

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MILENA RIBEIRO SILVA

**UTILIZAÇÃO DE IMAGENS DIGITAIS E CALIBRAÇÃO MULTIVARIADA PARA
A DETERMINAÇÃO DA PORCENTAGEM DE CACAU EM CHOCOLATES**

CAMPO MOURÃO

2023

MILENA RIBEIRO SILVA

**UTILIZAÇÃO DE IMAGENS DIGITAIS E CALIBRAÇÃO MULTIVARIADA PARA
A DETERMINAÇÃO DA PORCENTAGEM DE CACAU EM CHOCOLATES**

**Use of digital images and multivariate calibration to determine the percentage of cocoa
in chocolates**

Trabalho de Conclusão de Curso apresentado ao curso superior de Licenciatura em Química, da Universidade Tecnológica Federal do Paraná – UTFPR, Campus Campo Mourão, como requisito parcial para a obtenção do título de Licenciada em Química.

Orientadora: Prof. Dra. Patrícia Valderrama
Coorientadora: Prof. Dra. Vanessa Jorge dos Santos

CAMPO MOURÃO

2023



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

MILENA RIBEIRO SILVA

**UTILIZAÇÃO DE IMAGENS DIGITAIS E CALIBRAÇÃO MULTIVARIADA PARA
A DETERMINAÇÃO DA PORCENTAGEM DE CACAU EM CHOCOLATES**

Trabalho de Conclusão de Curso apresentado ao curso superior de Licenciatura em Química, da Universidade Tecnológica Federal do Paraná – UTFPR, Campus Campo Mourão, como requisito parcial para a obtenção do título de Licenciada em Química.

Data de aprovação: 29 de novembro de 2023

Patrícia Valderrama
Doutorado
Universidade Tecnológica Federal do Paraná

Vanessa Jorge dos Santos
Doutorado
Universidade Tecnológica Federal do Paraná

Gideã Taques Tractz
Doutorado
Universidade Tecnológica Federal do Paraná

Marcos Antônio Piza
Doutorado
Universidade Tecnológica Federal do Paraná

CAMPO MOURÃO

2023

Dedico este trabalho à minha família pela compreensão de
minha ausência e ao meu noivo pelo apoio incondicional.

AGRADECIMENTOS

Gostaria de expressar meus sinceros agradecimentos a todos que desempenharam um papel crucial na conclusão deste trabalho. Começo meus agradecimentos com uma profunda gratidão a Deus, que esteve ao meu lado ao longo desses difíceis e desafiadores anos de faculdade, me concedendo inteligência e sabedoria para concluir meus estudos.

Minha família desempenhou um papel central na minha jornada acadêmica, e não posso deixar de expressar minha gratidão a eles. Primeiramente, à minha mãe, Ana Maria e meu pai Vivaldo, que dedicaram todos os recursos possíveis, mesmo quando parecia impossível, para que eu pudesse entrar e concluir a faculdade. Eles não apenas me forneceram apoio financeiro, mas também estiveram presentes nos momentos em que eu acreditava que não conseguiria seguir adiante. A constante presença deles, apesar da distância, foi um alicerce inabalável.

Minha irmã, Viviane, desempenhou o papel de uma segunda mãe ao me receber em sua casa em Campo Mourão quando iniciei a faculdade. Ela me ensinou responsabilidade e independência, além de oferecer um refúgio nos momentos de desespero que todo calouro enfrenta. Graças a ela, cresci e me tornei a mulher independente que sou hoje.

Por último, mas não menos importante, quero agradecer profundamente ao meu noivo. Desde o segundo ano da faculdade até o momento presente, ele tem sido meu apoio constante. Ele me ajudou imensamente ao longo dos anos da faculdade. Seu apoio incondicional, sua capacidade de me confortar nos momentos difíceis foram essenciais.

Minhas orientadoras, Patrícia Valderrama, que me deu todo o suporte que precisei e que é um grande exemplo de professora e pesquisadora para mim, uma das melhores que já conheci, e a Vanessa Jorge dos Santos, que foi uma excelente professora para mim, uma ótima amiga e uma orientadora incrível, que me ajudou em tudo que precisei e me deu total suporte. Suas orientações e apoio foram essenciais para minha formação.

