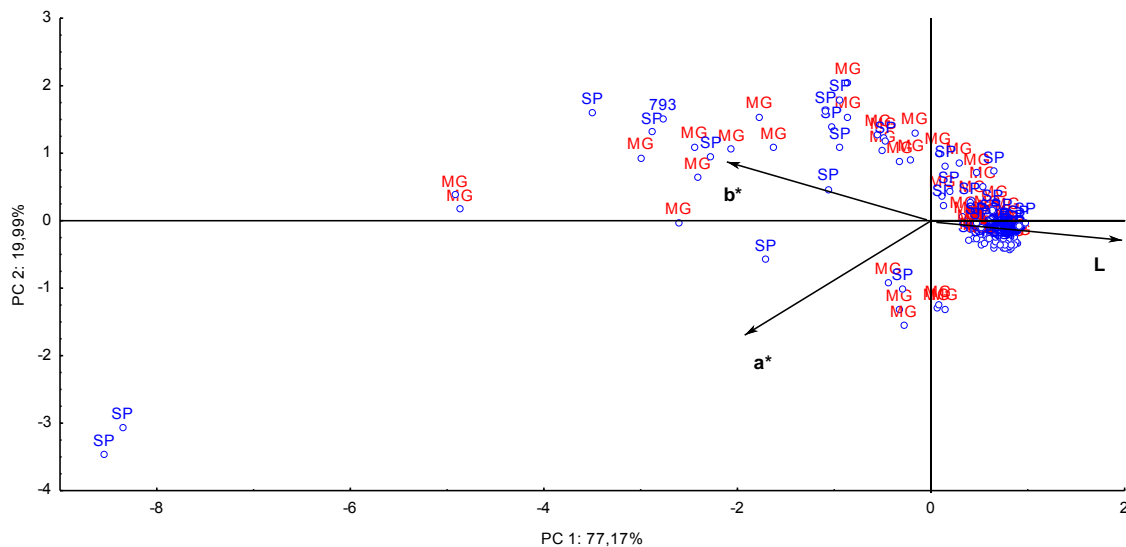


Figura 1. Distribuição das aguardentes de cana-de-açúcar produzidas no estado de São Paulo (SP) e Minas Gerais (MG) de acordo com os componentes principais, PC1, 77,17% de variância explicada e PC2, 19,99% de variância explicada e os vetores dos parâmetros de coloração (L, a* e b*).

Percebe-se a formação de um grupo de amostras ao lado direito do PC1 que foi constituído por amostras com valores de luminosidade (L*) mais elevados, correspondendo às amostras mais transparentes (Figura 1). Contudo, não foi observado agrupamento das amostras em função dos estados de origem (São Paulo e Minas Gerais), sendo observada uma sobreposição das mesmas indicando que a coloração foi bastante semelhante (Figura 1).

Por outro lado, percebe-se a presença de 26 amostras posicionadas no lado positivo do PC2, tendo sido influenciadas pelo parâmetro b*, e 9 posicionadas no lado negativo do PC2, influenciadas pelo parâmetro a* (Figura 1). Estas amostras constituem parte do universo amostral que apresentava coloração mais escura, possivelmente oriunda do processo de envelhecimento e/ou adição de aditivos (Figura 2).



Figura 2. Amostras de aguardente de cana-de-açúcar e seus respectivos valores de luminosidade (L*). *número da amostra conforme o descrito no Anexo 1.

A diferença de coloração entre as amostras pode ser devido ao processo de envelhecimento, em tonéis de madeiras, a que possivelmente elas foram submetidas, pois nesta etapa a bebida pode extrair compostos orgânicos coloridos da madeira (ODELLO et al., 2009). Além disso, o tempo de envelhecimento também pode promover um aumento significativo da coloração amarela, como demonstrado por Cardello e Faria (1998) e Bernardi et al. (2012).

Ao se analisar as amostras (Anexo 1), é possível verificar que as que apresentavam menor luminosidade (L*), Figura 2, ou seja, as mais escuras, apresentaram teor alcoólico inferior (38,3 %) quando comparadas a amostra 80 (44,1 %), esta diferença pode estar relacionada a prática do envelhecimento, que além de influenciar a coloração da bebida, pode reduzir o seu teor alcoólico (ODELLO et al., 2009).

3.2. Espectros NIR

Na Figura 3 podem ser observados os espectros NIR originais, sem pré-processamento, das 139 amostras de aguardente de cana-de-açúcar. Foram identificados picos nas faixas de 1.420 – 1.500 nm e 1.900 – 2070 nm, que estão relacionados ao primeiro sobreton de OH e a presença de moléculas de água, respectivamente (INÁCIO et al., 2013). Também é possível identificar picos nos comprimentos de ondas 1700 - 1800 nm, 2.280 – 2.370 nm e 2400 nm, que estão relacionados ao estiramento OH do etanol e metanol (FERNANDES et al., 2008), às combinações de CH e estiramentos de OH, e a presença de aromáticos e ao grupo CH₃ presente no metanol (COZZOLINO et al., 2011), respectivamente.

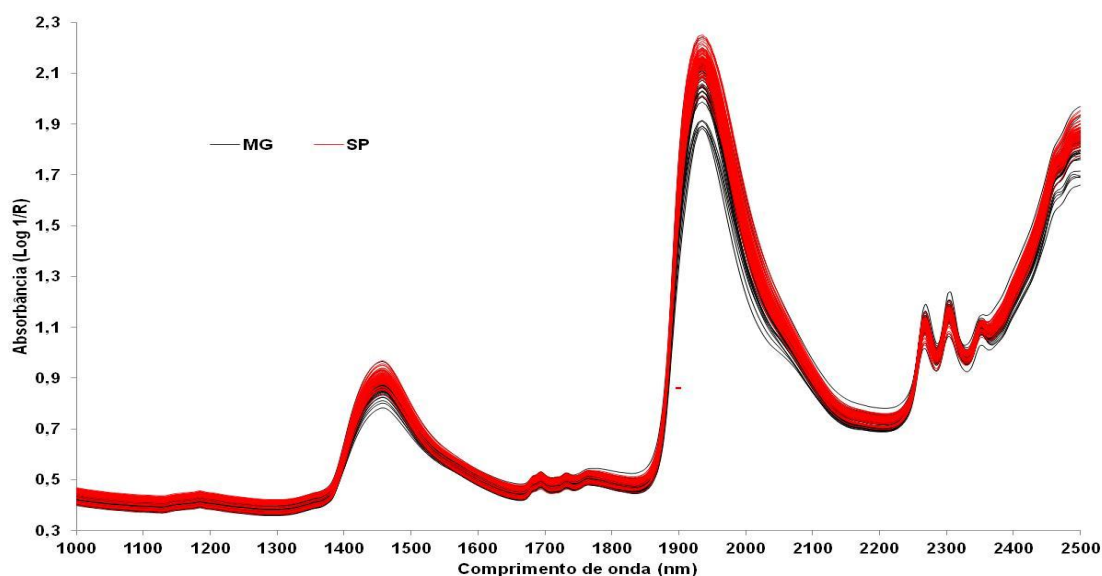


Figura 3. Espectros NIR das amostras de aguardente de cana-de-açúcar e cachaça provenientes dos estados de São Paulo e Minas Gerais.

Não foi possível observar diferenças espectrais entre os estados (Figura 3), com os espectros sendo muito semelhantes entre as amostras. Porém, foi

observado um pequeno efeito do espalhamento de luz, como pode ser visto na Figura 3. Este é um fenômeno que normalmente é ocasionado pela falta de homogeneidade ótica das amostras (MORGANO et al., 2005).

Com o objetivo de corrigir o efeito do espalhamento de luz presente nos espectros NIR, foi utilizado a correção do espalhamento multiplicativo de sinal (MSC). Este é minimizado com base no espalhamento médio de todos os espectros formadores na matriz de dados X (MARTENS & NAES, 1996). A correção MSC é semelhante à correção SNV, contudo, o MSC utiliza um modelo de regressão para corrigir os espectros (SABIN et al., 2004), enquanto que o SNV utiliza o desvio padrão para correção. Pode-se observar que os espectros NIR após a utilização do MSC (Figura 4) ficaram menos dispersos quando comparados aos espectros brutos (Figura 3).

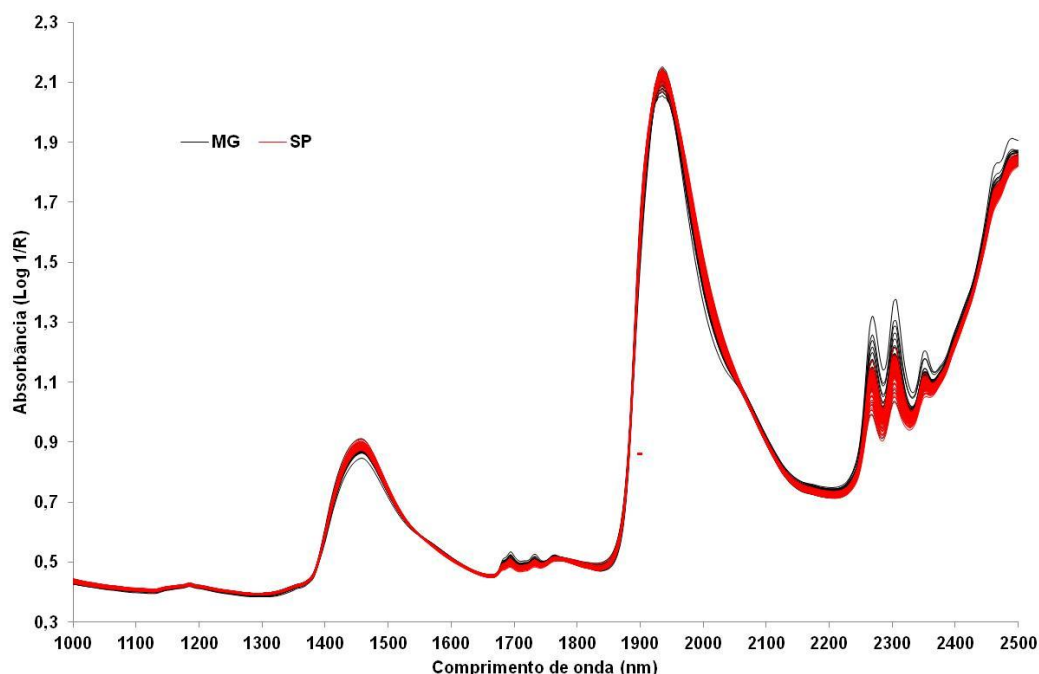


Figura 4. Espectros na região do infravermelho próximo (NIR) das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais após a utilização da correção do espalhamento multiplicativo de sinal (MSC).

Além do MSC, foi utilizada a segunda derivada de Savitsky-Golay (2ªD S-G) com um filtro de 4 pontos de suavização. A 2ªD S-G tem por objetivo remover as variações da linha de base entre as amostras, esse processo faz com que as informações contidas ao longo dos diferentes comprimentos de onda sejam acentuadas (Figura 5). Com a segunda derivada foi possível eliminar a variação linear da linha de base, normalmente devido ao efeito do espalhamento de luz.

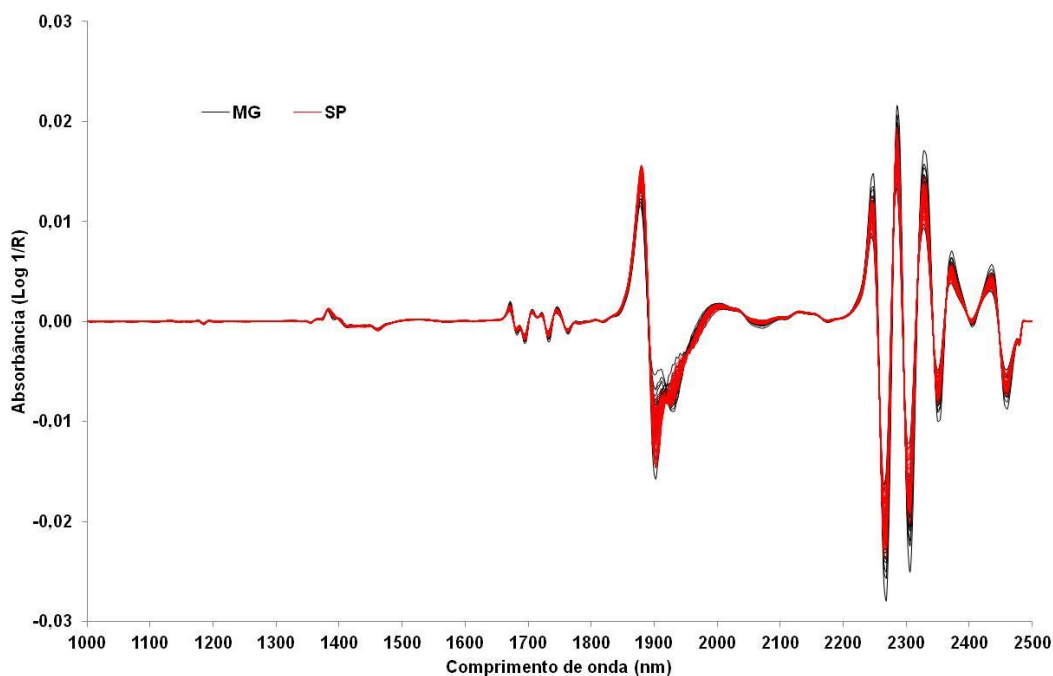


Figura 5. Espectros na região do infravermelho próximo (NIR) das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais após a transformação utilizando a segunda derivada de Savitsky-Golay (2ªD S-G) com janela de 4+4 pontos.

