

PROPOSIÇÃO DE ALTERNATIVAS DE TREINAMENTO DE DEGUSTADORES EM UM PAINEL SENSORIAL UTILIZANDO ÍNDICES DE DESEMPENHO

PROPOSITION OF TRAINING ALTERNATIVES FOR TASTERS IN A SENSORY PANEL USING PERFORMANCE INDEXES

¹ Luiz Antônio Nonenmacher Júnior

² Lucas Louzada Pereira

^{3*} Michel José Anzanello

¹Universidade Federal do Rio Grande do Sul – UFRGS. Email:ljuniornone@gmail.com

²Instituto Federal do Espírito Santo. E-mail:lucas.pereira@ifes.edu.br

³Universidade Federal do Rio Grande do Sul – UFRGS. E-mail: anzanello@producao.ufrgs.br

*Autor de correspondência

Artigo submetido em 24/09/2020, aceito em 28/11/2020 e publicado em 23/12/2020.

Resumo: Painéis sensoriais são muito utilizados na indústria de alimentos para controle de qualidade, classificação de produtos em categorias e pesquisas com diversos fins. Um painel sensorial é composto por um grupo de degustadores que confere notas a diferentes atributos sensoriais de amostras de produtos; em tal processo, espera-se que o painel seja composto por membros que avaliem produtos iguais de forma semelhante e consigam discriminar produtos diferentes. O objetivo deste artigo é aplicar e comparar três abordagens de treinamento de degustadores, com o objetivo de melhorar o desempenho do painel, medido através do Índice Alfa de Ledauphin. As três abordagens comparadas foram (i) treinamento de alguns degustadores para todos os atributos (treinamento de degustadores), (ii) treinamento de alguns atributos para todos os degustadores (treinamento de atributos), e (iii) combinação dessas duas técnicas (alguns atributos para alguns degustadores). O método combinado foi o que apresentou os melhores resultados, com uma necessidade de treinamento, para um mesmo desempenho do painel, 18,7% menor que o treinamento de degustadores e 51,4% menor que o treinamento de atributos.

Palavras-chave: painéis sensoriais; treinamento de degustadores; índice alfa de Ledauphin.

Abstract: Sensory panels are commonly utilized in the food industry for purposes of quality control, classification, and research. A sensory panel is composed of a group of graders that evaluate some products samples in a group of attributes. It is expected that the graders which compose a panel can evaluate equally the same products and capable of discriminating different ones. The objective of this article is to apply and compare three methods of training of graders that aim to increase the panel performance, evaluated by the Alpha Index of Ledauphin. The three compared methods are: (i) training of some graders in all attributes (graders training), (ii) training of some attributes for all graders (attributes training) and (iii) the combination of this techniques (some attributes for some graders). The combined method is the one that presented the best results, with a training need, for the same panel performance, 18.7% lower than the graders training and 51.4% lower than the attributes training.

Keywords: sensory panels, graders training, alpha index of Ledauphin.

1 INTRODUÇÃO

Na indústria de alimentos é comum o uso de painéis sensoriais para se avaliar um grupo de características sensoriais de um produto (WILDERJANS; CARIU, 2016) com vistas ao controle de qualidade e classificação dos produtos em faixas de atribuição de valor (por exemplo, faixas *premium* e regular), dentre outros (BANFIELD; HARRIES, 1975). Um painel é formado por um grupo de degustadores treinados que conferem notas a amostras de produtos (com a presença ou não de repetições) descritos por um conjunto de atributos, sendo os resultados tipicamente expressos numa escala numérica (ANZANELLO et al., 2011).

Para assegurar que um painel tenha um bom desempenho, é fundamental que seus integrantes sejam simultaneamente capazes de avaliar produtos de forma semelhante entre si e de discriminar produtos diferentes. Existem diferentes técnicas para mensurar o desempenho de um painel, sendo que nesse artigo é utilizado o Índice Alfa de Ledauphin (LEDAUPHIN et al., 2006). Tal índice é baseado no autovetor β , obtido a partir do maior autovalor de uma matriz S gerada a partir da comparação de matrizes que descrevem as notas individuais de cada degustador do painel sensorial. Essa técnica gera dois principais resultados: um índice que mensura a concordância do painel como um todo e uma série de valores que avalia o desempenho de cada um dos degustadores.

Este artigo propõe a comparação de três cursos de ação distintos com vistas ao treinamento de degustadores, monitorando seu resultado através do Alfa de Ledauphin do painel. Um desses métodos, mais utilizado na prática, consiste em selecionar um ou mais degustadores e treiná-lo em todos os atributos. Um dos métodos propostos consiste em selecionar um ou mais atributos e treiná-lo dentre todos os degustadores. Por fim, o último método proposto consiste em conduzir um

treinamento de forma combinada, selecionando apenas alguns atributos de alguns degustadores para realizar o treinamento. O impacto do treinamento sobre as notas de cada degustador será simulado com a utilização de uma matriz ponderada de notas e um índice de aprendizado.

Este artigo está organizado em cinco seções. Após a introdução, a seção 2 apresenta algumas técnicas utilizadas para avaliar o desempenho de painéis sensoriais, com enfoque no Índice Alfa de Ledauphin. Já a seção 3 apresenta o método proposto para o atingimento dos objetivos deste artigo, enquanto os resultados da aplicação e uma análise dos mesmos compõem a seção 4. Por fim, a seção 5 conclui o artigo, retomando os objetivos apresentados na introdução e sugerindo pesquisas futuras na área.

2 REFERENCIAL TEÓRICO

Essa seção inicia com a definição de painéis sensoriais e a apresentação de algumas técnicas utilizadas para avaliar o desempenho de tais painéis. Na sequência, as premissas do Índice Alfa de Ledauphin, utilizado neste trabalho, são descritas em maiores detalhes.