Agradeço aos amigos da faculdade que estiveram comigo do início ao fim desta jornada. André, Roberto, Ana Cristina, Rebeca e Renan, vocês compartilharam todas as dificuldades e nunca me deixaram desistir. Levarei essa amizade para a vida toda.

A todos os professores que contribuíram para meu crescimento acadêmico de alguma forma, meu profundo agradecimento.

Gostaria expressar minha profunda gratidão à UTFPR por todo o apoio recebido, assim como ao Departamento de Química (DAQUI). Obrigado a todos por fazerem parte desta jornada, e por tornarem este momento possível!

*“Até cortar os próprios defeitos pode ser perigoso.
Nunca se sabe qual é o defeito que sustenta nosso
edifício inteiro.” – Clarisse Lispector*

RESUMO

O chocolate é um dos alimentos mais consumidos do mundo, e sua produção é diversificada, apresentando produtos com diferentes teores de cacau. Por isso, este estudo teve por objetivo desenvolver e validar um modelo de calibração para a determinação da porcentagem de cacau em chocolates, integrando análise de imagem digital e regressão por mínimos quadrados parciais (PLS), como ferramenta quimiométrica. Um total de 1701 imagens foram adquiridas, empregando três diferentes *smartphones* em um ambiente com iluminação controlada. Foram consideradas as variedades de chocolate ao leite, meio amargo e amargo, de diferentes marcas e lotes. Como método de referência, a porcentagem de cacau foi obtida a partir das informações das embalagens. Os histogramas R,G,B foram separados em conjunto de calibração e validação externa. A escolha do número adequado de variáveis latentes foi feita por validação cruzada em blocos contínuos de 10 amostras. O modelo foi construído com 10 variáveis latentes, com os dados centrados na média. Os parâmetros de mérito como exatidão, ajuste, linearidade, inverso da sensibilidade analítica, limites de detecção e quantificação, e a relação de desempenho do desvio mostraram que o modelo está adequado para determinar a porcentagem de cacau em amostras de chocolates. A consideração dos coeficientes de regressão permitiu identificar a contribuição de cada canal no modelo PLS para a predição da porcentagem de cacau. O método proposto destaca-se pela sua eficiência, praticidade e potencial de aplicação na indústria alimentícia. Sua abordagem, aliada aos resultados de validação, sugere uma contribuição significativa para a análise de chocolates. Essa pesquisa não apenas apresenta uma solução para a determinação da porcentagem de cacau em chocolates, mas também abre caminho para futuras investigações e aplicações no campo da análise de alimentos por meio de técnicas avançadas de imagem e modelagem multivariada.

Palavras-chave: PLS; quimiometria; *smartphone*; teor de cacau.

ABSTRACT

Chocolate is one of the most consumed foods in the world, and its production is diverse, presenting products with different cocoa contents. Therefore, this study aimed to develop and validate a calibration model for determining the percentage of cocoa in chocolates by integrating a digital image and a chemometric tool of partial least squares (PLS) regression. A total of 1701 images were acquired, using three different smartphones in a controlled lighting environment. The varieties of milk, semi-sweet and bitter chocolate, from different brands and batches, were considered. As a reference method, the cocoa percentage was obtained from the information on the packaging. The R,G,B histograms were separated into a calibration and external validation set. The choice of the appropriate number of latent variables was made by cross-validation in continuous blocks of 10 samples. The model was built with 10 latent variables, with mean center preprocess. The parameters of merit such as accuracy, adjust, linearity, the inverse of analytical sensitivity, limits of detection and quantification, and the relation prediction deviation showed that the model is suitable for determining the percentage of cocoa in chocolate samples. The regression coefficients allowed to identify the contribution of each channel in the PLS model to predict the percentage of cocoa. The proposed method stands out for its efficiency, practicality, and potential for application in the food industry. Their approach, combined with validation results, suggests a significant contribution to the analysis of chocolates. This research not only presents a solution for determining the percentage of cocoa in chocolates but also paves the way for future investigations and applications in the field of food analysis through advanced imaging and multivariate modeling techniques.

Keywords: PLS; chemometrics; smartphone; cocoa content.