Ao se analisar os espectros NIR, não foi possível observar diferenças entre as amostras oriundas do estado de São Paulo e de Minas Gerais.

Portanto, os espectros foram submetidos a diferentes análises multivariadas no sentido de se obter a separação das amostras.

3.2.1. Análise por componentes principais - PCA (técnica não supervisionada)

A PCA foi empregada para observar diferenças espectrais entre as amostras produzidas em São Paulo e Minas Gerais. Com o objetivo de reduzir a dimensionalidade dos dados em um pequeno número de componentes principais, facilitando a observação de um possível agrupamento de amostras em função da sua origem geográfica ou a detecção de amostras anômalas (COZZOLINO et al., 2011), ou seja, que não pertence a nenhuma das categorias conhecidas (São Paulo e Minas Gerais).

Na Figura 6 pode-se observar os escores das duas primeiras componentes principais (PCs) obtidos com os espectros originais (brutos ou seja, sem pré-processamento).

Apesar das amostras de São Paulo se concentrarem à direita do PC1, não houve uma separação muito clara em relação às amostras de Minas Gerais, havendo algumas (17 amostras ou 24,6%) que se misturaram às daquele estado, estando concentradas à esquerda do PC1.

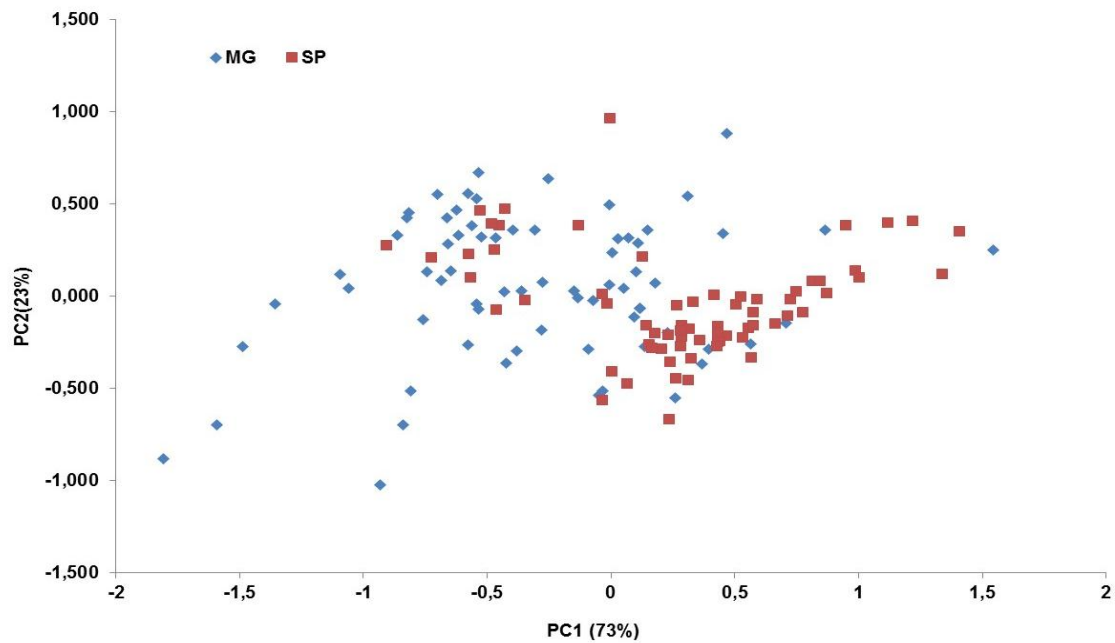


Figura 6. Escores do PC1 (73% de variância) e PC2 (23% de variância) obtidos com os espectros NIR (1000-2500 nm) sem pré-processamento das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais.

Na tentativa de se melhorar separação entre os grupos, foi utilizado os pré-processamentos MSC (Figura 7A) e a segunda derivada de Savitsky-Golay (2^aD S-G), Figura 7B.

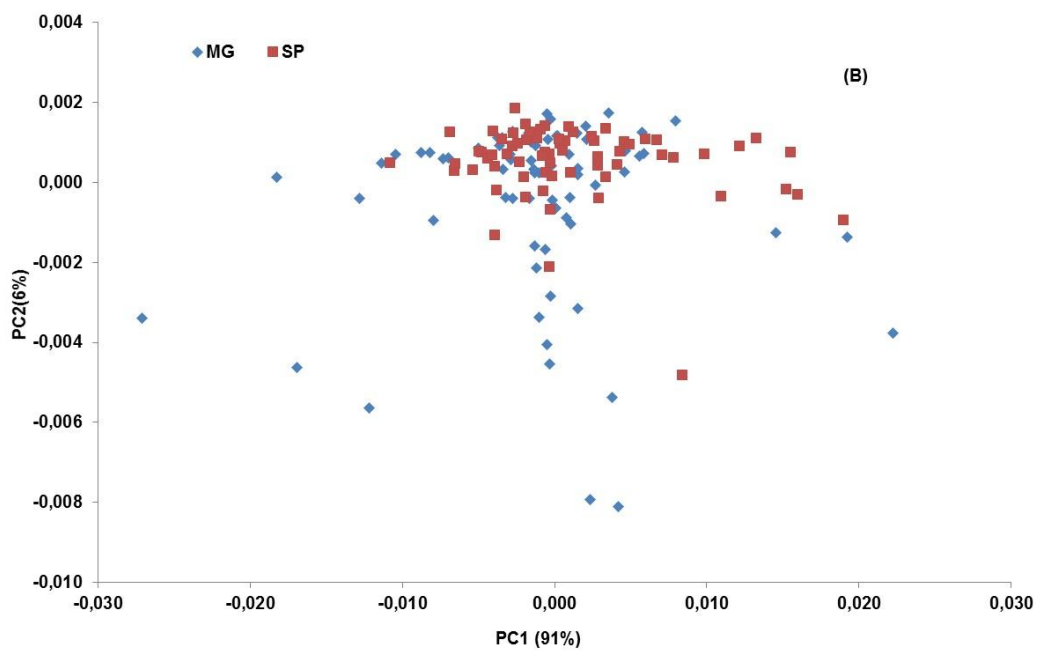
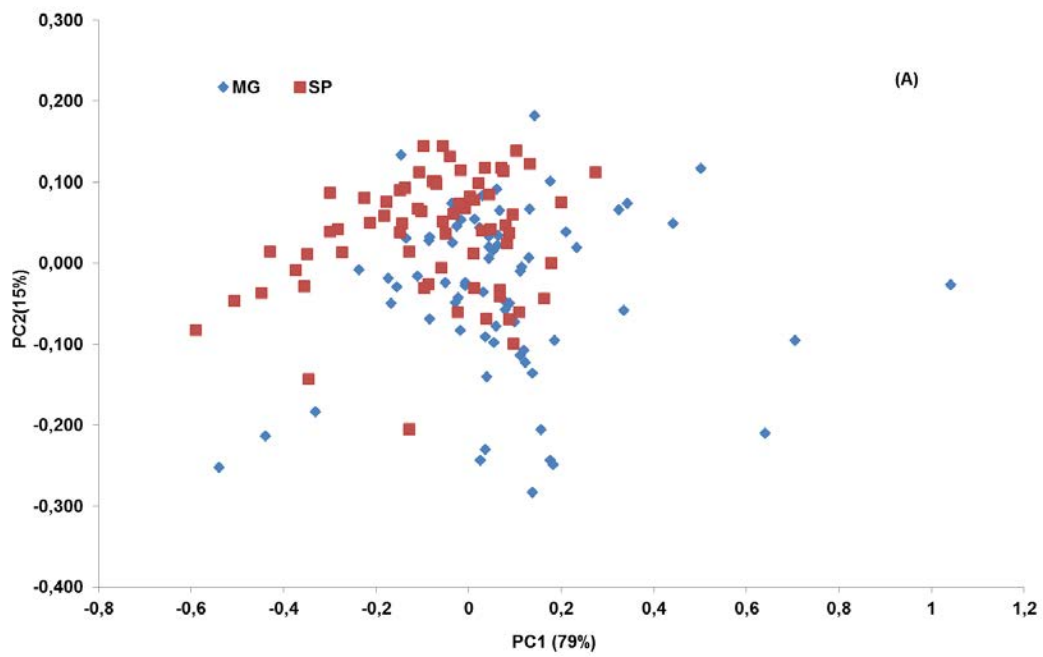


Figura 7. Escores do PC1 e PC2 obtidos com espectros NIR (1000-2500 nm) das amostras de aguardente de cana-de-açúcar provenientes dos estados de São Paulo e Minas Gerais após a utilização da correção do espalhamento multiplicativo de sinal (MSC) (A) e segunda derivada de Savitsky-Golay com janela de 4+4 pontos (B).

Pode-se observar que mesmo após o uso dos pré-processamentos não houve uma separação nítida entre as amostras (Figura 7), com a distribuição das amostras nas duas primeiras componentes principais dos PCAs apresentando sobreposições entre as classes, isto pode ser devido à similaridade entre as mesmas (COZZOLINO et al., 2006). Características da matéria-prima, fatores climáticos e o processo de produção, podem influenciar nessas similaridades e também nas diferenças das amostras, de acordo com o estado produtor (DANTAS et al., 2007).

Desta forma, o uso de técnicas de classificação supervisionadas é uma alternativa para a obtenção da separação desta bebida entre os estados de São Paulo e Minas Gerais.

3.2.2. Análises supervisionadas (PLS-DA / LDA / SVM)

As técnicas de classificação supervisionadas são aplicadas a uma ampla variedade de dados químicos com diversas finalidades, tais como identificar perfis, impressões digitais, autenticação, detecção de falsificação, avaliação da qualidade dos alimentos, interpretação de dados, etc. (LAVINE, 2006). A primeira técnica testada foi a análise discriminante utilizando o método de mínimos quadrados parciais (PLS PLS-DA). O modelo PLS-DA foi desenvolvido utilizando toda a faixa de comprimento de onda do NIR (1.000-2.500 nm) e também utilizando os espectros sem e com pré-processamento (MSC e 2ªD S-G 4+4 pontos de suavização), foram utilizados 7 variáveis latentes para cada modelo (Tabela 2).

Os modelos desenvolvidos na etapa de calibração apresentaram valores de classificação satisfatórios, sendo o melhor modelo o obtido com os

espectros sem pré-processamento, apresentando uma classificação correta de 81%, ou seja, das 100 amostras utilizadas na calibração do modelo, apenas 19 foram classificadas incorretamente, destas 5 pertenciam ao estado de São Paulo e 14 ao de Minas Gerais.

Tabela 2. Desempenho de classificação e predição dos modelos de análise discriminante com método de regressão por mínimos quadrados parciais (PLS-DA) desenvolvidos com e sem pré-processamento no comprimento de onda de 1000 -2500 nm para aguardente de cana de açúcar em função do estado produtor, estado de São Paulo e Minas Gerais.

Pré-processadas	Calibração (n=100)			Predição (n=39)		
	Classificadas incorretamente		Acurácia (%)	Classificadas incorretamente		Acurácia (%)
	SP (n=45)	MG (n=55)		SP (n=24)	MG (n=15)	
Sem Processamento	5	14	81	4	2	84,6
MSC	7	13	80	4	1	87,2
2ª D SG	9	14	77	5	4	76,9

Pré Proces: pré-processamento, Sem Proces: sem processamento, MSC: correção do espalhamento multiplicativo de sinal, 2ª D SG: segunda derivada de Savitsky-Golay com janela de 4+4 pontos, SP: São Paulo e MG: Minas Gerais

Em relação ao desempenho na classificação do conjunto externo de amostras o melhor resultado foi obtido com o modelo PLS-DA desenvolvido com os espectros NIR pré-tratados com MSC, que apresentou uma classificação correta de 87,2%. Por outro lado, o uso da 2ª D SG apresentou o pior resultado de classificação do grupo de predição (76,9%, Tabela 2).

Das amostras classificadas incorretamente no conjunto de predição, quatro em 24 amostras pertencentes ao estado de São Paulo foram classificadas como sendo do estado de Minas Gerais, o que levou a uma acurácia de 83,3% de amostras corretamente classificadas como sendo deste

estado. Se tratando de Minas Gerais, uma em 15 amostras foram incorretamente classificadas como sendo deste estado, o que gerou um resultado de 93,3% de acertos. Estes resultados foram superiores aos 76% de classificação obtidos por Cozzolino et al (2011) ao utilizarem modelos PLS-DA visando a discriminação de amostras de vinhos Sauvignon Blanc produzidos na Austrália e Nova Zelândia.