2.1 ABORDAGENS PARA AVALIAÇÃO DE DESEMPENHO DE PAINÉIS SENSORIAIS

Um painel sensorial que avalia um produto alimentício é composto por um grupo de degustadores treinados que graduam, em geral numa escala numérica, um conjunto de atributos para diferentes amostras de um grupo de produtos. Para que o painel sensorial tenha sucesso em sua implementação, é desejado que os degustadores sejam capazes de (i) repetir consistentemente suas avaliações ao longo

do tempo, (ii) avaliar produtos de maneira semelhante aos demais membros do painel, e (iii) discriminar produtos com características sensoriais diferentes (LATREILLE et al., 2006). Existem diferentes metodologias para fazer a avaliação de desempenho de um painel e de seus membros, conforme apresentadas a seguir.

Uma das abordagens básicas para avaliação de desempenho de painéis sensoriais é utilizar a variância dos dados para calcular índices numéricos que têm sua significância testada através de métodos estatísticos como a ANOVA ou o Teste T. Na sequência, os índices podem formar um espaço bidimensional e os degustadores, atributos ou produtos podem ser analisados graficamente. Um conjunto desses índices é utilizado por Rossi (2012) para avaliar degustadores. São utilizados dois indicadores: índice de repetitividade (*repeatability*), que avalia a capacidade do degustador de avaliar repetições do mesmo produto de forma igual, e o índice de reprodutibilidade (*reproducibility*), que mede a capacidade de um degustador, na média, ter uma mesma avaliação que outros membros do painel. Este é um dos métodos comparados por Pinto et al. (2014) para classificar degustadores de acordo com seu desempenho.

Outro conjunto de índices foi proposto por Bi (2007) para avaliar a qualidade de um painel sensorial. Ele utiliza o coeficiente de concordância e o de confiabilidade (*agreement and reliability*), sendo que o coeficiente de concordância avalia o consenso entre os diferentes degustadores para um mesmo produto e o de confiabilidade mede a capacidade de discriminação do painel. Com propósitos semelhantes, Talsma (2016) utiliza a teoria da generalização para avaliar o desempenho de um painel sensorial e de seus membros. A variância é decomposta em sete elementos (avaliando a variância do produto, do degustador, da réplica e suas interações de segunda e terceira ordem); o

indicador de confiabilidade, por sua vez, é calculado através da divisão da variância de cada componente pela variância total.

Uma técnica frequentemente utilizada para avaliação de desempenho de painéis é a análise de componentes principais (PCA). A técnica visa a transformar as variáveis originais em um conjunto reduzido de componentes principais (combinações lineares das variáveis originais) com o objetivo de explicar o máximo possível da variabilidade dos dados originais (BORGOGNONE et al., 2001). O PCA forma a base da análise de consonância (*consonance analysis*), técnica proposta por Dijksterhuis (1995), que avalia o desempenho dos degustadores através da comparação do primeiro autovalor com os demais. A análise de consonância também é um dos métodos utilizados por Pinto et al. (2014) para classificar degustadores de acordo com seu desempenho.

O PCA pode ser utilizado conjuntamente com outras técnicas com vistas à avaliação de degustadores. Anzanello et al. (2011) utilizaram PCA e *k-Nearest Neighbor* (*k-NN*) para encontrar um conjunto ótimo de atributos e degustadores que melhor diferenciam as amostras analisadas, enquanto que Sinesio et al. (1990) usaram PCA, análise generalizada de Procustes (GPA) e regressão de mínimos quadrados parciais (PLSR) para avaliar a capacidade de um grupo de degustadores para avaliar a rancidez oxidativa em amostras de alimentos. Por fim, Sivertsen e Risvik (1994) combinaram PCA com GPA e ANOVA para analisar diferenças entre vinhos produzidos em regiões diferentes e avaliar o desempenho dos avaliadores.

A Análise Generalizada de Procustes ou *Generalized Procustes Analysis* (GPA) é uma técnica estatística que compara o formato de diferentes objetos. Essa comparação é feita através da translação dos dados de cada degustador para uma origem comum (neutralizando o efeito de notas individuais nos níveis da escala), de uma rotação e, por fim, de uma contração ou

expansão (para ajustar diferenças de escala entre os degustadores) (SINESIO et al., 1990). Essa técnica pode ser utilizada para comparar degustadores, revelando aqueles que estão percebendo as amostras diferentes do restante do painel (BANFIELD; HARRIES, 1975), ou para comparar o desempenho de diferentes painéis avaliando produtos iguais (MCEWAN et al., 2002).

Outras técnicas utilizam conceitos de clusterização para avaliar painéis sensoriais. A clusterização de variáveis em torno de componentes latentes cria um grupo de *clusters* descrito por uma variável latente de tal forma que a variância em cada grupo seja minimizada. Ao analisar o componente latente de cada *cluster*, pode-se comparar os atributos de determinado produto (VIGNEAU et al., 2003). No entanto, essa análise só é feita em duas dimensões (atributos \times produtos), sendo necessário fazer a média entre os degustadores, perdendo essa informação. Para suprir esta carência, Wilderjans e Cariou (2016) propuseram o método CLV3W, uma extensão do CLV para dados em três dimensões (degustadores \times atributos \times produtos); tais autores usaram esta técnica para simultaneamente comparar os atributos e o desempenho de degustadores.

Outro modelo utilizado para espaços de mais de duas dimensões é a análise de fatores paralelos, *Parallel factor analysis* (PARAFAC). Trata-se de um método de decomposição multivariado originado no campo de psicometria e que pode ser visto como uma generalização da análise de componentes principais (PCA) para ordens superiores (BRO, 1997). O modelo gera parâmetros que representam a variabilidade nas dimensões de interesse (por exemplo, produtos, degustadores e atributos) e permitem sua análise. Cocchi et al. (2006) utilizaram o PARAFAC para analisar os efeitos do envelhecimento de amostras de Aceto Balsâmico Tradizionale di Modena (ABTM) sobre suas características sensoriais.