LISTA DE ILUSTRAÇÕES

Figura 1 - Decomposição da matriz X e o vetor y em <i>scores</i> e <i>loadings</i>	19
Quadro 1 - Equações para calcular os parâmetros de mérito	21
Figura 2 - Leitura da embalagem e rótulo das amostras	22
Figura 3 – Caixa para aquisição das imagens digitais	23
Figura 4 - Canais R,G,B das amostras de chocolate	25
Figura 5 - Número de variáveis latentes	26
Figura 6 - Ajuste do modelo multivariado para determinação da porcentagem de cacau das amostras de calibração	28
Figura 7 - Resíduos de Student para o modelo multivariado na determinação da porcentagem de cacau.....	29
Figura 8 - Coeficiente de regressão obtidos para os modelos PLS	30

LISTA DE TABELAS

Tabela 1 - Diferenças entre os tipos de chocolates mais comercializados	15
Tabela 2 - Parâmetros de mérito do modelo PLS	27

LISTA DE ABREVIATURAS E SIGLAS

ANVISA	Agência Nacional De Vigilância Sanitária
ASTM	Annual Book Of Standards
LD	Limite De Detecção
LQ	Limite De Quantificação
MLR	Regressão Linear Multipla
PCR	Regressão Por Componentes Principais
PLS	Mínimos Quadrados Parciais
RDC	Resolução Da Diretoria Colegiada
RGB	Vermelho (Red), Verde (Green) E Azul (Blue)
RMSEC	Raiz Quadrada Do Erro Médio Quadrático De Calibração
RMSECV	Raiz Quadrada Do Erro Médio Quadrático De Validacao Cruzada
RMSEP	Raiz Quadrada Do Erro Médio Quadrático De Previsão
RPD	Relação De Desempenho Do Desvio
VLs	Variáveis Latentes

SUMÁRIO

1	INTRODUÇÃO	12
2	OBJETIVOS	14
2.1	Objetivos específicos	14
3	FUNDAMENTAÇÃO TEÓRICA	15
3.1	Chocolate	15
3.2	Imagens digitais	16
3.3	Quimiometria	17
3.3.1	Calibração multivariada	17
3.3.1.1	<u>Regressão por mínimos quadrados parciais (PLS)</u>	18
3.4	Validação	19
4	METODOLOGIA	22
4.1	Amostragem	22
4.2	Verificação da porcentagem de cacau	22
4.3	Aquisição das imagens digitais	23
4.4	Calibração multivariada	23
4.5	Validação	24
5	RESULTADOS E DISCUSSÃO	25
6	CONCLUSÃO	31
	REFERÊNCIAS	32

1 INTRODUÇÃO

A indústria de chocolate é um setor que apresenta grande relevância econômica em todo o mundo. De acordo com Brasil (2023), a indústria apresentou um crescimento de 11,43% na produção no primeiro semestre de 2022, em comparação com o mesmo período do ano anterior, foram produzidas 370 mil toneladas em contraste com as 332 mil toneladas de 2021.

O setor de chocolate possui grande relevância econômica globalmente, e o Brasil se destaca como um dos principais produtores e processadores de cacau (Brasil, 2023). A Alemanha é a principal produtora de chocolate no mundo. De acordo com Leite (2018), o Brasil se destaca como um dos poucos países que concentram a produção de cacau e o processamento de chocolates em seu território, permitindo que se posicione como um dos principais produtores de chocolate globalmente.

Dentre os variados tipos de chocolates, o chocolate ao leite é o mais consumido no Brasil (Brasil, 2022). Além disso, o Brasil contém inúmeras indústrias neste ramo, sendo a Nestlé® a maior marca vendida, representando cerca de 34% do mercado brasileiro (Mercado do Cacau, 2018).

O chocolate é produzido pela mistura de diversos ingredientes como derivados de cacau, massa de cacau, cacau em pó e ou manteiga de cacau, dentre outros. Em virtude da grande importância do cacau na formulação do chocolate este pode ser classificado em ao leite, meio amargo e amargo, diferenciando-se com base na porcentagem de cacau presente. Dentre esses tipos o chocolate ao leite é o mais consumido mundialmente (Brasil, 2022).