A representação gráfica do desempenho de predição das classes de aguardente de cana de açúcar em função do estado de origem utilizando do melhor modelo PLS-DA (pré-processado com MSC) pode ser vista na Figura 8.

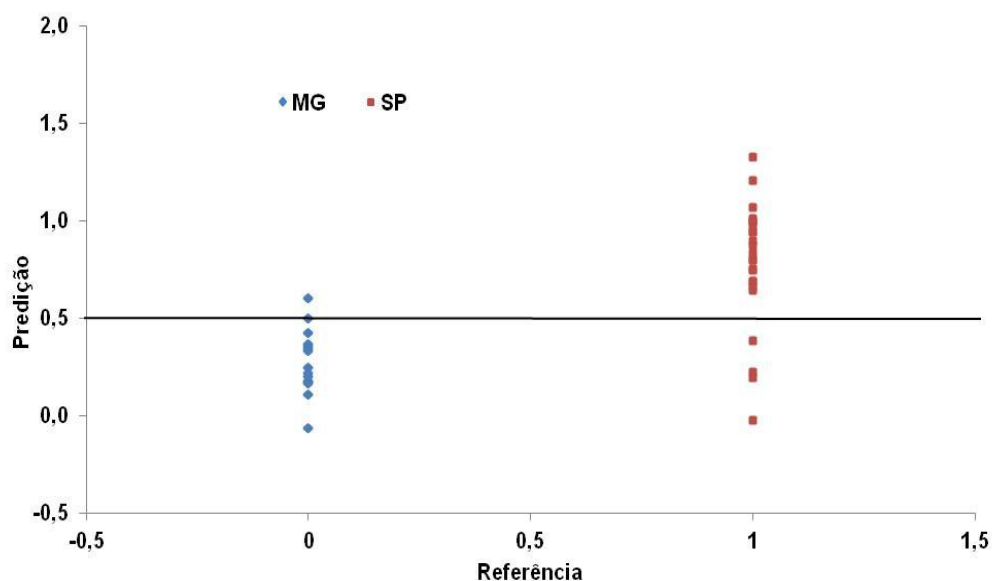


Figura 8. Valores de predição e referência de classes de aguardentes de cana-de-açúcar produzidas nos estados de São Paulo (1) e Minas Gerais (0) utilizando o modelo PLS-DA pré-processados com MSC e valor de corte de 0,5.

Ao utilizar o valor de corte de 0,5, como proposto por Cunha Júnior et al. (2015), nota-se que para a classe Minas Gerais, apenas uma amostra ficou

acima do valor e para a classe do estado de São Paulo quatro ficaram abaixo deste valor (Figura 8).

Das amostras de São Paulo classificadas como de Minas Gerais, três delas foram doadas pelo Laboratório de Análise Físico-Química de Aguardente de cana (LAFQA), portanto não há conhecimento da cidade de origem delas, apenas o estado (São Paulo), entretanto uma amostra classificada incorretamente é oriunda da cidade de Limeira - SP, enquanto a amostra de Minas Gerais classificada como São Paulo é da cidade de São Thiago - MG, ambas as amostras em seus rótulos são classificadas como aguardente de cana.

Na tentativa de se melhorar a classificação das amostras, foi realizada a análise linear discriminante (LDA), Tabela 3.

Os resultados obtidos com os modelos do conjunto de calibração foram semelhantes, apresentando uma classificação correta máxima de 82%, tanto para o modelo desenvolvido com espectros sem e com o pré-processamento MSC. Enquanto o modelo utilizando espectros pré-processados com a 2ª derivada de Savitsky Golay obteve um valor de acurácia de calibração 81%, superior ao encontrado no modelo PLS-DA (77%).

A análise discriminante linear (LDA) se baseia no uso das componentes principais (PCs), desta forma, os modelos apresentados na Tabela 3 foram desenvolvidos com a utilização de 10 PCs e isso explica uma maior variabilidade dos dados, gerando modelos mais robustos. Modelos mais simples podem ser obtidos com menor número de PCs. Liu et al. (2006), ao utilizar a LDA para a classificação de vinhos Tempranillo produzidos na Austrália com apenas 3 PCs relataram que 72% dos vinhos australianos foram

corretamente classificados, sendo esta porcentagem maior para os vinhos espanhóis, ou seja 85%. Nota-se que a acurácia destes modelos foram inferiores aos apresentados na Tabela 3, o que pode ser um indicativo que o número de PCs é importante para a performance do modelo (CHEN et al., 2008).

Tabela 3. Desempenho de classificação e predição dos modelos de análise linear discriminante (LDA) desenvolvidos com e sem pré-processamento no comprimento de onda de 1000 -2500 nm para aguardente de cana-de-açúcar em função do estado produtor, São Paulo e Minas Gerais.

Pré-Proces.	Calibração (n=100)			Predição (n=39)		
	Classificadas incorretamente		Acurácia (%)	Classificadas incorretamente		Acurácia (%)
	SP (n=45)	MG (n=55)		SP (n=24)	MG (n=15)	
Sem proces.	15	3	82	4	2	84,6
MSC	14	4	82	4	1	87,2
2ª D SG	14	5	81	6	1	82,1

Pré-Proces: pré-processamento, Sem proces: sem processamento, MSC: correção do espalhamento multiplicativo de sinal, 2ª D SG: segunda derivada de Savitsky-Golay com janela de 4+4 pontos, SP: São Paulo e MG: Minas Gerais.

O uso da LDA não proporcionou grandes modificações aos resultados de predição, com exceção do modelo desenvolvido com espectros processados com a 2ªD S-G que aumentou a classificação correta do grupo de predição de 76% do modelo PLS-DA (Tabela 2) para 82,1% no modelo LDA (Tabela 3).

Além dos modelos PLS-DA e LDA, foi também testado a classificação das amostras usando a técnica da máquina de vetores de suporte (SVM), Tabela 4. Para se obter um bom modelo SVM, o tipo de função kernel (linear,

polinomial, sigmoide e função de base radial) e os parâmetros da função devem ser devidamente determinados (MA & YANG, 2011; DONG et al., 2014). A função kernel reduz a complexidade computacional (DONG et al., 2014), aumentando o espaço dimensional para projetar os dados e transformando os dados linearmente inseparáveis em separáveis (LI et al., 2009).

Foi testada a função kernel linear e de base radial, optou-se pela função kernel linear, pois esta apresentou o melhor resultado quando comparada a função kernel de base radial (RBF), que é considerada a primeira escolha (HSU et al., 2010). Na kernel RBF todas as amostras foram consideradas como sendo de uma mesma classe, o valor de C de 5 permitiu a construção de um modelo de fácil interpretação (SILVA et al., 2014).

Os resultados dos modelos SVM na etapa de calibração foram inferiores ao atingidos nos outros modelos (PLS-DA, LDA) nos três modelos construídos, o modelo com os espectros sem pré-processamentos apresentou uma classificação correta de 78% das amostras de aguardente de cana-de-açúcar em função do estado de origem.

Tabela 4. Desempenho de classificação e predição dos modelos desenvolvidos pela técnica de máquina de vetores de suporte (SVM), com e sem pré-processamento dos espectros no infravermelho próximo (1000 -2500 nm) de aguardentes de cana-de-açúcar em função do estado produtor, estado de São Paulo e Minas Gerais.

Pré-Proces.	Calibração (n=100)			Predição (n=39)		
	Classificadas incorretamente		Acurácia (%)	Classificadas incorretamente		Acurácia (%)
	SP (n=45)	MG (n=55)		SP (n=24)	MG (n=15)	
Sem proces.	13	9	78	4	1	87,2
MSC	24	8	68	11	0	71,8
2ª D SG	45	0	55	0	24	38,5

Pré-Proces: pré-processamento, Sem proces: sem processamento, MSC: correção do espalhamento multiplicativo de sinal, 2ª D SG: segunda derivada de Savitsky-Golay como janela de 4 pontos, SP: São Paulo e MG: Minas Gerais

O modelo SVM com os espectros brutos apresentou o melhor resultado com 78 % e 87,2 % de classificação corretas para calibração e predição, respectivamente, classificando corretamente 34 em 39 amostras do conjunto de predição, Silva et al. (2014), utilizando os espectros brutos de NIR e a função kernel linear, classificaram corretamente 11 das 12 amostras de gasolina.

O modelo desenvolvido utilizando a máquina de vetores de suporte (SVM) em espectros, apresentou o menor número de vetores de suporte (VS = 70), isto pode ser uma indicação de um modelo mais robusto como descrito por Alves et al. (2014). Contudo, o uso de modelos SVM também não melhoram acentuadamente os resultados em relação a classificação das amostras (Tabela 4).

Para este modelo SVM desenvolvido com toda a faixa espectral (1000-2500 nm), o modelo construído com espectros sem pré-processamento foi o que apresentou a melhor classificação na predição, 87,2%, (Tabela 4), sendo considerado juntamente com os modelos PLS-DA e LDA, na mesma faixa de comprimento de onda, pré-processados com MSC, técnicas adequadas para a classificação de amostras de aguardentes de cana de açúcar em função do seu origem

Vale enfatizar que foram investigadas a origem amostras de aguardentes de cana que não foram corretamente classificadas, em todos os modelos no grupo de predição, pois poderia ter havido um erro em relação ao estado que elas foram produzidas, São Paulo e Minas Gerais. Isso não ocorreu, ou seja, todas as amostras realmente pertenciam a suas classes.

4. CONCLUSÕES

As diferenças de coloração encontradas nas amostras de aguardente de cana-de-açúcar possivelmente são provenientes do processo de envelhecimento e/ou adição de aditivos a que elas possivelmente foram submetidas.

A espectroscopia NIR combinada às técnicas quimiométricas proporcionaram uma correta classificação (87,2%) de aguardente de cana-de-açúcar produzidas em regiões distintas, São Paulo e Minas Gerais. Esta metodologia pode ser uma alternativa aos métodos analíticos tradicionais que apresentam altos custos e requerem longos períodos de análise.

No entanto, os modelos de classificação apresentam algumas limitações, tais como o número de amostras utilizadas para a construção dos mesmos, semelhanças entre algumas amostras de aguardente de cana-de-açúcar em função das condições edafoclimáticas. Portanto, é necessário melhorar a acurácia dos modelos de predição antes destes serem empregados como técnica analítica para o controle de qualidade de aguardente de cana-de-açúcar

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ALVES, J.C.; HENRIQUE, C.; POPPI, R.J. Classification of diesel pool refinery streams through near infrared spectroscopy and support vector machines using C-SVC and m-SVC. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 117, p. 389–396, 2014.
- BECHE, R. V.; FRANCO, D. W. Distinção entre cachaças destiladas em alambiques e em colunas usando quimiometria. **Química Nova**, São Paulo, v. 32, n. 2, p. 332-336, 2009.
- BERNARDES, C.D.; FIGUEIREDO, M.C.P. e BARBEIRA, P. J.S. Developing a PLS model for determination of total phenolic content in aged cachaças. **Microchemical Journal**, v. 116, pp. 173-177, 2014.
- BEVIN, C. J.; DAMBERGS, R. G.; FERGUSON, A.J.; COZZOLINO, D. Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. **Analytica chimica acta**, v.621, p.19–23, 2006.
- BRASIL. Decreto número 4062, de 21 de dezembro de 2001. Define as expressões “cachaça”, “Brasil” e “cachaça do Brasil” como indicações geográficas e dá outras providências. **Diário Oficial da República federativa do Brasil**. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.
- BRASIL. Decreto número 6871, de 4 de junho de 2009. Regulamenta a lei número 8.918 de julho de 1994, sobre a padronização, a classificação, o registro, a inspeção, a produção e a fiscalização de bebidas. **Diário Oficial**

da República federativa do Brasil. Disponível em:
<<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

BRASIL. Instrução Normativa número 13, de 29 de junho de 2005. Aprova o regulamento técnico para fixação dos padrões de identidade e qualidade para aguardente de cana e para cachaça. **Diário Oficial da República federativa do Brasil.** Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

BRASIL. Lei número 9279, de 14 de maio de 1996. Dispõe sobre direitos e obrigações relativos à propriedade industrial. **Diário Oficial da República Federativa do Brasil.** Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

BRERETON, R.; **Chemometrics for Pattern Recognition**, John Wiley & Sons: Chichester, 2007

BYLESJO, M, RANTALAINEN, M.; CLOAREC, O; NICHOLSON, J.K.; HOLMES, E.; TRYGG, J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification, v.20, p. 341-351, 2006.

CARDELLO, H. M. A. B.; FARIA, J. B. Análise descritiva quantitativa da aguardente de cana durante o envelhecimento em tonel de carvalho (*Quercus alba L.*). *Ciência e Tecnologia de Alimentos* v. 18, n. 2, p.169-175, 1998.