Por fim, também encontram espaço na literatura abordagens apoiadas no cálculo de um índice único para avaliar os desempenhos de um painel e de cada degustador. Um desses índices é o Alfa de Cronbach descrito por Pinto et al. (2014) e baseado no coeficiente Alfa de Cronbach (CRONBACH, 1951), o qual é calculado utilizando a variância total de cada avaliação e a covariância média entre avaliações. Já Ledauphin et al. (2006) propuseram um coeficiente calculado através do autovetor correspondente ao maior autovalor de uma matriz de similaridade obtida através da comparação par a par das avaliações normalizadas de cada degustador, sendo que este indicador varia entre +1 e -1, com +1 representando a consistência total (do degustador ou do painel) e -1 a discordância total. Por se tratar do índice abordado neste estudo, a próxima seção traz a fundamentação matemática e premissas do índice.

2.2. ALFA DE LEDAUPHIN

Dado um painel sensorial no qual I degustadores avaliam N produtos em M atributos, pode-se montar I matrizes do tipo \mathbf{Y}_i com cada matriz tendo uma dimensão $n \times m$. Com essa notação, o elemento Y_{im} representa a avaliação feita pelo degustador i em um produto n em relação ao atributo m .

O primeiro passo para obtenção do índice é remover os vieses das avaliações dos degustadores. Para remover o efeito da utilização de diferentes níveis da escala (tipicamente verificado em avaliações), o primeiro procedimento consiste em centrar na origem cada coluna referente a cada degustador. Isso é realizado subtraindo a média de cada coluna, gerando as matrizes centradas $\mathbf{Y}C_1, \mathbf{Y}C_2, \dots, \mathbf{Y}C_I$. Outra possível fonte de erro entre os degustadores é uso de intervalos diferentes da escala. Para corrigir isso é necessário multiplicar cada matriz centrada por um escalar θ_i , tal que $\theta_i = \frac{1}{\sqrt{f_i}}$, onde f_i é a soma dos quadrados de todos os elementos de cada matriz centrada $\mathbf{Y}C_i$.

Com a multiplicação de cada matriz $Y C_i$ por θ_i , geram-se as I matrizes corrigidas Z_i . O próximo passo é comparar cada par de matrizes Z_i através do cálculo de d_{i1i2} , como na equação (1).

$$d_{i1i2} = \text{traço}(Z_{i1}^T Z_{i2}) \quad (1)$$

Com o cálculo de d_{i1i2} , pode-se definir a similaridade entre cada par de degustadores por s_{i1i2} , com $s_{i1i2} = \frac{1+d_{i1i2}}{2}$, podendo d_{i1i2} ser entendido como a correlação entre as matrizes Z_{i1} e Z_{i2} consideradas como vetores pelo rearranjo dos termos em uma só coluna. Por isso, d tem características similares ao coeficiente de correlação, variando entre -1 e +1, com s variando entre 0 e 1. Quando $s = 0$ ($d = -1$), os degustadores estão em total desacordo, de tal forma que $Z_{i1} = -Z_{i2}$. De maneira semelhante, $s = 1$ ($d = +1$) ocorre quando os degustadores estão em total acordo, tal que $Z_{i1} = Z_{i2}$. Por fim, quando não há correlação entre Z_{i1} e Z_{i2} , $s = 0,5$ ($d = 0$).

Através dos valores de s pode-se montar a matriz S , de dimensão $I \times I$, com cada elemento ($i1, i2$) da matriz sendo igual ao valor de s_{i1i2} . Essa matriz é simétrica, pois $s_{i1i2} = s_{i2i1}$, e cujos termos da diagonal principal são iguais a 1 (pois $s_{i1i2} = 1$ sempre que $i1 = i2$). Deve-se então extrair de S o autovetor $\beta = (\beta_1, \beta_2, \dots, \beta_I)^T$ correspondente ao maior autovalor de S . Esse vetor β deve ser normalizado, de forma que a soma de todos os seus elementos β_i seja igual a 1. Os coeficientes β_i variam entre 0 e 1 e representam o grau de concordância do degustador i com o restante do painel. Assim, é possível definir a matriz C , que representa uma média ponderada de notas do painel, através da equação (2).

$$C = \sum_{i=1}^I \beta_i Z_i \quad (2)$$

A partir do cálculo da matriz C , é possível, através da equação (3), calcular um indicador α_i que avalia o desempenho de cada degustador.

$$\alpha_i = \frac{\text{traço}(Z_i^T C)}{\sqrt{\text{traço}(C^T C)}} \quad (3)$$

O coeficiente α_i pode ser entendido como o coeficiente de correlação entre as matrizes Z_i e C consideradas como vetores e, portanto, pode variar entre -1 e +1. Um valor negativo de α_i indica que o degustador i julga os produtos de forma substancialmente diversa dos demais degustadores. De forma contrária, quanto mais próximo α_i estiver de 1, mais o degustador está em acordo com o grupo. Através da média dos valores de α_i pode-se obter um índice de desempenho global para o painel, γ . Quanto maior o valor de γ , maior é a concordância do painel, sendo que este índice só pode ser igual a 1 se, e somente se, todas as matrizes Z_i forem iguais. Além de avaliar o desempenho do painel na sua configuração original, γ pode também ser utilizado para avaliar como se altera o desempenho do painel com alterações nos dados (retirada de degustador, de atributos ou mudanças de notas).

3 PROCESSOS METODOLÓGICOS: MATERIAIS E MÉTODOS

O objetivo deste estudo é analisar o efeito esperado do treinamento de degustadores sobre o desempenho do painel sensorial através do Índice Alfa de Ledauphin. Foram comparadas três alternativas de treinamento: treinamento de um atributo específico para todos os degustadores (treinamento dos atributos), treinamento de um determinado degustador para todos os atributos (treinamento dos degustadores) e treinamento conjunto de um atributo específico para determinado degustador (treinamento combinado).