A porcentagem de cacau desempenha um papel crucial na definição das características sensoriais do chocolate, influenciando diretamente seu sabor, aroma, textura e intensidade. Portanto, a avaliação precisa do teor de cacau no chocolate é fundamental para assegurar a qualidade e a uniformidade na produção, prevenindo possíveis práticas fraudulentas. Nas indústrias, a vigilância rigorosa do controle de qualidade é imperativa devido à relevância do aspecto visual e do sabor para a comercialização do produto.

Nas indústrias, a determinação desse parâmetro envolve a utilização de equipamentos de alto custo, como espectrofotômetros e cromatógrafos. Esses dispositivos, embora eficazes, têm a desvantagem de consumir quantidades significativas de solventes tóxicos e demandar um tempo considerável para a conclusão das análises (Botelho *et al.*, 2014).

Nesse contexto, o uso de imagens digitais, por meio de *smartphones*, oferece uma alternativa de baixo custo, ambientalmente correta e de análise instantânea. No entanto, estes não fornecem informações claras, necessitando do auxílio de ferramentas quimiométricas para

interpretar os resultados. Um dos métodos quimiométricos mais utilizados com imagens digitais é a calibração multivariada. Inúmeros estudos podem ser encontrados envolvendo a aplicação de imagens digitais com quimiometria, como Vittorazzi *et al.* (2020) em classificação de cédulas brasileiras (R\$) e como Botelho *et al.* (2014), Fulgêncio *et al.* (2020), Botelho *et al.* (2017), Curbani *et al.* (2020) e Assad *et al.* (2002) em amostras de alimentos.

Nesse cenário, a necessidade de métodos mais eficientes e acessíveis para a determinação da porcentagem de cacau é evidente. A aplicação de imagens digitais, aliada à ferramenta da calibração multivariada, oferece uma resposta promissora a esse desafio. Através do uso de *smartphones*, essa abordagem proporciona uma análise instantânea, econômica e ambientalmente correta, tornando-se uma alternativa viável e eficaz para as indústrias de chocolates.

2 OBJETIVOS

Construir e validar um modelo de calibração multivariada para a determinação da porcentagem de cacau em chocolates usando imagens digitais.

2.1 Objetivos Específicos

- Realizar uma amostragem representativa de chocolates comerciais, incluindo chocolates ao leite (com teor de cacau variando de 20% a 39%), meio amargo (com teor de cacau variando de 40% a 55%) e amargo (com teor de cacau variando de 56% a 100%);
- Observar o rótulo de cada amostra, focando na porcentagem de cacau presente;
- Utilizar três *smartphones* de marcas diferentes para adquirir as imagens digitais dos chocolates;
- Carregar e cortar as imagens utilizando o *software* Matlab;
- Construir um modelo de calibração multivariada através da regressão por mínimos quadrados parciais (PLS) para a porcentagem de cacau;
- Validar o modelo citado anteriormente.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Chocolate

O chocolate, um produto derivado do cacau, é obtido principalmente a partir da utilização de massa de cacau ou pó de cacau, manteiga de cacau e uma combinação de outros ingredientes, como leite, açúcar, emulsificantes e aromatizantes (Chevalley, 1994), contendo, no mínimo, 25 % (g/100 g) de sólidos totais de cacau.

O Brasil é o quinto maior produtor e consumidor de chocolate do mundo. O doce brasileiro fica atrás apenas dos Estados Unidos, Alemanha, Rússia e Reino Unido (Brasil, 2022).

De acordo com os dados da Associação Brasileira de Indústrias de Chocolates, Amendoins e Balas (Brasil, 2023), o setor de chocolates fechou 2021 com um volume de produção de 693 mil toneladas, crescimento de 35,9% em relação ao mesmo período de 2020. Seu faturamento no Brasil, em 2021, foi de R\$ 13 bilhões.

A receita do chocolate pode variar bastante de país para país, em função das diferentes preferências e regulamentações que existem. Em termos legais, são estabelecidos limites para as quantidades de cacau, sólidos do leite e tipos de gorduras vegetais que podem ser adicionados ao produto. Essas gorduras incluem a manteiga de cacau, a gordura do leite e as gorduras vegetais em si (Martin, 1994).

Atualmente o Brasil tem uma grande variedade de chocolates, no qual os mais consumidos são: chocolate ao leite, meio amargo, amargo e branco. Na Tabela 1 é descrito os ingredientes e teor de cacau referente a cada tipo de chocolate.