CHEN Y, XIE MY, YAN Y, ZHU SB, NIE SP, LI C, WANG YX, GONG XF. Discrimination of *Gonoderma lucidum* according to geographical origin with near infrared diffuse reflectance spectroscopy and pattern recognition techniques. **Analytica Chimica Acta**: pp.121-130, 2008.

CORTES, C.; VAPNIK, V. N.. **Machine Learning**, v. 20, p.273–297. 1995

- COZZOLINO, D.; CYNKAR, W.; SHAH, N.; SMITH, P.A. Can spectroscopy geographically classify Sauvignon blanc wines from Australia and New Zealand? **Food Chemistry**, n. 126, P.673-678, 2011.
- COZZOLINO, D. Recent Trends on the Use of Infrared Spectroscopy to Trace and Authenticate Natural and Agricultural Food Products. **Applied Spectroscopy Reviews**, v. 47, n.7, P.518-530, 2012.
- CUNHA JUNIOR, L. C.; NARDINI, V.; KHATIWADA, BED P.; Teixeira, G. H. A.; WALSH, KERRY B. Classification of intact açai (*Euterpe oleracea* Mart.) and juçara (*Euterpe edulis* Mart) fruits based on dry matter content by means of near infrared spectroscopy. **Food Control**, v. 50, p. 630-636, 2015
- DANTAS, H.J.; VILAR, F.A.; SILVA, F. L. H.SILVA, A.S. Avaliação da influência da velocidade de destilação na análise físico-química de aguardente de cana-de-açúcar. **Revista Brasileira de Produtos Agroindustriais**, Campina Grande, v.9, n.2, pp.101-109, 2007.
- DONG, J.J.; LI, Q.L.; YIN, H.; ZHONG, C.; HAO, J.G.; YANG, P.F.; TIAN, Y.H.; JIA, S.R. Predictive analysis of beer quality by correlating sensory evaluation with higher alcohol and ester production using multivariate statistics methods. **Food Chemistry**, v. 161, pp. 376–382, 2014.
- FERNANDES, H.L.; RAIMUNDO JR, IVO, M.; PASQUINI, C.; ROHWEDDER J. Simultaneous determination of methanol and ethanol in gasoline using NIR spectroscopy: Effect of gasoline composition. **Talanta**, v. 75, pp.804–810, 2008.
- HARTIGAN, J.A. *Clustering algorithms*. New York: Wiley, 1975. 351p.

- HSU, C.W.; Chang, C.C.; LIN, C.J.. A practical guide to support vector classification. **Technical report**, Department of Computer Science, National Taiwan University. July, 2009.
- INÁCIO, M.R.C.; LIMA, K.M.G.; LOPES, V.G.; PESSOA, J.D.C.; TEIXEIRA, G.H.A. Total anthocyanin content determination in intact açai (*Euterpe oleracea* Mart.) and palmitero-juçara (*Euterpe edulis* Mart.) fruit using near infrared spectroscopy(NIR) and multivariate calibration. *Food Chemistry*, v.136, 1160-1164, 2013.
- INPI. Indicação Geográfica-IG, 2014. Disponível em: <http://www.agricultura.gov.br/desenvolvimento-sustentavel/indicacao-geografica>. Acesso em: 09 de maio de 2015.
- ISAKSSON, T., TORMOD, N. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR Spectroscopy. **Applied Spectroscopy**, v.42, n. 7, p 1273- 1284, 1988
- KENNARD, R. W. e STONE, L. A. Computer aided design of experiments. **Technometrics**, v.11, p. 137–148, 1969.
- LAVINE, B.K. Pattern recognition. **Critical Reviews in Analytical Chemistry**, n.36, p. 153–161, 2006.
- LI, H.; LIANG Y.; XU, Q. Support vector machines and its applications in chemistry. **Chemometrics and Intelligent Laboratory Systems** 95, p.188–198, 2009.
- LIU, L., COZZOLINO, D., CYNKAR, W.U., DAMBERGS, R.G., JANIK, L., O'NEILL, B.K., COLBY, C.B., AND GISHEN, M. Geographic classification of Spanish and Australian Tempranillo red wines by visible and near infrared

- spectroscopy combined with multivariate analysis. **Food Chem.**, v. 106:781, 2006.
- LIU, L.; COZZOLINO, D.; CYNKAR W.U.; DAMBERGS R.G.; JANIK L.; O'NEILL B.K.; COLBY C.B.; GISHEN M.. Preliminary study on the application of visible–near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. **Food Chemistry**, v.106, p. 781-786, 2008.
- LLÉTI, R.; MELÉNDEZ E.; ORTIZ M.C.; SARABIA L.A.; SÁNCHEZ, M.S. Outliers in partial least squares regression. Application to calibration of wine grade with mean infrared data. **Analytica Chimica Acta**, v. 544, pp. 60–70, 2005.
- McGUIRE, R.G. Reporting of objective color measurements. **HortScience**, v.27, p. 254-255, 1992.
- MARTENS, H., NAES, T. **Multivariate Calibration**. New York: John Wiley & Sons, 1996.
- MIRALBÉS, C. Discrimination of European wheat varieties using near infrared reflectance spectroscopy. **Food Chemistry**, v. 106, pp.386–389, 2008
- NAES, T.; ISAKSSON, T.; FEARN T.; DAVIES, T. *A User-friendly Guide to Multivariate Calibration and Classification*, **NIR Publications**, Chichester, UK, 2002.
- PASQUINI, C. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. **Journal of Brazilian Chemistry Society**, v.14, n.2, p.198-219, 2003

- SABIN.J.G.; FERRAO, M.F.; FURTADO, J.C. Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA. **Revista Brasileira de Ciências Farmacêuticas**, v. 40, n. 3, 2004.
- SAVITSKY, A., GOLAY, M.J.E. Smoothing and differentiation by simplified least squares procedures. **Analytical Chemistry**, v.36, p.1627-1632, 1964.
- SILVA, M.P.; BRITO, L.R.; HONORATO, F.A.; PAIM, A.P.; PASQUINI, C.; PIMENTEL, M.F. Classification of gasoline as with or without dispersant and detergent additives using infrared spectroscopy and multivariate classification. **Fuel**, v. 116, p.151–157, 2014.
- StatSoft, Inc. (2004). Statistica (data analysis software system), version 7. www.statsoft.com.
- SUSI, H. e BYLER, D. M. Biochem. Biophys. **Res. Commun.** 115, 391-397, 1983.
- WONG, K.H.; RAZMOVSKI-NAUMOVSKI, V.; LI, K.M.; LI, G.Q.; CHAN, K. Differentiation of Pueraria lobata and Pueraria thomsonii using partial least square discriminant analysis (PLS-DA). **Journal of Pharmaceutical and Biomedical Analysis**, v. 84, pp.– 13, 2013.

Uso da espectroscopia do infravermelho próximo (NIR) como método não destrutivo para a determinação do teor de etanol de aguardentes de cana de açúcar

RESUMO

A aguardente de cana-de-açúcar é uma bebida típica do Brasil obtida pela destilação do fermentado do mosto de cana-de-açúcar. Por definição, aguardente de cana é a bebida destilada que tem o teor alcoólico entre 38 a 54% em volume (v/v) a 20°C, e cachaça é a bebida destilada cujo teor alcoólico varia entre 38 a 48% (v/v) a 20°C. Como o teor alcoólico é um importante atributo de qualidade, o objetivo deste trabalho foi usar a espectroscopia do infravermelho próximo (NIR) como um método não destrutivo para prever o teor alcoólico de aguardentes de cana-de-açúcar. Foram utilizadas 139 amostras provenientes do estado de São Paulo e Minas Gerais, divididas em conjunto de calibração (n = 100) e predição (n = 39) pelo algoritmo clássico Kennard-Stone. Os espectros NIR foram obtidos utilizando o espectrômetro FT-IR Spectrum 100N como $\log 1/R$, na faixa de número de ondas de 4.000 a 10.000 cm^{-1} (1.000-2.500 nm) com 64 scans a uma resolução espectral de 16 cm^{-1} , com intervalo de 8 cm^{-1} . O método de referência para determinação do teor de etanol foi o densimétrico, seguindo as recomendações da NBR 59992 (2008) da ABNT. Empregou-se a regressão por mínimos quadrados parciais (PLSR) para desenvolver os modelos de predição de etanol e foi empregada a seleção de comprimentos de onda em função dos *loadings*, para melhorar o

acurácia dos modelos. Ao se utilizar toda a faixa espectral (1000 - 2500 nm), o modelo obtido com os espectros NIR sem pré-processamento apresentou um $RMSE_{CV}$ de 1,67% e R^2_{CV} de 0,86 na etapa de calibração cruzada. Na etapa de predição, o melhor modelo foi construído abrangendo os comprimentos de onda 1393-1541 nm e 1843-2166 nm, atingindo um erro de predição de 1,98% e RPD de 2,02. A espectroscopia NIR e a análise de regressão PLSR demonstraram potencial para predizer o teor de etanol de amostras de aguardente de cana-de-açúcar, oferecendo a possibilidade de uso como uma ferramenta para determinar rapidamente e de forma não destrutiva o teor de etanol

ABSTRACT

Sugarcane aguardente is a typical Brazilian spirit produced by the distillation of fermented sugarcane most. By definition aguardente is the sugarcane distillate which has alcoholic content between 38 to 54% in volume (v/v) at 20°C, and cachaça is the sugarcane distillate with alcoholic content between 38 to 48% (v/v) at 20°C. As ethanol content is an important quality attributes, the objective of this study was to use NIR spectroscopy as a non-destructive method to predict ethanol content in Sugarcane aguardente. A total of 139 samples of sugarcane aguardentes from Minas Gerais (n=70) and São Paulo (n=69) were collected. The samples were divided into calibration set (n = 100) and prediction set (n= 39) using Kennard-Stone algorithm. NIR spectra were collected using a FT-NIR spectrometer over the range of 4,000 to 10,000 cm^{-1} (1,000 to 2,500 nm) at a spectral resolution of 16 cm^{-1} with interval of 8 cm^{-1} and 64 scans per spectra. The reference method for alcoholic content determination was the densimetric procedure. The least partial square (PLS) regression was used to develop the ethanol prediction models and a selection wavelength was done depending on the loadings to improve the accuracy of the models. When the whole spectral range (1000 - 2500 nm) was used with NIR spectra without preprocessing NIR the RMSE_{CV} was of 1.67% and the R^2_{CV} of 0.86 in cross-calibration step. For prediction, the best model was built covering the wavelengths of 1393-1541 nm and 1843-2166 nm, reaching a prediction error of 1.98% and RPD of. 2.02. PLSR was a good non-destructive method to predict ethanol content in sugarcane aguardente.

1.INTRODUÇÃO

A produção anual de aguardente de cana-de-açúcar no Brasil está estimada em 1,2 bilhões de litros anuais, sendo produzidos em aproximadamente 12 mil estabelecimentos. Porém, legalmente registrados no Ministério de Agricultura, Pecuária e Abastecimento (MAPA) e Receita Federal, são menos de 2.000 estabelecimentos, representados por 4.182 marcas, sendo os principais estados produtores São Paulo, Rio Grande do Sul, Ceará, Minas Gerais e Rio de Janeiro (IBRAC, 2013).

Os termos aguardente de cana-de-açúcar e cachaça são definidos Decreto número 6871, de 4 junho de 2009, do Ministério da Agricultura, Pecuária e Abastecimento (BRASIL, 2009), que os define como destilados de cana-de-açúcar fermentado. Por este Decreto, a aguardente de cana é o destilado de cana-de-açúcar que possui graduação alcoólica entre 38 e 54% em volume a 20°C, e a cachaça é o destilado de cana-de-açúcar fermentado com graduação alcoólica entre 38 e 48% em volume a 20 °C, ainda conforme este decreto há também a classe de destilado alcoólico simples de cana-de-açúcar, este apresenta graduação alcoólica entre 54 e 70% em volume a 20 °C.

Alguns dos requisitos de qualidade dessa bebida, incluem a concentração dos componentes secundários, que devem apresentar concentrações superiores a 200 mg.100 mL⁻¹ e inferiores a 650 mg.100 mL⁻¹ de álcool anidro, com os seguintes limites máximos para cada componente: 150 mg.100 mL⁻¹ de álcool anidro para acidez volátil (expressa em ácido acético),

200 mg.100 mL⁻¹de álcool anidro de ésteres, 30 mg.100 mL⁻¹ de álcool anidro de aldeídos totais, 5 mg.100 mL⁻¹ de álcool anidro da soma de furfural e hidroximetilfurfural e 360 mg.100 mL⁻¹ de álcool anidro de álcoois superiores (soma de álcoois n-propílico, isobutílico e isoamílico), (BRASIL, 2005).

Muitos métodos têm sido utilizados para a determinação da graduação ou teor alcoólico de bebidas. Entre os descritos pela “Association of Official Analytical Chemists” (AOAC) constam os métodos refratométricos e os métodos de densidade que utilizam picnômetros ou densímetros para medir o teor alcoólico de bebidas destiladas (BARBOZA & POPPI, 2003). Outros métodos podem ser encontrados na literatura, tais como os cromatográficos, métodos eletroquímicos e a espectroscopia do infravermelho próximo (NIR) (BARBOZA & POPPI, 2003; MENDES et al., 2003).