Como notação, considera-se um painel sensorial no qual I degustadores avaliam N produtos em M atributos, gerando I matrizes Y_i de dimensão $N \times M$ nas quais as linhas representam os produtos e as colunas representam os atributos. Para o cálculo dos efeitos do treinamento é necessário gerar uma matriz de notas

ponderada, L , calculada pela ponderação de cada uma das I matrizes de notas de cada degustador pelos autovetores normalizados, $\beta = (\beta_1, \beta_2, \dots, \beta_I)T$, conforme a equação (4).

$$L = \sum_{i=1}^I \beta_i Y_i \quad (4)$$

Percebe-se que o cálculo de L é análogo ao cálculo da matriz C apresentada na seção 2.2.1, com a diferença de que C utiliza as matrizes corrigidas Z_i , ao passo que L utiliza as matrizes originais Y_i . Para simular o efeito do treinamento de um degustador, parte-se da premissa que o treinamento tende a aproximar a nota do degustador da nota ponderada do painel, sendo a magnitude dessa aproximação dada pela qualidade do treinamento. Definindo it como índice de treinamento (variando entre 0 e 1) e W como a matriz de avaliações após treinamento, obtêm-se W através da equação (5):

$$W = (1 - it) * Y + it * L \quad (5)$$

Pela equação 5, pode-se verificar que quanto maior o valor de it , mais a nota do degustador se aproxima da nota ponderada do painel após treinamento. Para a comparação do desempenho das diferentes alternativas de treinamento, este artigo testará os valores de it iguais a 1, 0,75, 0,5 e 0,25 (quando menor it , menor a assimilação do treinamento por parte do degustador).

Definido o procedimento para incorporação do aprendizado do degustador, parte-se para as etapas remanescentes do método, as quais apoiam-se no: (i) cálculo do Alfa de Ledauphin inicial; (ii) treinamento dos atributos; (iii) treinamento dos degustadores; (iv) treinamento combinado, e (v) comparação dos resultados. Vale ressaltar que as etapas (ii), (iii) e (iv) são realizadas de forma paralela e não sequencial; depois de simular o treinamento dos atributos na etapa (ii), retoma-se o banco inicial para a etapa (iii) e o mesmo ocorre para a etapa (iv). Tais etapas são agora detalhadas.

A primeira etapa tem como objetivo calcular o Alfa de Ledauphin, índice que

será utilizado como medida de desempenho do painel e que é obtido através da média dos índices α_i de cada degustador. Além disso, nessa etapa é calculada a matriz L , que objetiva simular o efeito do treinamento sobre as notas dos degustadores. Os índices α_i e β_i obtidos nessa seção são calculados através dos procedimentos descritos na seção 2.2. Essa etapa cumpre a função de estabelecer valores de referência para as próximas análises.

A segunda etapa visa a avaliar o efeito de treinamento dos atributos para todos os degustadores e sugerir a melhor sequência de treinamento (sequência de atributos que maximiza o desempenho do painel para uma determinada quantidade de treinamentos). O treinamento de um atributo m e a consideração das potenciais melhorias em termos de homogeneização dos pareceres emitidos pelos degustadores treinados é realizado pela substituição, para cada degustador, das notas conferidas a um atributo, conforme a equação (6), na qual it (índice de treinamento) varia entre 0 e 1 e representa o quanto o treinamento do painel foi efetivo (quando it é 1, a nota do degustador torna-se igual a nota ponderada do painel):

$$W_m = (1 - it) * Y_m + it * L_m \quad (6)$$

Inicialmente faz-se o treinamento de cada um dos M atributos, um de cada vez, avaliando o desempenho desse treinamento pelo cálculo do Índice Alfa do painel e sua comparação com o índice inicial obtido na etapa (i). Define-se então o atributo que gerou o maior alfa como m_1 , o qual deve ser o primeiro a ser treinado nos degustadores para maximizar o desempenho do painel.

Na sequência, substitui-se permanentemente o atributo m_1 nas matrizes de cada degustador, e realiza-se o treinamento de cada um dos $M - 1$ atributos restantes. Cada treinamento tem seu desempenho avaliado e define-se m_2 como o atributo que gerou o melhor resultado. Dessa forma, m_1 e m_2 representam os dois

atributos que devem ser treinados para se obter a maior consistência do painel. A próxima etapa então substitui os atributos m_1 e m_2 nas matrizes Y_i , e avalia o treinamento dos $M - 2$ atributos restantes, gerando, de forma análoga aos casos anteriores, m_3 . Tal processo iterativo é realizado M vezes, gerando uma sequência de treinamento dos atributos m_1, m_2, \dots, m_M .

A terceira etapa, realizada de forma paralela à segunda etapa, tem o objetivo de avaliar o efeito de se treinar um degustador para todos os atributos e sugerir a melhor sequência de degustadores a serem treinados. O treinamento de um degustador i consiste em substituir todas as suas avaliações por aquelas após treinamento, conforme equação (7)

$$W_i = (1 - it) * Y_i + it * L \quad (7)$$

Treinando e avaliando o desempenho de cada um dos I degustadores, define-se o degustador que gerou o melhor desempenho como i_1 . A partir disso, substitui-se a matriz Y_{i_1} por W_{i_1} de forma definitiva e calcula-se o efeito do treinamento dos $I - 1$ degustadores restantes, obtendo i_2 . O processo então continua por I etapas, de forma análoga ao apresentado na segunda etapa, obtendo-se ao final uma sequência de degustadores a serem treinados i_1, i_2, \dots, i_I .