Tabela 1- Diferenças entre os tipos de chocolates mais comercializados

Tipo de chocolate	Ingredientes utilizados	Teor de massa de cacau (%)
Ao leite	Açúcar, massa e manteiga de cacau, leite em pó e condensado	20-39
Amargo	Massa e manteiga de cacau, pouco açúcar	56-85
Meio Amargo	Massa e manteiga de cacau, pouco açúcar	40-55

Fonte: (Souza, 2010, p. 26)

O cacau é um ingrediente fundamental na produção de chocolate e sua porcentagem influencia diretamente nas características e qualidade do produto final. Existem diversos tipos de chocolates disponíveis no mercado, com uma ampla variedade de percentuais de cacau. Essa variação de teor de cacau resulta em diferenças significativas de sabor, aroma, textura e intensidade do chocolate. A porcentagem mínima de cacau em um produto para ser chamado de chocolate no Brasil é de 25% (Anvisa, 2005).

No Brasil e em alguns outros países, os chocolates com menor teor de cacau, como os chocolates ao leite, são mais populares e amplamente consumidos. Esses chocolates têm um perfil de sabor mais suave e doce, porém as preferências em relação ao teor de cacau variam de acordo com o país e a cultura. No Canadá e em outros países europeus, chocolates com teores mais elevados de cacau são mais apreciados e consumidos (Mordor Intelligence, 2023).

Essa diversidade de preferências reflete-se não apenas nos hábitos alimentares, mas também nas técnicas de análise utilizadas para diferenciar as porcentagens de cacau nos chocolates, e com a evolução tecnológica que se teve com passar dos anos, as análises tradicionais poderiam ser substituídas por processamento de imagens digitais de alta precisão, que estão sendo cada vez mais utilizadas em diversos campos científicos e técnicos.

3.2 Imagens digitais

Os avanços na microeletrônica na última década permitiram que telefones celulares estivessem disponíveis imediatamente, com especificações cada vez melhores e a um custo reduzido (GRUDPAN *et al.*, 2015). Isso resultou na ampla utilização de imagens digitais obtidas por esses dispositivos em diversos campos científicos e técnicos (Lopez-Molinero; Jimenez-Lamana, 2016).

Botelho, Dantas e Sena (2014) afirmaram que a objetificação da cor para fins de informação química tornou-se mais fácil devido ao aumento da disponibilidade de dispositivos de imagem portáteis e de baixo custo, tais como *scanners*, câmeras digitais, *webcams*, *tablets* e telefones celulares, e isso contribuiu para a democratização da ciência da medição.

A decomposição de imagens digitais em um sistema de cores é a forma mais comum de extrair informações. O sistema RGB (Red, Green e blue) é um sistema aditivo que utiliza a combinação de vermelho, verde e azul para formar diferentes tonalidades de cores. Cada pixel é formado por uma combinação dessas cores, e a intensidade de cada cor é medida em canais, onde 0 significa ausência total de cor e 255 representa a intensidade máxima. A combinação desses canais cria diferentes cores, que podem ser analisadas por um histograma de frequência.

Esse histograma pode ser usado para desenvolver modelos quimiométricos, onde as variáveis RGB são fundidas com outros parâmetros de cor (Botelho; Assis; Sena, 2014).

Imagens digitais apresentam dados complexos, especialmente em grandes volumes, exigindo ferramentas de análise. A quimiometria é uma abordagem amplamente empregada nesse contexto, integrando a objetificação da cor por meio de imagens digitais com métodos matemáticos e estatísticos.

3.3 Quimiometria

De acordo com Otto (1999), a quimiometria é a aplicação de métodos matemáticos e estatísticos para tratar dados químicos com o objetivo de obter mais informações e melhores resultados analíticos.

A quimiometria desenvolveu novas técnicas para tratar dados com várias aplicações, incluindo otimização de processos, classificação de dados e determinação quantitativa. A quimiometria abrange muitas áreas, e uma delas é a calibração multivariada, que na química analítica é a principal linha de pesquisa (Valderrama, 2005).