A espectroscopia NIR tem se mostrado muito interessante, devido as suas características de versatilidade, flexibilidade e eficácia (MARTELO-VIDAL & VASQUEZ, 2014), sendo considerada uma técnica rápida, segura de operação, não destrutiva, relativamente barata, proporcionando automação dos processos de controle de qualidade nos produtos da agroindústria (PASQUINI, 2003).

Os espectros NIR constituem um conjunto de dados que contêm muitas informações, e um dos métodos quimiométricos que tem auxiliado no desenvolvimento de equações, para a extração dessas informações, é a regressão por mínimos quadrados parciais (PLSR) (URBANO CUADRADO et al., 2005).

A técnica NIR associada ao PLSR, tem sido muito utilizada em bebidas alcoólicas, promovendo a sua caracterização e avaliação de alguns

parâmetros, entre eles o teor de etanol, como na indústria de vinhos (URBANO-QUADRADO et al., 2004; URTUBIA et al.,2004; COZZOLINO et al.,2006; MARTELLO-VIDAL e VÁSQUEZ, 2014), cervejas (CASTRITIUS et al., 2010) ou em vodcas (KOLOMIETS et al., 2010).

O etanol, como descrito por Cozzolino et al (2006), tem um forte sinal de absorbância na região do infravermelho próximo em bebidas alcoólicas, com uma intensidade menor apenas que a água, porém a precisão e robustez da calibração pode ser limitada pela variação das amostras e variações na concentração de açúcar.

Na literatura há poucos trabalhos relatando a aplicação do NIR para a determinação teor de etanol em aguardentes de cana de açúcar. Deste modo, objetivo deste trabalho foi desenvolver modelos de predição do teor de etanol de aguardente de cana-de-açúcar, utilizando a espectroscopia NIR e a regressão por mínimos quadrados múltiplos (PLSR).

2. Material e métodos

2.1. Obtenção das amostras

Para se produzir um modelo robusto, o mesmo deve incluir a variabilidade natural da propriedade de interesse e também um número mínimo de amostras empregadas na calibração (50 a 100), Pasquini (2003). Assim, foram utilizadas 139 amostras de aguardentes de cana-de-açúcar, provenientes dos estados de São Paulo e Minas Gerais (Anexo 1). As amostras foram cedidas pelo Centro de Pesquisa e Desenvolvimento Qualidade da

Cachaça da Universidade Estadual Paulista, Faculdade de Ciências Farmacêuticas (UNESP-FCFAR, n = 83), pelo laboratório de Análises Físico-Químicas de Aguardente (LAFQA) da Universidade Federal de Lavras (UFLA) (n = 49); e algumas foram compradas no comércio de Jaboticabal - SP (n = 7).

Destas amostras, 100 amostras foram selecionadas para o conjunto de calibração e 39 para o conjunto de predição, utilizando o algoritmo de seleção Kennard-Stone clássico aos espectros NIR (KENNARD & STONE, 1969).

2.2. Aquisição dos espectros NIR

De cada amostra foram obtidos cinco espectros na região do infravermelho próximo, o que totalizou 695 espectros. Os espectros foram obtidos na região NIR utilizando um espectrômetro FT-IR Spectrum 100N (PerkinElmer, Shelton, EUA), na geometria óptica de reflectância como $\log 1/R$, na faixa de número de ondas de 4.000 a 10.000 cm^{-1} (1.000-2.500 nm) com 64 scans a uma resolução espectral de 16 cm^{-1} , com intervalo de 8 cm^{-1} , totalizando cinco espectros de cada amostra.

2.3. Método de referência para a determinação de etanol

O grau alcoólico real foi determinado pelo método densimétrico seguindo as recomendações da NBR 5992.2008 da ABNT (2005), com uso de um alcoômetro centesimal (Alcoômetro de Gay Lussac – ALLA Brasil), indicando somente a concentração do álcool em volume. O grau alcoólico foi corrigido para a temperatura de 20°C, na qual as determinações do alcoômetro são exatas (ANVISA, 2010). Foram observados valores máximos de 57,1 e mínimos de 30,3 % (Tabela 1).

Tabela 1. Teor de etanol (% v/v) em aguardentes de cana-de-açúcar nas populações obtidas após divisão pelo algoritmo de Kennard-Stone.

População	N	Estado	Etanol % v.v. ⁻¹ 20°C			
			Média	DP	Máximo	Mínimo
Pop1	69	São Paulo (i)	40,1a	3,86	47,6	30,3
Pop2	70	Minas Gerais (ii)	44,6b	3,66	57,1	35,3
Test F			**			
Calibração	100	*i+ii	42,5c	4,48	57,1	30,3
Predição	39	*i+ii	42,0c	4,12	49,1	33,3
Test F						

Dividido por algoritmo de seleção Kennard Stone clássico. N - número de amostras. DP- desvio padrão. Médias seguidas de pelo menos uma letra comum nas colunas, para cada elemento, não diferem entre si pelo teste de Tukey (P<0,05). ** = significativo a 1% de probabilidade pelo teste F. NS = não significativo.

2.4. Análises quimiométricas

Os cinco espectros coletados de cada amostra (n = 695) foram reduzidos a sua média (n = 139), sendo utilizado o programa Unscrambler®, versão X.3 (CAMO, Oslo, Noruega) para a construção dos modelos de predição quanto ao teor de etanol.

2.4.1. Pré-processamento

Os espectros foram pré-processados com o objetivo de melhorar o desempenho dos modelos removendo ou minimizando informações indesejáveis, como ruído sistemático indesejável, espalhamento de luz e aumentando a contribuição da composição química para a construção do modelo. Desta forma foram empregados diferentes pré-processamentos, como:

- Multiplicative Scatter Correction ou correção do espalhamento multiplicativo de sinal – MSC (ISAKSSON & TORMOD, 1988) O MSC foi

utilizado para minimizar o os efeitos aditivos e/ou multiplicativos nos dados espectrais causados pelo espalhamento de luz. (SOUZA & POPPI, 2012).

- Variação Normal Padrão – SNV (BARNES et al., 1993; NAES et al., 2002). Similarmente ao MSC, o SNV foi utilizado visando a remoção das interferências de espalhamento de luz e os efeitos do espalhamento multiplicativo entre diferentes amostras (AZZOUZ et al., 2003).

- Segunda Derivada de Savitsky-Golay - 2ª D SG (SAVITSKY & GOLAY, 1964). Foi utilizada para reduzir o “drift” da linha de base e aumentar o caráter espectral (NAES et al., 2002).

Além dos pré-processamentos anteriormente descritos, foi testado também a combinação dos mesmos, ou seja: SNV + 2ª Derivada de Savitsky-Golay (SAVITSKY & GOLAY 1964).

2.4.2. Construção dos modelos de regressão por mínimos quadrados múltiplos

Para a determinação quantitativa do teor de etanol a partir dos espectros NIR, foi utilizado o método por regressão de mínimos quadrados parciais (PLSR), que estabelece a relação entre os sinais espectrais e o método de referência para a construção dos modelos matemáticos (SIMÕES, 2008).

Como qualquer outro método de regressão, o PLSR tem o objetivo de encontrar uma relação entre uma matriz X com m colunas, que corresponde ao sinal espectral (absorbância) para cada número de onda, e n linhas, representando cada amostra (padrão), com uma matriz Y , definida com k colunas, com as informações químicas e com n linhas, correspondendo a cada amostra (MORGANO et al., 2005; SOUZA & FERRÃO, 2006), gerando

variáveis latentes ou fatores que descrevem as estruturas de covariância destas duas matrizes (DONG et al., 2014).

Assim, é assegurado que as variáveis latentes sejam ordenadas de acordo com a sua relevância para prever as variáveis na matriz Y (COZZOLINO, 2011).

Para escolher o número de ideal de variáveis latentes do modelo, utilizou-se método de validação cruzada total, denominada de “*Leave one out*”. Neste método uma das amostras do conjunto de calibração é excluída da fase de desenvolvimento do modelo, sendo reservada como elemento de previsão. Este processo é repetido n vezes, de maneira a permitir que todos os padrões de calibração (n) participem como elementos de previsão. Finalmente, o erro de previsão ($RMSE_{CV}$) é obtido comparando-se a concentração prevista para cada padrão com o seu valor verdadeiro, valor este que é apresentado em função do número de VL utilizada na elaboração dos modelos (FERREIRA et al., 1999).

O melhor modelo de calibração foi escolhido através do valor de R^2_{CV} e $RMSE_{CV}$, ou seja, o valor do coeficiente de determinação, que deve ser o valor mais próximo de 1 possível, e o valor do erro de validação cruzada total, que deve ser o mais próximo de zero (MAZUR, 2012).

O desempenho dos modelos de predição foi avaliado calculando os erros de predição $RMSE_P$ (root mean square error of prediction) e pelo cálculo do desvio do resíduo de predição (RPD), como a equação abaixo:

$$RPD = \frac{\text{Desvio padrão do grupo de validação}}{RMSE_P \text{ ou } SE_P}$$

Desta forma, os resultados encontrados nos modelos PLSR foram avaliados pelas variáveis $RMSE_c$, $RMSE_{cv}$, R^2_c e R^2_{cv} , e também pelo RPD conforme o sugerido por NICOLAÏ et al. (2007).

2.5. Análise Univariada

Os dados do teor de etanol foram submetidos à análise de variância (ANOVA), testando as diferenças oriundas dos estados de origem (Minas Gerais e São Paulo) e também entre os conjuntos de calibração e predição. Foi utilizado o teste de Tukey a 5% de significância e a análise estatística foi feita utilizando o programa Statistica, versão 7.0 (STATSOFT, 2004).

3. Resultados e Discussão

3.1 Análise Univariada

Os valores obtidos para o teor de etanol variaram de 30,3 a 57,1 % (v/v) a 20°C (Tabela 1). As amostras oriundas do estado de Minas Gerais apresentaram teor de etanol (44,6%) significativamente maiores ($p < 0,05$) que as obtidas no São Paulo (40,1 %).

Baseados na legislação vigente (BRASIL, 2009), das 70 amostras oriundas do estado de Minas Gerais, 60 foram classificadas como cachaças e aguardente de cana-de-açúcar, apresentando o teor de etanol entre 38% a 48%, enquanto 6 apresentaram teor de etanol superior a 48% e inferior a 54%, portando foram classificadas somente como aguardente de cana-de-açúcar. Apenas 1 amostra foi classificada como destilado alcoólico simples de cana-de-

açúcar, pois esta apresentou teor de etanol superior a 54% e inferior a 70%, as amostras restantes (n = 3) não receberam classificação alguma, pois o Decreto número 6871, de 4 de junho de 2009 (BRASIL, 2009), não classifica amostras com teores de etanol inferiores à 38%.

Em relação às amostras oriundas do estado de São Paulo (n = 69), 49 foram classificadas como cachaças ou aguardentes de cana-de-açúcar, por apresentarem um teor de etanol entre 38% a 48%, o restante (n = 20) apresentaram teor de etanol inferior a 38%, não se enquadrando em nenhuma classe (cachaça, aguardente de cana-de-açúcar e destilado de cana-de-açúcar) segundo a legislação em vigor (BRASIL, 2009).

Desta forma, das 139 amostras analisadas, 16,5% (n = 23), não atenderam a legislação vigente quanto ao teor de etanol. Este resultado foi semelhante ao relatado por Caruso et al. (2008) que ao determinarem o teor alcoólico de amostras de diferentes estados brasileiros (n = 60) verificaram que 17% (n = 10) apresentaram teor de etanol abaixo da especificação. Isto indica possíveis problemas no processo de destilação da bebida ou dificuldade na padronização do teor de etanol do produto final (CARUSO et al., 2008).

Os teores máximos (57,1 %), mínimos (35,3%) e médios (44,6%) de etanol das amostras do estado de Minas Gerais, foram muito semelhantes ao encontrado por Labanca e Glória (2006) para este mesmo estado, enquanto Barcelos et al. (2007) e Marinho et al. (2009) encontraram valores superiores para o teor mínimo, 40,8% e 40%, respectivamente e inferiores para o teor máximo de etanol, 43,25% e 42,33%.

Odello et al. (2009) ao determinarem o teor de etanol de 20 amostras de cachaças, sendo 16 oriundas do estado de São Paulo, e o restante de outros

estados, obteve um faixa de 30 a 44% v/v, sendo, portanto, inferior ao encontrado neste estudo, que compreendia a faixa 30,3 a 57, 1% v/v para os estados de São Paulo e Minas Gerais.

As amostras do estado de São Paulo apresentaram valor médio de teor de etanol menor que as amostras do estado de Minas Gerais, este resultado pode estar ligado a tradições culturais da produção, como exemplo temos as cachaças do Sul de Minas que apresentam teor alcoólico menor quando comparadas as produzidas em outras regiões de Minas Gerais (BARCELOS et al., 2007).