A etapa (iv) é similar às etapas (ii) e (iii). No entanto, ao invés de se treinar um atributo para todos os degustadores ou treinar um degustador em todos os atributos, treina-se, a cada passo, apenas um atributo para um determinado degustador. O treinamento de um atributo m para um degustador i é feito através da equação (8):

$$W_{im} = (1 - it) * Y_{im} + it * L_m \quad (8)$$

De forma semelhante às etapas anteriores, escolhe-se então a combinação de atributo e degustador que gerou o maior Alfa, substitui-se esse valor permanentemente e testam-se as $(M \times I - 1)$ combinações restantes. Esse processo ocorre

$M \times I$ vezes, até que todos os valores tenham sido substituídos.

Por fim, a etapa (v) tem como objetivo comparar os resultados das etapas (ii), (iii) e (iv), avaliando o tipo de treinamento que resulta no melhor resultado (avaliado através do aumento do Alfa de Ledauphin). Os resultados serão comparados para quatro valores de índice de treinamento: 1, 0,75, 0,5 e 0,25. Para que a comparação seja adequada, serão comparados também o número de valores substituídos nas matrizes X para cada treinamento; o treinamento de um atributo para todos os degustadores substitui $I \times N$ valores, o treinamento de um degustador para todos os atributos substitui $M \times N$ valores e o treinamento de apenas um atributo para um degustador envolve N valores.

4 RESULTADOS E DISCUSSÃO

O banco de dados analisado refere-se a um painel sensorial de amostras de café especiais processados por via-úmida. Nele, 11 degustadores avaliam 9 produtos descritos por 10 atributos: fragrância/aroma, uniformidade, ausência de defeitos, doçura, sabor, acidez, corpo, finalização, equilíbrio e final, atribuindo notas de 0 a 10, numa escala do tipo maior-é-melhor. Os 11 degustadores são certificados pelo Coffee Quality Institute (CQI) e o painel sensorial seguiu o protocolo da Specialty Coffee Association of America (SCA, 2018).

Na sequência são apresentados os resultados obtidos com a aplicação do método descrito na seção 2.3. Essa seção segue uma divisão por subseções de acordo com as cinco etapas do método proposto.

4.1 CÁLCULO DO ALFA DE LEDAUPHIN INICIAL

Aplicando os procedimentos descritos na etapa (i), obteve-se para cada degustador, o seu α_i , e o autovetor associado, β_i , utilizado para cálculo da

matriz ponderada **Z**, conforme apresentado na Tabela 2.1.

Tabela 2.1: Índices Iniciais

Degustador	Alfa	Beta
1	0,748	0,097
2	0,768	0,098
3	0,529	0,088
4	0,539	0,089
5	0,496	0,087
6	0,763	0,098
7	0,674	0,094
8	0,494	0,087
9	0,538	0,089
10	0,664	0,094
11	0,282	0,078
Geral	0,590	-

Através do Alfa calculado para cada degustador, encontra-se o Alfa médio do painel, 0,590, número que será utilizado como parâmetro de referência nas etapas seguintes. Pode-se perceber que o degustador 11 tem um desempenho substancialmente baixo quando comparado aos demais integrantes do painel (Alfa = 0,282), sugerindo a falta de concordância do mesmo com o restante do painel. Por outro lado, o degustador com melhor desempenho é o degustador 2, com um índice de 0,763. Analisando os valores de Beta, pode-se perceber que eles são proporcionais ao valor de Alfa, com o pior avaliador apresentando o menor Beta e o melhor apresentando o Beta mais elevado. Isso torna a ponderação menos sujeita a erros causados por degustadores inconsistentes.

4.2 TREINAMENTO DOS ATRIBUTOS

Seguindo os procedimentos da etapa (ii), foi realizado o treinamento dos atributos com um total de 10 etapas, fixando a cada etapa a configuração anterior e selecionando o melhor atributo a ser gerado. A Tabela 2.2 traz o Alfa obtido com o

treinamento de cada atributo a cada etapa, utilizando *it* igual a 1. Depois que um atributo é escolhido como o melhor de determinada etapa, ele não é mais treinado, o que é representando por “-” na tabela.

Com base em tais resultados, percebe-se que o Alfa cresce a cada etapa, chegando, como esperado, ao valor 1 no final do treinamento. A ordem ótima de atributos a ser treinada é: doçura, ausência de defeitos, acidez, sabor, uniformidade, final, finalização, fragrância/aroma e equilíbrio. Os resultados da Tabela 2.2 podem ser utilizados pelo gestor do painel sensorial na obtenção de um Alfa-meta ou para treinar apenas um número definido de atributos (dada uma limitação de recursos).

No caso de se buscar um Alfa-meta, encontra-se na tabela o primeiro valor de Alfa maior que a meta e faz-se o treinamento dos atributos dessa etapa e das etapas anteriores. Para a situação em que há um limite de atributos a serem treinados, simplesmente treina-se os atributos das etapas 1 até o limite.

4.3 TREINAMENTO DOS DEGUSTADORES

De forma análoga os procedimentos da etapa (iii) para a avaliação do impacto de se treinar um determinado degustador para todos os atributos sobre o desempenho do painel. Esse treinamento ocorreu em 11 etapas, valor igual a quantidade de degustadores. Os resultados são apresentados na Tabela 2.3, que segue o mesmo formato da Tabela 2.2.

Pode-se perceber, novamente, que o valor de Alfa aumenta a cada etapa, com o valor final sendo igual a 1. A ordem dos degustadores a ser treinada é: 11, 5, 8, 4, 3, 9, 7, 10, 1, 2 e 6. Vale ressaltar que essa ordem de degustadores é similar, mas não igual, à ordem decrescente de degustadores por seu índice Alfa de Ledaphin (conforme apresentado na Tabela 2.1), o que justifica o uso dessa ferramenta.

Os resultados da Tabela 2.3 podem ser utilizados pelo gestor do painel de forma semelhante aos resultados da Tabela 2.2, para a definição dos degustadores prioritários com vistas a atingir-se um Alfa-meta ou uma limitação de degustadores.