3.3.1 Calibração multivariada

A calibração multivariada desempenha um papel crucial na análise química, pois estabelece uma relação matemática entre as medidas instrumentais e os valores da propriedade de interesse. É uma etapa fundamental que proporciona resultados precisos e confiáveis (Valderrama, 2005). Um modelo de calibração é uma função matemática que relaciona duas variáveis: a variável dependente (Y) e a variável independente (X), como demonstrada pela Equação 1.

$$Y = f(X) = Xb \quad (1)$$

De acordo com Santos (2021), a calibração consiste em duas etapas: a modelagem dos dados, que seleciona as amostras de forma representativa para o conjunto de calibração para posteriormente construir um modelo que relaciona matematicamente as duas variáveis (X e Y). A segunda etapa corresponde a validação, na qual as variáveis independentes são combinadas com o coeficiente de regressão para calcular os valores previstos da variável dependente em um novo conjunto de amostras (Valderrama, 2005).

Durante a calibração, um conjunto de dados específico, conhecido como conjunto de calibração, é usado para ajustar os parâmetros do modelo, representados pelos coeficientes de regressão (b), através de cálculos matemáticos utilizando os dados experimentais (Martens; Naes, 1996).

Na calibração multivariada, ocorre a relação entre mais de uma resposta instrumental e a propriedade de interesse. Esses métodos de calibração permitem a análise mesmo quando há interferentes presentes, desde que esses interferentes estejam presentes nas amostras usadas para construir o modelo de calibração. Além disso, a calibração multivariada oferece possibilidades adicionais, como determinações simultâneas e análises mesmo sem resolução (Valderrama, 2005; Martens; Naes, 1996).

De acordo com Valderrama (2005), na área da química analítica, uma variedade de métodos de regressão tem sido aplicada para construir modelos de calibração multivariada. Dentre esses métodos, os mais utilizados são a regressão linear múltipla (MLR), a regressão por componentes principais (PCR) e a regressão por mínimos quadrados parciais (PLS).

O PLS, que é amplamente utilizado na química analítica, destaca-se como o método de regressão mais empregado, pois esse método é especialmente indicado quando X contém variáveis altamente correlacionadas, como no presente trabalho. Outra vantagem desse método é que pode ser usado mesmo quando as amostras contêm interferentes que podem estar presentes nas amostras do conjunto de calibração (Morgano, 2007).

. Soma-se a isto que existem muitos trabalhos que utilizam essa ferramenta quimiométrica para obter dados de imagens digitais como os de Valderrama, (2005); Botelho *et al.* (2014), Botelho *et al.* (2017), Rosa *et al.* (2022), Damasceno *et al.* (2015) e Beltrame *et al.* (2016).

3.3.1.1 Regressão por mínimos quadrados parciais (PLS)

O método de calibração multivariada conhecido como regressão por mínimos quadrados parciais (PLS) estabelece uma relação entre a matriz X e a matriz Y , utilizando suas decomposições simultâneas. Esse método gera um produto de duas matrizes: *scores* (T e U) e *loadings* (P , T e QT). Permite a comparação de duas variáveis, a independentes (X) e dependentes (Y), como pode ser observado nas Equações 2 e 3 e na Figura 1.

$$X = TP^T + E_X = \sum t_{AP} t_{A}^T + E_X \quad (2)$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{E}_Y = \sum \mathbf{u}_A \mathbf{q}_A^T + \mathbf{E}_Y \quad (3)$$

Onde, as matrizes \mathbf{X} e \mathbf{Y} representam as respostas instrumentais e a propriedade de interesse, respectivamente. Além disso, existem as matrizes \mathbf{E}_X e \mathbf{E}_Y , que correspondem à parte não modelada, ou seja, aos resíduos (Santos, 2021).

Figura 1. Decomposição da matriz \mathbf{X} e o vetor \mathbf{y} em scores e loadings

$$\begin{array}{c} \begin{array}{ccc} & p & \\ \begin{array}{|c|} \hline \mathbf{X} \\ \hline \end{array} & = & \begin{array}{|c|} \hline \mathbf{T} \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{P}^T \\ \hline \end{array} + \begin{array}{|c|} \hline \mathbf{E} \\ \hline \end{array} \\ n & & n & d & p & n \end{array}$$

$$\begin{array}{c} \begin{array}{ccc} & m & \\ \begin{array}{|c|} \hline \mathbf{y} \\ \hline \end{array} & = & \begin{array}{|c|} \hline \mathbf{U} \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{Q}^T \\ \hline \end{array} + \begin{array}{|c|} \hline \mathbf{F} \\ \hline \end{array} \\ n & & n & d & m & n \end{array}$$

Fonte: (Otto, 1999).