3.2. Espectros NIR

Conjuntamente a obtenção dos espectros NIR das amostras de aguardente de cana-de-açúcar, foram coletados espectros NIR em amostras puras de água, etanol, metanol (Figura 1). A esta figura foram incluídos também espectros NIR de aguardente de cana-de-açúcar com baixo e alto teor de etanol visando a melhor observação dos grupos funcionais.

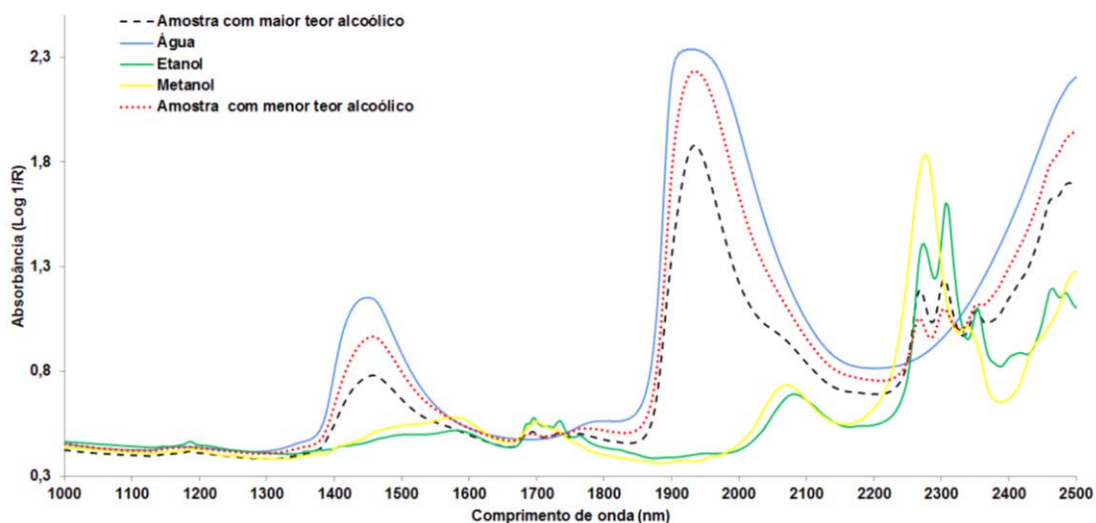


Figura 1. Espectros na região do infravermelho próximo de amostras puras de água, etanol, metanol e amostras aguardentes de cana-de-açúcar.

Na Figura 1 pode-se observar que os espectros NIR das amostras de aguardente de cana-de-açúcar e água pura apresentaram picos nas faixas de 1.420 – 1.500 nm e 1.900 – 2070 nm, que estão relacionados ao primeiro sobreton de OH e a presença de moléculas de água, respectivamente (INÁCIO et al., 2013).

Ainda com relação à Figura 1, é possível identificar picos das amostras pura de etanol e metanol, nos comprimentos de ondas 1700 - 1800 nm, que estão relacionados ao estiramento OH do etanol e metanol (FERNANDES et al., 2008). Os picos presentes na faixa de 2000 - 2100 nm estão relacionados às moléculas que possuem ligações C-H e C=C (SMYTH et al., 2008) e por fim os picos 2.280 – 2.370 nm e 2400 nm, se correlacionam às combinações de CH e estiramentos de OH, e a presença de aromáticos e ao grupo CH₃ presente no metanol (COZZOLINO et al., 2011).

Com isso, foi possível constatar que a amostra com maior teor alcoólico possuiu os picos relacionados ao etanol e metanol mais acentuados, enquanto os picos relacionados à presença da molécula de água, 1.420 – 1.500 nm e 1.900 – 2070 nm, foram menos intensos. Por outro lado, a amostra com menor teor alcoólico apresentou picos relacionados ao etanol menos acentuados e os relacionados à presença de água mais intensos (Figura 1).

3.3. Modelos PLSR para a predição do teor de etanol

Para construção dos modelos utilizou-se os espectros brutos e pré-processados com MSC, SNV e 2ª derivada de Savitisky-Golay (2ªD S-G) e a combinação de SNV+2ªD S-G (Tabela 2). Para o desenvolvimento dos modelos as amostras foram divididas em conjunto de calibração (n = 100) e de predição (n = 39) utilizando o algoritmo de Kennard-Stone clássico (KENNARD & STONE, 1969.).

O conjunto de calibração e predição apresentou teor médio de etanol de 42,5% (v/v) e 42% (v/v), respectivamente, não havendo diferença significativa entre elas ($p < 0,05$), Tabela 1.

Tabela 2. Desempenho da etapa de calibração e predição de modelos de regressão dos mínimos quadrados parciais (PLS) para determinação do teor alcoólico de aguardentes de cana-de-açúcar com diferentes pré-processamentos de espectros de reflectância na região do infravermelho próximo (NIR) (1000-2500nm) de aguardentes de cana de açúcar dos estados de São Paulo e Minas Gerais.

Pré- Processamento	Calibração com validação cruzada total					Predição		
	VL	R ² c	R ² cv	RMSE _c	RMSE _{cv}	R ² p	RMSE _p	RPD
Sem processamento	4	0,87	0,86	1,59	1,67	0,72	2,20	1,88
MSC	3	0,89	0,87	1,71	1,79	0,71	2,30	1,79
2 ^a D S-G	5	0,83	0,80	1,85	2,00	0,69	2,33	1,78
SNV	3	0,85	0,84	1,71	1,79	0,71	2,29	1,80
SNV+2 ^a D S-G	6	0,88	0,83	1,60	1,87	0,76	2,10	1,98

MSC: correção do espalhamento multiplicativo de sinal. 2^a D S-G: 2^a derivada de Savistky Golay com janela de 4 pontos. SNV: variação normal padrão. VL: variável lanete. R²c: coeficiente de regressão linear da etapa de calibração. R²cv: coeficiente de regressão linear da etapa de validão cruzada. RMSE_c: erro de calibração. RMSE_{cv}: erro de validação cruzada R²p: coeficiente de regressão linear da etapa de predição. RMSE_p: erro de predição. RPD: desvio do resíduo de predição.

Ao se utilizar toda a faixa espectral (1000 - 2500 nm), o modelo obtido com os espectros NIR sem pré-processamento apresentou o melhor resultado para calibração, utilizando 4 variáveis latentes (VLs) e um RMSE_{cv} de 1,67% e R²cv de 0,86 (Tabela 2).

Segundo Urbano-Cuadrado et al. (2004), valores de R²cv maiores que 0,90 indicam uma excelente precisão do modelo, assim como valores de R²cv entre 0,70 e 0,90 indicam uma boa precisão; e por outro lado R²cv menores que 0,70 indicam que o modelo pode ser apenas utilizado para distinção entre baixos, médios e altos valores da medida em questão. Desta forma, o valor de R²cv de 0,86 indica uma boa precisão do modelo.

Ao se avaliar a performance dos modelos para o conjunto de predição também utilizando toda a faixa espectral (1.000 – 2.500 nm), o melhor resultado foi obtido quando os espectros NIR foram submetidos aos pré-processamentos SNV+2^aD S-G (Tabela 2). Este modelo gerou o menor RMSEP (2,10%) e o maior R²_p (0,76) com um RPD de 1,98. Segundo Smith et al. (2008), quanto maior o valor do RPD melhor a probabilidade do modelo predizer a composição química com precisão de amostras externas ao do conjunto de calibração. Para Nicolaï et al. (2007), um valor de RPD entre 1,5 e 2 significa que o modelo pode discriminar valores baixos de altos da variável resposta, um valor entre 2 e 2.5 indica que as previsões quantitativas grosseiras são possíveis e um valor entre 2,5 e 3 ou acima corresponde de uma boa a excelente acurácia de predição, respectivamente. Assim, o melhor resultado obtido (RPD de 1,98) permite a discriminação de aguardentes de cana-de-açúcar com altos teores de etanol em relação as de baixos teores.

Na tentativa de se melhorar o desempenho dos modelos de predição, foram selecionados diferentes comprimentos de ondas em função dos *loadings* (Figura 2), sendo utilizados os espectros NIR sem pré-processamento e os processados com SNV, 2^aD S-G e SNV+2^aD S-G (Tabela 3).

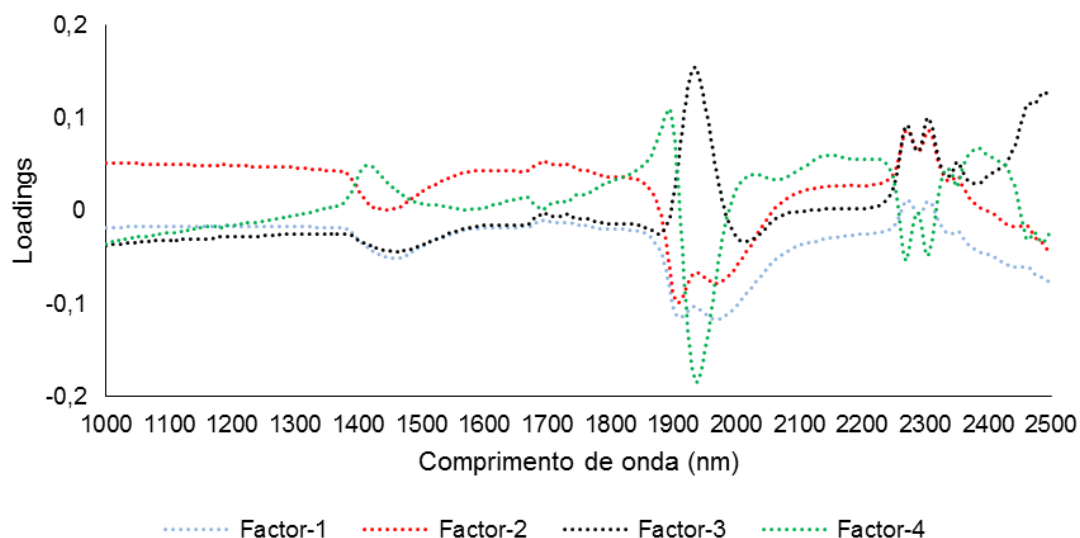


Figura 2: *Loadings* da análise de regressão PLS, utilizados para prever o teor de etanol em amostras de aguardentes de cana de açúcar por espectroscopia do infravermelho próximo. *loadings* dos espectros sem pré-processamento

Os *loadings* dos modelos PLSR normalmente se assemelham ao espectro NIR original, na medida em que expressam as correlações entre cada ponto do comprimento de onda contendo depressões e picos, o que pode ser atribuído a bandas harmônicas, relacionadas com componentes das amostras (COZZOLINO et al., 2008). Os modelos foram desenvolvidos abrangendo duas diferentes faixas espectrais: 1393-1541 nm e 1843-2166 nm; 1097-1278 nm e 1393-1541 nm.

Esses comprimentos de ondas selecionados estão associados com bandas de absorção O- H ou segundo sobreton (1500nm), com estiramento OH (1600-1900nm) e C=O (1900nm), com o primeiro sobreton do O-H (1420-1950nm) e com tons de C-H (2200nm) e C=C (2000-2100nm) Estas regiões espectrais são características tanto de absorção de água (O-H sobretons) como o de teor de etanol ou açúcares (COZZOLINO et al., 2003; LIU et al., 2008).

Os melhores resultados foram obtidos utilizando a combinação de duas faixas de comprimentos de ondas (1.393 -1.541 + 1.843 – 2.166 nm) com os espectros NIR sem pré-processamento (Tabela 3). Para o grupo de calibração foi possível a obtenção de um RMSE_C de 1,47%, R²_C de 0,89. Por outro lado, ao se utilizar o grupo de predição, que não foi incluído ao modelo de calibração, houve uma piora na performance do modelo, pois o RMSE_P aumentou para 1,98% e o R²_P diminuiu para 0,76, com um RPD de 2,02 (Tabela 3).

Tabela 3. Resultado de calibração e predição de modelos de regressão dos mínimos quadrados parciais (PLS), em diferentes regiões do comprimento de onda, para determinação do teor alcoólico de aguardentes de cana-de-açúcar com diferentes pré-processamentos de espectros de reflectância na região do infravermelho próximo (NIR) de aguardentes de cana de açúcar dos estados de São Paulo e Minas Gerais.