4.4 TREINAMENTO COMBINADO

Nessa etapa propõe-se o treinamento de apenas um atributo de um degustador. Esse treinamento apoia-se em 110 etapas,

geradas pela multiplicação do número de degustadores (11) pela quantidade de atributos (10). A partir desse treinamento, gera-se uma tabela semelhante às Tabelas 2.2 e 2.3, com cada linha representando uma combinação entre degustador e atributo. A Tabela 2.4 ilustra os resultados gerados com a utilização de *it* igual a 1; a tabela em seu tamanho original é suprimida por limitações de espaço.

Tabela 2.2: Treinamento dos atributos

Atributo	Etapa 1	Etapa 2	Etapa 3	Etapa 4	Etapa 5	Etapa 6	Etapa 7	Etapa 8	Etapa 9	Etapa 10
Fragrância/aroma	0,610	0,667	0,713	0,748	0,784	0,820	0,862	0,906	-	-
Uniformidade	0,608	0,666	0,715	0,751	0,788	-	-	-	-	-
Ausência de defeitos	0,626	0,689	-	-	-	-	-	-	-	-
Doçura	0,645	-	-	-	-	-	-	-	-	-
Sabor	0,614	0,672	0,720	0,755	-	-	-	-	-	-
Acidez	0,616	0,674	0,722	-	-	-	-	-	-	-
Corpo	0,609	0,666	0,712	0,747	0,782	0,820	0,861	0,904	0,951	-
Finalização	0,611	0,669	0,716	0,751	0,787	0,823	0,865	-	-	-
Equilíbrio	0,607	0,663	0,710	0,745	0,780	0,816	0,857	0,902	0,948	1,000
Final	0,610	0,668	0,716	0,752	0,788	0,826	-	-	-	-
Máximo	0,645	0,689	0,722	0,755	0,788	0,826	0,865	0,906	0,951	1,000

Tabela 2.3: Treinamento dos degustadores

Degustador	Etapa 1	Etapa 2	Etapa 3	Etapa 4	Etapa 5	Etapa 6	Etapa 7	Etapa 8	Etapa 9	Etapa 10	Etapa 11
1	0,615	0,685	0,737	0,783	0,826	0,865	0,902	0,935	0,966	-	-
2	0,613	0,681	0,731	0,779	0,823	0,861	0,899	0,932	0,962	0,984	-
3	0,637	0,703	0,754	0,799	0,841	-	-	-	-	-	-
4	0,636	0,704	0,756	0,802	-	-	-	-	-	-	-
5	0,640	0,710	-	-	-	-	-	-	-	-	-
6	0,614	0,681	0,730	0,776	0,821	0,861	0,898	0,931	0,962	0,983	1,000
7	0,624	0,692	0,744	0,790	0,836	0,874	0,913	-	-	-	-
8	0,639	0,708	0,758	-	-	-	-	-	-	-	-
9	0,636	0,702	0,752	0,798	0,841	0,880	-	-	-	-	-
10	0,625	0,691	0,742	0,787	0,832	0,871	0,909	0,944	-	-	-
11	0,659	-	-	-	-	-	-	-	-	-	-
Máximo	0,659	0,710	0,758	0,802	0,841	0,880	0,913	0,944	0,966	0,984	1,000

Tabela 2.4: Treinamento combinado - parcial

Degustador	Atributo	Etapa 1	Etapa 2	Etapa 3	...	Etapa 74	...	Etapa 108	Etapa 109	Etapa 110
1	Fragrância/aroma	0,592	0,606	0,615	...	0,000	...	-	-	-
2	Fragrância/aroma	0,589	0,603	0,612	...	0,942	...	-	-	-
3	Fragrância/aroma	0,592	0,606	0,615	...	0,000	...	-	-	-
4	Fragrância/aroma	0,593	0,607	0,616	...	0,000	...	-	-	-
5	Fragrância/aroma	0,599	0,613	0,622	...	0,000	...	-	-	-
6	Fragrância/aroma	0,589	0,603	0,612	...	0,941	...	0,993	-	-
...
6	Finalização	0,590	0,604	0,613	...	0,942	...	0,991	0,995	1,000
7	Finalização	0,590	0,604	0,614	...	0,943	...	-	-	-
...
11	Final	0,596	0,609	0,618	...	0,000	...	-	-	-
	Máximo	0,599	0,613	0,622	0,636	0,943	0,993	0,995	1,000	1,000

Assim como nos casos anteriores, o valor de Alfa aumenta a cada etapa e o valor após a última etapa é igual a 1. A tabela 2.4 completa pode ser utilizada pelo gestor do painel para alcançar os mesmos objetivos já explicados nas etapas anteriores.

4.5 COMPARAÇÃO DOS RESULTADOS

Esta etapa visa identificar o tipo de treinamento – atributo de forma isolada, degustador de forma isolada ou combinado – que demanda um menor número de substituições nas avaliações (representando o efeito de treinamentos). Cada etapa de treinamento por atributo gera a necessidade de 99 substituições (11×9), cada etapa do treinamento do degustador necessita de 90 substituições (10×9) e cada etapa do combinado exige 9 substituições.

No entanto, para uma comparação correta, seria necessário manter o número de substituições para cada método, o que torna a comparação do treinamento por atributo e por degustador inviável. Para solucionar o problema, decidiu-se gerar uma interpolação linear entre os valores de substituição ausentes. Além disso, a necessidade de substituição de cada método foi dividida por 9 (número de produtos comum a todos os métodos para facilitar a exibição). Fazendo essa divisão e a interpolação, gerou-se a Figura 2.1, que compara os três métodos utilizando um valor de it igual a 1.