Comprimento de onda (nm)	Pré-processamento	Calibração com validação cruzada total					Predição		
		VL	R ² c	R ² cv	RMSEc	RMSE _{cv}	R ² p	RMSEp	RPD
1000-2500 1393-1541		4	0,87	0,86	1,59	1,67	0,72	2,20	1,88
+1843-2166	Sem processamento	6	0,89	0,87	1,47	1,60	0,76	1,98	2,02
1097-1278									
+1393+1541		4	0,86	0,84	1,65	1,76	0,72	2,15	1,92
1000-2500 1393-1541		3	0,85	0,84	1,71	1,79	0,71	2,29	1,80
+1843-2166	SNV	3	0,84	0,83	1,77	1,84	0,65	2,37	1,74
1097-1278									
+1393+1541		3	0,84	0,83	1,77	1,84	0,65	2,37	1,74
1000-2500 1393-1541		6	0,88	0,83	1,60	1,87	0,76	2,10	1,98
+1843-2166	SNV+2 ^a D S-G	4	0,84	0,82	1,74	1,92	0,71	2,19	1,88
1097-1278									
+1393+1541		2	0,83	0,82	1,80	1,88	0,60	2,60	1,58

2^a D S-G: 2^a derivada de Savistky Golay com janela de 4 pontos. SNV: variação normal padrão. VL: variável lanete. R²c: coeficiente de regressão linear da etapa de calibração. R²cv: coeficiente de regressão linear da etapa de validação cruzada. RMSEc: erro de calibração. RMSE_{cv}: erro de validação cruzada R²p: coeficiente de regressão linear da etapa de predição. RMSEp: erro de predição. RPD: desvio do resíduo de predição.

A espectroscopia NIR vem sendo utilizada para a determinação do teor de etanol em várias bebidas. Urbano-Cuadrado et al. (2004) desenvolveu modelos de predição de etanol em diferentes amostras de vinhos ($n = 150$) e relataram menores valores de $RMSE_P$, ou seja, 0,24% (v/v) com R^2 de 0,98. Contudo, o teor alcóolico das amostras de vinho era bem inferior (9,58 – 15,15%) ao das amostras de destilados de cana-de-açúcar (Tabela 1). Por outro lado, Mendes et al. (2003) usando espectros FT-NIR de diferentes bebidas alcoólicas (cerveja, whisky e cachaça) na faixa espectral de 7.824 – 9.113 cm^{-1} e 5.529 – 6.025 cm^{-1} , desenvolveram modelos PLSR com $RMSE_P$ de 0,42% e R^2 de 0,99. Apesar do $RMSE_P$ ter sido menor que os apresentados na Tabela 3, os modelos desenvolvidos por foram feitos com apenas 12 amostras de calibração e 12 de predição.

Resultados mais robustos em relação a determinação do teor de etanol em aguardente de cana-de-açúcar foram relatados por Barboza e Poppi (2003). Estes utilizaram a faixa espectral de 700 – 1.100 nm, denominado ondas curtas do infravermelho próximo (SW-NIR), para a determinação do teor de etanol em amostras de aguardente de cana-de-açúcar ($n = 51$) e relataram um $RMSE_{cv}$ de 0,15%. O interessante deste trabalho é a sua preocupação com a acurácia dos modelos com relação a temperatura, pois segundo Barboza e Poppi (2003), um aumento de temperatura provoca a diminuição do número de grupos e de um aumento de grupos OH livres. Em virtude dessa mudança nos espectros vibracionais, um modelo multivariado não terá boa capacidade de previsão caso não seja realizada a correção da temperatura.

Barboza e Poppi (2003) descrevem duas estratégias para realizar a correção de temperatura. A primeira estratégia é eliminar as diferenças entre os espectros em função da variação na temperatura das amostras pelo desenvolvimento de calibrações multivariadas robustas, a segunda estratégia é baseado na padronização de espectros, o que significa para modelar situações diferentes diretamente.

É importante destacar que os erros obtidos em nosso trabalho foram superiores aos encontrados por Barboza e Poppi (2003), pois as variações de aguardentes de cana-de-açúcar utilizadas para a construção dos nossos modelos de regressão PLS eram grandes, esta variação não só consistia no teor de etanol, mas também na diversidade da origem das mesmas (São Paulo e Minas Gerais), no processo de envelhecimento que algumas delas foram submetidas, nas diferenças de concentração dos componentes secundários e no seu processo de produção.

Mesmo com toda a variação presente no universo amostral, o valor de RPD de 2,02 indica que o modelo tem capacidade de prever a composição química com precisão de amostras fora do conjunto de calibração (SMITH et al., 2008), apesar de Nicolai et al. (2007) relatarem que este valor permite apenas previsões quantitativas grosseiras.

Na Figura 3 podem ser observados os valores de etanol obtidos pelo método de referência e os preditos com o modelo PLS usando os espectros NIR sem pré-processamento e as faixas de comprimentos de ondas (1.393 - 1.541 + 1.843 – 2.166 nm) como apresentado na Tabela 3.

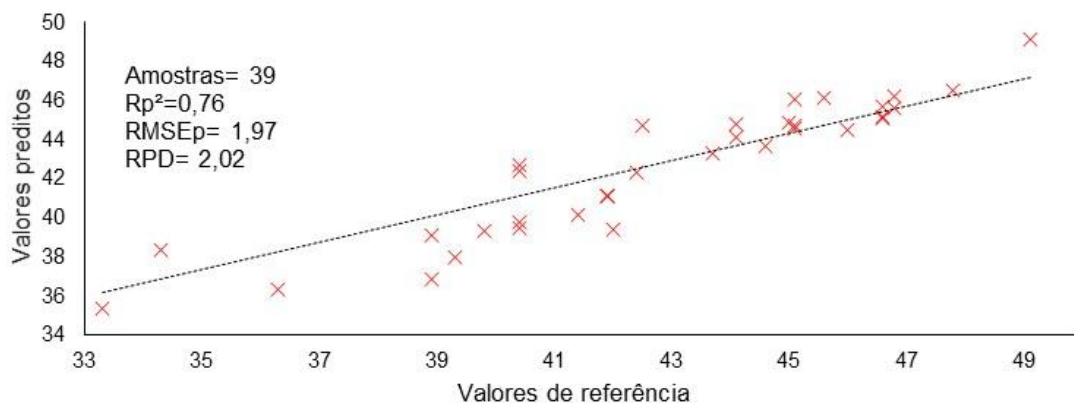


Figura 3. Teor de etanol (% v/v) obtidos pelo método de referência e os teores preditos pelo modelo PLS usando os espectros NIR sem pré-processamento e as faixas de comprimentos de ondas (1.393 -1.541 + 1.843 – 2.166 nm).

4. Conclusões

A técnica da espectroscopia do infravermelho próximo (NIR) e análise quimiométrica de regressão por mínimos quadrados parciais (PLSR) demonstraram o potencial da espectroscopia NIR para prever o teor de etanol de amostras de aguardente de cana-de-açúcar oferecendo a possibilidade de uso como uma ferramenta para determinar rapidamente e de forma não destrutiva o teor de etanol.

Os comprimentos de ondas selecionados e utilizados para a construção do modelo PLSR de calibração contribuíram para a precisão dos modelos desenvolvidos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ABNT. Associação Brasileira de Normas Técnicas. Determinação da massa Específica e do Teor Alcoólico do Álcool Etilico e suas Misturas com Água – Método de Ensaio: NBR 5992.2008^a
- ANVISA. Agência Nacional da Vigilância Sanitária. **Métodos Gerais.** Farmacopeia Brasileira, 5ª edição, v.1, 2010.
- AZZOUZ, T.; PUIGDOMÉNECH, A.; TAULER, R. Comparison between different data pré-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method. **Analytica Chimica Acta**, v. 484, p. 121-134, 2003.
- BARBOSA, F.D.; POPPI, R.J. Determination of alcohol content in beverages using short-wave near-infrared spectroscopy and temperature correction by transfer calibration procedures. **Analytical and Bioanalytical Chemistry**.v.377, pp.695-701, 2003.
- BARCELOS, L.V.F.; CARDOSO, M.G.; VILELA, F.J.; ANJOS, J.P. Teores de carbamato de etila e outros componentes secundários em diferentes cachaças produzidas em três regiões do estado de minas gerais: zona da mata, sul de Minas e Vale do Jequitinhonha
- BRASIL. Decreto número 6871, de 4 de junho de 2009. Regulamenta a lei número 8.918 de julho de 1994, sobre a padronização, a classificação, o registro, a inspeção, a produção e a fiscalização de bebidas. **Diário Oficial da República federativa do Brasil.** Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

BRASIL. Instrução Normativa número 13, de 29 de junho de 2005. Aprova o regulamento técnico para fixação dos padrões de identidade e qualidade para aguardente de cana e para cachaça. **Diário Oficial da República Federativa do Brasil**. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em 09 de maio de 2015.

CASTRITIUS, S.; KRON, A.; SCHAFER, T.; RADLE, M.; HARMS, D. Determination of Alcohol and Extract Concentration in Beer Samples Using a Combined Method of Near-Infrared (NIR) Spectroscopy and Refractometry. **Journal Agricultural Food Chemistry**, v.58, pp.12634-12641, 2010.

CARUSO, Miriam Solange Fernandes; NAGATO, Leticia Araújo Farah; ALABURDA, Janete. Evaluation of alcoholic strength and congeners in cachaças. **Rev. Inst. Adolfo Lutz (Impr.)**, São Paulo, v. 67, n. 1, 2008

CIENFUEGOS, F.; VAITSMAN, D. **Análise Instrumental**. São Paulo: Interciência, 2000, 606 p.

COZZOLINO, D.; SMYTH, H.E.; GISHEN, M. Feasibility Study on the Use of Visible and Near-Infrared Spectroscopy Together with Chemometrics To Discriminate between Commercial White Wines of Different Varietal Origins **J. Agric. Food Chem.**, v.51, p. 7703-7708, 2003.

COZZOLINO, D.; DAMBERGS, R.G.; JANIK, L.; CYNKAR, W.U.; GISHEN, W. Review: Analysis of grapes and wine by near infrared spectroscopy. **Journal of Near Infrared Spectroscopy**. v.14, pp. 279-289, 2006.

COZZOLINO, D.; KWIATKOWSKI, M.J.; DAMBERGS, R.G.; CYNKAR, W.U.; JANIK, L.J.; SKOUROUMOUNIS J.; GISHEN, M.; Analysis of elements in wine using near infrared spectroscopy and partial least squares regression. **Talanta**, v. 74, p. 711–716, 2008.

- COZZOLINO, D.; CYNKAR, W.; SHAH, N.; SMITH, P.A. Can spectroscopy geographically classify Sauvignon blanc wines from Australia and New Zealand? **Food Chemistry**, n. 126, P.673-678, 2011.
- DONG, J.J.; LI, Q.L.; YIN, H.; ZHONG, C.; HAO, J.G.; YANG, P.F.; TIAN, Y.H.; JIA, S.R. Predictive analysis of beer quality by correlating sensory evaluation with higher alcohol and ester production using multivariate statistics methods. **Food Chemistry**, v. 161, pp. 376–382, 2014.
- FERNANDES, H.L.; RAIMUNDO JR, IVO, M.; PASQUINI, C.; ROHWEDDER J. Simultaneous determination of methanol and ethanol in gasoline using NIR spectroscopy: Effect of gasoline composition. **Talanta**, v. 75, pp.804–810, 2008.
- FERREIRA, Márcia M. C. et al . Quimiometria I: calibração multivariada, um tutorial. **Quím. Nova**, v. 22,n. 5, p. 724-731,1999
- IBRAC. Mercado interno.2008. Disponível em: <http://www.ibrac.net/index.php/servicos/estatisticas/mercado-interno>. Acesso em 10 de junho de 2015.
- IBRAC. Estabelecimentos e marcas registradas no Ministério da Agricultura (Cachaça e Aguardente de cana-de-açúcar). 2013. Disponível em: <http://www.ibrac.net/index.php/servicos/estatisticas/estabelecimentosregistrados/estabelecimentos-por-estado> Acesso em: 18 de maio de 2015.
- INÁCIO, M.R.C.; LIMA, K.M.G.; LOPES, V.G.; PESSOA, J.D.C.;TEIXEIRA, G.H.A. Total anthocyanin content determination in intact açaí (*Euterpe oleracea* Mart.) and palmitero-juçara (*Euterpe edulis* Mart.) fruit using near infrared spectroscopy(NIR) and multivariate calibration.**Food Chemistry**, v.136, 1160-1164, 2013.