A Figura 2.1 permite concluir que treinamentos de forma combinada (ou seja,

focando em atributos específicos de determinados degustadores) conduz a valores superiores de Alfa para quase todo o espaço de substituições. Além disso, percebe-se que o valor máximo do Alfa (quando todos os degustadores e atributos foram treinados) é igual a 1. Geram-se então as Figuras 2.2, 2.3 e 2.4, que apresentam os resultados para valores de índice de aprendizado iguais a 0,75, 0,5 e 0,25, respectivamente.

Figura 2.1: Comparação dos métodos com $it = 1$

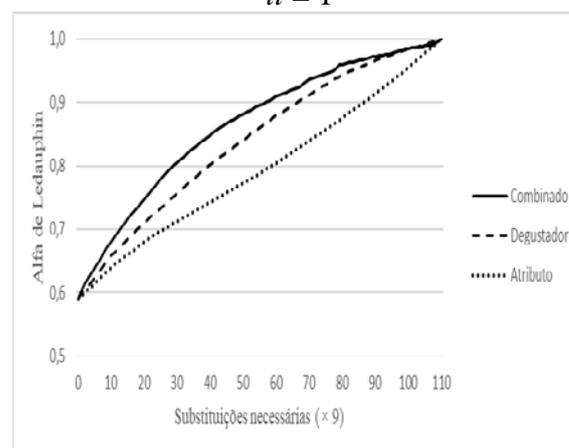


Figura 2.2: Comparação dos métodos com $it = 0,75$

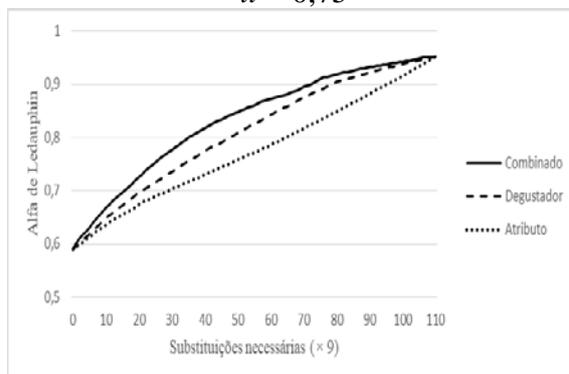
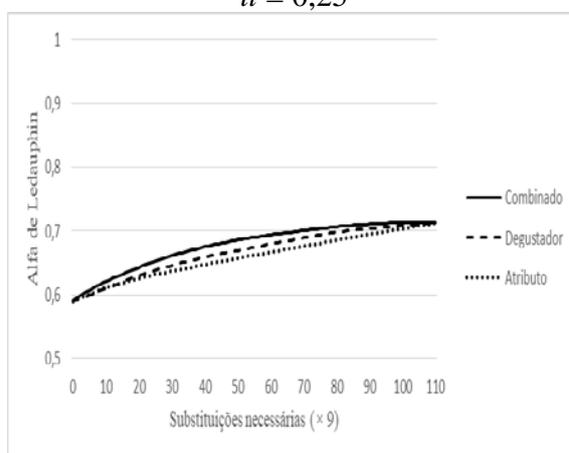


Figura 2.3: Comparação dos índices com $it = 0,25$



A primeira conclusão obtida das Figuras 2.1, 2.2, 2.3 e 2.4 é que o método combinado apresenta melhor desempenho para todos os valores de it testados. No entanto, a diferença entre o desempenho do método combinado e dos demais métodos reduz-se com a redução dos valores de it . Para visualizar esses resultados de forma numérica, montou-se a Tabela 2.5, que mostra, para cada método de treinamento, o Alfa Médio e o Alfa Máximo.

Tabela 2.5: Índices Alfa

Método	$it = 1$	$it = 0,75$	$it = 0,5$	$it = 0,25$	Média
Combinado	0,868	0,836	0,765	0,679	0,787
Degustador	0,843	0,813	0,746	0,669	0,768
Atributos	0,796	0,777	0,726	0,660	0,740
Alfa Máximo	1	0,953	0,841	0,713	

A partir dos resultados da Tabela 2.5, pode-se concluir que o ordenamento dos métodos em função de seu desempenho mantém-se o mesmo independentemente do valor de it , com o treinamento combinado apresentando o melhor resultado, seguido do treinamento por degustador. Também se pode notar a redução do Alfa Máximo com a redução de it , o que é justificado pelo fato de que menores valores de it fazem com que o treinamento tenha menor efeito, e as notas dos degustadores mantenham-se mais próximas às originais.

Outra forma de se comparar o desempenho de cada método consiste em avaliar a quantidade de substituições necessárias. Para isso, parte-se do Alfa mínimo (0,590) e vai-se até o Alfa máximo, em incrementos de 0,05, encontrando para cada valor de Alfa a quantidade de substituições que cada método necessita até atingir tal valor de Alfa (por exemplo, para um valor de it igual a 1, o método combinado necessita de 8 substituições para atingir um Alfa de 0,66, enquanto o método por degustador exige 11 e o método por atributos exige 15). Fazendo a média desses valores para cada método e valor de it gera-se a Tabela 2.6.

Tabela 2.6: Média de substituições necessárias

Método	$it = 1$	$it = 0,75$	$it = 0,5$	$it = 0,25$	Média
Combinado	37,46	37,01	35,61	32,78	35,72
Degustador	44,35	44,31	44,16	43,00	43,95
Atributos	57,02	55,44	53,14	50,70	54,07

Os resultados da Tabela 2.6 demonstram que o treinamento combinado exige, em média, 18,7% menos treinamento do que o treinamento por degustador e 33,9% a menos que o treinamento por atributos, enquanto que o treinamento por degustador necessita de 18,7% menos treinamento que o método por atributos (as diferenças entre as médias foram testadas por uma ANOVA e apresentam o valor de p menor do que 5%). Além disso, percebe-se uma redução das substituições necessárias

com valores menores de *it*; isso é explicado pela convergência mais rápida dos Alfas para valores menores de *it*, conforme pode ser visualizado nas Figuras 2.1 a 2.4.

5 CONCLUSÕES

Três métodos para se treinar degustadores de um painel sensorial foram aplicados, avaliando o ganho de desempenho do painel e a quantidade de treinamento necessária. Todos os três métodos, conforme esperado, melhoraram o desempenho do painel a cada iteração. No entanto, perceberam-se diferenças na quantidade de treinamento demandado por cada método para obter os mesmos ganhos.

O método proposto de treinamento combinado entre degustadores e atributos apresentou os melhores resultados para todos os valores de índice de aprendizado testados, gerando um ganho de desempenho com uma necessidade menor de treinamento. Depois desse método, o método tradicional de se treinar alguns degustadores para todos os atributos mostrou-se melhor do que a opção de treinar apenas alguns atributos para todos os degustadores.

Como o treinamento de degustadores está associado a custos financeiros e de tempo, reduzir o treinamento é uma das preocupações de gestores de painéis sensoriais. Por isso, a principal contribuição deste trabalho foi apresentar um método que permite aumentar o desempenho de um painel às custas de menor volume de treinamento.

Diversas propostas de estudos futuros são delineadas. O efeito dos treinamentos, neste estudo, foi simulado através da substituição das notas a serem treinadas pela matriz ponderada **Z** e por um índice de treinamento; sugere-se, para futuros trabalhos, utilizar as mesmas técnicas aqui apresentadas com base no efetivo treinamento dos degustadores

(coletando as notas antes e depois do mesmo). Outro desdobramento consiste em aplicar e comparar as três técnicas de treinamento avaliadas (treinamento de atributos, treinamento de degustadores e treinamento combinado) para outros painéis sensoriais, tanto de café quanto de outros gêneros alimentícios. Dessa forma, pode-se dar mais robustez aos resultados deste estudo.

AGRADECIMENTOS

Os autores agradecem a Universidade Federal do Rio Grande do Sul, Departamento de Engenharia de Produção e Transportes, pelo suporte ao desenvolvimento deste estudo, bem como ao Instituto Federal do Espírito Santo, campus Venda Nova do Imigrante pela realização dos estudos sensoriais e pelo suporte via edital PRPPG No. 10/2019 – Programa de Pesquisador Produtividade e ao CNPq (469058/2014-5), pelo financiamento desta pesquisa.

REFERÊNCIAS

- ANZANELLO, M.; FOGLIATTO, F.; ROSSINI, K. Data mining-based method for identifying discriminant attributes in sensory profiling. **Food Quality and Preference**, v. 22, 2011.
- BANFIELD, C; HARRIES, J. A technique for comparing judges' performance in sensory tests. **International Journal of Food Science & Technology**, v. 10, 1975.
- BI, J. Agreement and reliability assessments for performance of sensor descriptive panel. **Journal of Sensory Studies**, v. 18, 2007.
- BORGOGNONE, M.; BUSSI, J.; HOUGH, G. Principal component analysis in sensory

- analysis: covariance or correlation matrix? **Food Quality and Preference**, v. 12, 2001.
- BRO, R. PARAFAC. Tutorial and applications. **Chemometrics and Intelligent Laboratory Systems**, v.38, 1997.
- COCCHI, M.; BRO, R.; DURANTE, C.; MANZINI, D.; MARCHETTI, A.; SACCANI, F.; SIGHINOLFI, S.; ULRICI, A. Analysis of sensory data of Aceto Balsamico Tradizionale di Modena (ABTM) of different ageing by application of PARAFAC models. **Food Quality and Preference**, v. 17, 2006.
- CRONBACH, L.; Coefficient alpha and the internal structure of tests. **Psychometrika**, v. 16, 1951.
- DIJKSTERHUIS, G. Assessing panel consonance. **Food Quality and Preference**, v. 6, 1995.
- LATREILLE, J.; MAUGER, E.; AMBROISINE, L.; TENENHAUS, M.; VINCENT, M.; NAVARRO, S.; GUINOT, C. Measurement of the reliability of sensory panel performances. **Food Quality and Preference**, v. 17, 2006.
- LEDAUPHIN, S.; HANAFI, M.; QANNARI, E. Assessment of the agreement among the subjects in fixed vocabulary profiling. **Food Quality and Preference**, v.17, 2006.
- MCEWAN, J.; HUNTER, E.; VAN GEMERT, L.; LEA, P. Proficiency testing for sensory profile panels: measuring panel performance. **Food Quality and Preference**, v. 13, 2002.
- PINTO, F.; FOGLIATTO, F.; QANNARI, E. A method for panelists' consistency assessment in sensory evaluations based on the Cronbach's alpha coefficient. **Food Quality and Preference**, v. 32, 2014.
- ROSSI, F. Assessing sensory panelist performance using repeatability and reproducibility measures. **Food Quality and Preference**, v. 12, 2001.
- SCA, 2018. **Specialty Coffee Association Protocols**. 2018. Disponível em: <<https://sca.coffee/research/protocols-best-practices>>. Acesso em 4 set. 2018.
- SINESIO, F.; RISVIK, E.; RODBOTTEN, M. Evaluation of panelist performance in descriptive profiling of rancid sausages: a multivariate study. **Journal of Sensory Studies**, v. 5, 1990.
- SIVERTSEN, H.; RISVIK, E. A Study of Sample and Assessor variation: A multivariate study of wine profiles. **Journal of Sensory Studies**, v. 9, 1994.
- TALSMA, P. Assessing sensory panel performance using generalizability theory. **Food Quality and Preference**, v. 47, 2016.
- VIGNEAU, E; QANNARI, E. Clustering of variables around Latent Components. **Communications in Statistics- Simulation and Computation**, v. 32, 2003.
- WILDERJANS, T.; CARIOU, V. CLV3W: A clustering around latent variables approach to detect panel disagreement in three-way conventional sensory profiling data. **Food Quality and Preference**, v. 47, 2016.