- ISAKSSON, T., TORMOD, N. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR Spectroscopy. **Applied Spectroscopy**, v.42, n. 7, p 1273- 1284, 1988
- KENNARD, R. W. e STONE, L. A. Computer aided design of experiments. **Technometrics**, v.11, p. 137–148, 1969.
- KOLOMIETS, O.A.; LACHENMEIER D.W.; HOFFMANN U.; SIESLER W. Quantitative determination of quality parameters and authentication of vodka using near infrared spectroscopy. **Journal of Near Infrared Spectroscopy**, v. 18, p.59–67, 2010.
- MARINHO A.V.; MACEDO, J.P.R.; SIQUEIRA, M.I.D. Avaliação da acidez volátil, teor alcoólico e de cobre em cachaças artesanais. **Estudos**, v. 36, n. 1/2, p. 75-93, 2009.
- MAPA, Ministério da Agricultura, Pecuária e Abastecimento. Indicação Geográfica-IG. (2015). Disponível em: <http://www.agricultura.gov.br/desenvolvimento-sustentavel/indicacao-geografica>. Acesso em: 12 de junho de 2015.
- MARTELO-VIDAL, M.J.; VÁZQUEZ, M. Evaluation of Ultraviolet, Visible, and Near Infrared Spectroscopy for the Analysis of Wine Compounds. **Czech Journal of Food Sciences**. v.32, n.1, pp.37-47, 2014.
- MAZUR, L. Aplicação de metodologia NIR para determinação de metilxantinas presentes na erva-mate (*Ilex paraguariensis*)80f, Dissertação (Mestrado em Engenharia de Alimentos), Universidade Federal do Paraná, Curitiba, 2012
- MENDES, L.S.; OLIVEIRA, F.C.C.; SUAREZ, P.A.Z., RUBIM, J.C. Determination of ethanol in fuel ethanol and beverages by Fourier transform

- (FT)-near infrared and FT-Raman spectrometries. **Analytica Chimica Acta** v.493, p.219–231, 2003.
- MORGANO, M.A.; FARIA, C. G.; FERRÃO, M.F.; BRAGAGNOLO, N.; FERREIRA, M.C.F. Determinação de proteína em café cru por espectroscopia NIR e regressão PLS. **Ciência e Tecnologia de Alimentos**. v. 25, n.1. p. 25-31, 2005.
- NAES, T.; ISAKSSON, T.; FEARN T.; DAVIES, T. *A User-friendly Guide to Multivariate Calibration and Classification*, **NIR Publications**, Chichester, UK, 2002.
- NICOLAI, B. M.; BEULLENS, K.; BOBELYN, E.; PEIRS, A.; SAEYS, W.; THERON, K. I. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. **Postharvest Biology and Technology**, v. 46, p.99-118, 2007.
- ODELLO, L.; BRACESCHI, G.P.; SEIXAS, F.R.F.; SILVA, A.A.; GALINARO, C, A.; FRANCO, D.W. Avaliação sensorial de cachaça. **Química Nova**, v. 32, n. 7, p. 1839-1844, 2009 .
- PASQUINI, C. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. **Journal of Brazilian Chemistry Society**, v.14, n.2, p.198-219, 2003
- SAVITSKY, A., GOLAY, M.J.E. Smoothing and differentiation by simplified least squares procedures. **Analytical Chemistry**, v.36, p.1627-1632, 1964.
- SIMÕES, SILVA F.S., Potencialidades da Espectroscopia de infravermelho com Transformadas de Fourier na caracterização de aguardentes víquicas. **Dissertação de mestrado**. Universidade de Aveiro, 2008.

- SOUZA, J. S.; FERRÃO, M. F. . Aplicações da espectroscopia no infravermelho no controle de qualidade de medicamentos contendo diclofenaco de potássio. Parte I: Dosagem por regressão multivariada. **RBCF. Revista Brasileira de Ciências Farmacêuticas**, São Paulo, v. 42, n. 3, p. 437-445, 2006.
- SOUZA, A.M.; POPPI, R.J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: um tutorial, parte I. **Química Nova**, v. 35, n. 1, pp. 223-229, 2012
- SMYTH, H.E.; COZZOLINO, D.; CYNKAR, W.U.; DAMBERGS, R.G.; SEFTON, M.; GISHEN, M. Near infrared spectroscopy as a rapid tool to measure volatile aroma compounds in Riesling wine: possibilities and limits. **Analytical and Bioanalytical Chemistry**, v.308, pp.1911-1916, 2008.
- URBANO-QUADRADO, M.;LUQUE de CASTRO, M.D.; PEREZ, P.M.; GÓMEZ NIETO, M.A. Near infrared reflectance spectroscopy and multivariate analysis in enology Determination or screening of fifteen parameters in different types of wines. **Analytica Chimica Acta**, v.527, pp.81-88, 2004.
- URBANO-QUADRADO, M.;LUQUE de CASTRO, M.D.; PEREZ, P.M.; GÓMEZ NIETO, M.A. Comparison and joint use of near infrared spectroscopy and Fourier transform mid infrared spectroscopy for the determination of wine parameters. **Talanta**, v.66, pp.218-224, 2005.

ANEXO

Anexo 1. Nome, estado produtor e teor alcoólico das aguardente de cana-de-açúcar de São Paulo e Minas Gerais.

Amostra	Nome	Estado	Teor de etanol (%v/v)
1	Sabor da Estância	São Paulo	32,3
2	Cachaça Barra Grande	São Paulo	45,6
3	Ziegen Hof Microdestilada	São Paulo	42,0
4	Cachaça Fuzuê Branca	São Paulo	40,4
5	Cachaça Cortarelli	São Paulo	42,4
6	Cachaça Artesanal	São Paulo	34,3
7	Cachaça do Santo	São Paulo	36,3
8	Portal da Baraceia	São Paulo	36,3
9	Schiermann Cachaça Prata	São Paulo	39,3
10	Cachaça Quizumba na Senzala	São Paulo	39,3
11	Cachaça R&R	São Paulo	41,4
12	Cachaça Formosa	São Paulo	35,3
13	Cachaça Bueno	São Paulo	41,4
14	Cachaça de Porto	São Paulo	32,3
15	Cachaça São Francisco	São Paulo	41,9
16	Cachaça Engenho Spolador	São Paulo	41,9

17	Aguardente de Cana Sertaneja	São Paulo	41,4
18	Cachaça do Horto	São Paulo	33,3
19	Cachaça Boa Vista	São Paulo	40,4
20	Cachaça Dujac	São Paulo	38,8
21	Cachaça Empório da cana	São Paulo	39,8
22	Cachaça Boa Esperança	São Paulo	39,8
23	Nossa Cachaça	São Paulo	36,3
24	Cachaça Zé da Varge	São Paulo	30,3
25	Cachaça Bassi	São Paulo	41,4
26	Cachaça Mulata	São Paulo	45,6
27	Saudades de Matao	São Paulo	39,3
28	Cachaça Morano	São Paulo	33,3
29	Cachaça Cabriuvinha	São Paulo	38,9
30	Cachaça Chora Menina	São Paulo	42,0
31	Cachaça Pequeno Engenho	São Paulo	36,8
32	Cachaça João Roberto	São Paulo	36,8
33	Fina Flor	São Paulo	36,8
34	Adega do Italiano	São Paulo	41,0
35	Cachaça Senhora	São Paulo	37,9
36	Cachaça Força Oculta	São Paulo	36,8
37	Cachaça do Santo	São Paulo	42,0
38	Aguardente de Cana Saúva	São Paulo	38,9
39	Cachaça Sermaria	São Paulo	34,8

40	Engenho São Luiz	São Paulo	41,0
41	Cachaça Fuzuê Envelhecida	São Paulo	41,4
42	Sagatiba	São Paulo	38,3
43	Portal da Baraceia	São Paulo	37,9
44	Portal da Baraceia Envelhecida	São Paulo	40,4
45	Cachaça Polacchini	São Paulo	37,9
46	Cachaça Fuzuê	São Paulo	40,4
47	Cachaça Polacchini	São Paulo	37,9
48	Portal da Baraceia Envelhecida	São Paulo	40,4
49	Portal da Baraceia	São Paulo	40,4
50	Aguardente de Cana JP	São Paulo	42,0
51	Cachaça Morano Prata	São Paulo	43,0
52	Cachaça Empório da cana	São Paulo	44,1
53	Aguardente Bico Doce	São Paulo	45,1
54	Sabor da Estância Descansada	São Paulo	43,9
55	Cachaça Itupeva	São Paulo	43,5
56	Velha Engenhoca	São Paulo	46,1
57	Cachaça Mazzaropi	São Paulo	36,3
58	Cachaça Ribeiro da Silva	São Paulo	42,4
59	Sabor do Andes	São Paulo	43,5
60	Cachaça das Torres	São Paulo	42,4
61	Sebastiana	São Paulo	42,4
62	Cachaça Bebedouro	São Paulo	44,6

63	Sabor da Estância	São Paulo	36,3
64	Amostras cedidas pelo LAFQA	São Paulo	38,3
65	Amostras cedidas pelo LAFQA	São Paulo	46,6
66	Amostras cedidas pelo LAFQA	São Paulo	47,6
67	Amostras cedidas pelo LAFQA	São Paulo	45,6
68	Amostras cedidas pelo LAFQA	São Paulo	44,6
69	Amostras cedidas pelo LAFQA	São Paulo	47,6
70	Cachaça Terra de Ouro	Minas Gerais	42,0
71	Bocaína	Minas Gerais	35,3
72	Cachaça da Pedra	Minas Gerais	41,4
73	Famigerada	Minas Gerais	50,1
74	Cachaça Magestade	Minas Gerais	39,8
75	Cachaça Tabúa	Minas Gerais	41,9
76	Cachaça Puricana	Minas Gerais	39,9
77	Cachaça Fascinação	Minas Gerais	39,9
78	Rainha Vale	Minas Gerais	43,0
79	Cachaça do Dedé	Minas Gerais	44,1
80	Vale Verde Prata	Minas Gerais	44,1
81	Cachaça Coqueiro	Minas Gerais	46,2
82	Cachaça Terra Forte	Minas Gerais	43,0
83	Cachaça Dona Branca	Minas Gerais	45,1
84	Ouro 1	Minas Gerais	44,1
85	Ouro Mineiro Umburama	Minas Gerais	45,1

86	Cachaça Taverna de Minas	Minas Gerais	42,0
87	Cachaça Lua Nova	Minas Gerais	45,1
88	Arte Mineira	Minas Gerais	45,1
89	Cachaça Aroeirinha	Minas Gerais	43,0
90	Si Sobra Nois Vendi	Minas Gerais	42,0
91	Confraria Chico Mineiro	Minas Gerais	43,0
92	Velha Aroeira	Minas Gerais	44,1
93	Cachaça Rainha da Cana	Minas Gerais	44,1
94	Cachaça Januária	Minas Gerais	44,1
95	Buriti	Minas Gerais	42,5
96	Rainha do Vale	Minas Gerais	44,1
97	Sagarana	Minas Gerais	46,6
98	Porteirinha	Minas Gerais	43,5
99	Rainha do Norte	Minas Gerais	45,1
100	Embaia Saia	Minas Gerais	52,7
101	Cachaça Boutique	Minas Gerais	42,4
102	Cachaça Raiz	Minas Gerais	47,6
103	Balalaica	Minas Gerais	51,7
104	Cristalina do Picão	Minas Gerais	47,6
105	Amostras cedidas pelo LAFQA	Minas Gerais	46,2
106	Amostras cedidas pelo LAFQA	Minas Gerais	46,2
107	Amostras cedidas pelo LAFQA	Minas Gerais	44,6
108	Amostras cedidas pelo LAFQA	Minas Gerais	35,8

109	Amostras cedidas pelo LAFQA	Minas Gerais	46,4
110	Amostras cedidas pelo LAFQA	Minas Gerais	38,3
111	Amostras cedidas pelo LAFQA	Minas Gerais	46,6
112	Amostras cedidas pelo LAFQA	Minas Gerais	43,5
113	Amostras cedidas pelo LAFQA	Minas Gerais	42,4
114	Amostras cedidas pelo LAFQA	Minas Gerais	43,7
115	Amostras cedidas pelo LAFQA	Minas Gerais	44,6
116	Amostras cedidas pelo LAFQA	Minas Gerais	44,6
117	Amostras cedidas pelo LAFQA	Minas Gerais	44,6
118	Amostras cedidas pelo LAFQA	Minas Gerais	42,4
119	Amostras cedidas pelo LAFQA	Minas Gerais	46,6
120	Amostras cedidas pelo LAFQA	Minas Gerais	43,5
121	Amostras cedidas pelo LAFQA	Minas Gerais	45,0
122	Amostras cedidas pelo LAFQA	Minas Gerais	45,0
123	Amostras cedidas pelo LAFQA	Minas Gerais	46,8
124	Amostras cedidas pelo LAFQA	Minas Gerais	42,7
125	Amostras cedidas pelo LAFQA	Minas Gerais	57,1
126	Amostras cedidas pelo LAFQA	Minas Gerais	47,8
127	Amostras cedidas pelo LAFQA	Minas Gerais	52,7
128	Amostras cedidas pelo LAFQA	Minas Gerais	47,8
129	Amostras cedidas pelo LAFQA	Minas Gerais	46,8
130	Amostras cedidas pelo LAFQA	Minas Gerais	49,1
131	Amostras cedidas pelo LAFQA	Minas Gerais	48

132	Amostras cedidas pelo LAFQA	Minas Gerais	49,1
133	Amostras cedidas pelo LAFQA	Minas Gerais	41,9
134	Amostras cedidas pelo LAFQA	Minas Gerais	43,9
135	Amostras cedidas pelo LAFQA	Minas Gerais	46,0
136	Amostras cedidas pelo LAFQA	Minas Gerais	46,0
137	Amostras cedidas pelo LAFQA	Minas Gerais	46,0
138	Amostras cedidas pelo LAFQA	Minas Gerais	39,8
139	Amostras cedidas pelo LAFQA	Minas Gerais	37,8