Porém, o PLS estabelece relações complexas entre matrizes de dados, e a validação se torna crucial para avaliar a capacidade do modelo em lidar com novos dados. Ao comparar as previsões do modelo com valores conhecidos, obtidos através de metodologias oficiais para a propriedade de interesse, a validação garante a generalização e a confiabilidade do modelo.

3.4 Validação

Na etapa de validação, são comparados os valores previstos pelo modelo com aqueles conhecidos e obtidos por meio de uma metodologia oficial de análise para a propriedade de interesse. Essa comparação visa avaliar o desempenho do modelo de calibração proposto, garantindo a confiabilidade e precisão das previsões. (Valderrama, 2005). Segundo Botelho *et al.* 2013, a validação envolve a determinação de parâmetros específicos, os quais variam dependendo do método, objetivo e órgão regulatório.

Baqueta *et al.* (2019, 2020) e Santos *et al.* (2020) abordaram que os principais parâmetros avaliados em metodologias envolvendo calibração multivariada são exatidão, limite

de detecção (LD), limite de quantificação (LQ), linearidade, sensibilidade, inverso da sensibilidade analítica e relação de desempenho do desvio (RPD).

A exatidão analisa a concordância entre os valores estimados pelo modelo de calibração multivariada e os valores de referência. A sensibilidade refere-se à capacidade do método de distinguir, com um certo nível de confiança, duas concentrações próximas. O inverso da sensibilidade analítica permite diferenciar a menor concentração que o método analítico pode detectar na faixa de concentração utilizada. LD representa o valor mínimo do analito na amostra que pode ser detectado, porém não necessariamente quantificado, pois detectar o analito significa reconhecer sua presença, enquanto a quantificação requer uma medida precisa e repetitiva, muitas vezes baseada em uma curva de calibração. Já o LQ é o valor mínimo do analito na amostra que pode ser quantificado de forma precisa e reprodutível nas condições estabelecidas.

A avaliação da linearidade é realizada qualitativamente por meio da análise dos resíduos gráficos para observar a aleatoriedade dos dados e também pelo teste estatístico de Jarque-Bera (Jarque; Bera, 1987). O RPD é empregado para estimar a capacidade preditiva do modelo, e conforme apontado por Botelho, Mendes e Sena (2013), os valores que se situam entre 2,4 e 1,5 são considerados satisfatórios.

E a avaliação do ajuste envolve a análise da correlação entre os valores de referência e os valores estimados pelos modelos de calibração. Os cálculos destes parâmetros estão descritos no Quadro 1.

Quadro 1 - Equações para calcular os parâmetros de mérito

Parâmetros de Mérito	Equação
Exatidão	$RMSEP = \sqrt{\frac{\sum_{i=1}^{nv} (y_i - \hat{y}_i)^2}{nv}}$
	$RMSEC = \sqrt{\frac{\sum_{i=1}^{nv} (y_i - \hat{y}_i)^2}{nc}}$
	$RMSECV = \sqrt{\frac{\sum_{i=1}^{nv} (y_i - \hat{y}_i)^2}{ncv}}$
Sensibilidade	$S\hat{E}N = \frac{1}{\ b\ }$
Sensibilidade Analítica	$Sensibilidade\ analítica = \frac{sensibilidade}{\ \delta x\ }$
Sensibilidade Analítica ⁻¹	$Sensibilidade\ Analítica^{-1} = \frac{1}{sensibilidade\ analítica}$
Limite de Detecção	$LD = 3\delta x \ b_k\ = 3\delta x \frac{1}{S\hat{E}N}$
Limite de Quantificação	$LQ = 10\delta x \ b_k\ = 10\delta x \frac{1}{S\hat{E}N}$
RPD	$RPD_{cal} = \frac{DP_{cal}}{RMSECV}$
	$RPD_{val} = \frac{DP_{val}}{RMSEP}$

Raiz quadrada do erro médio quadrático de calibração (RMSEC), previsão (RMSEP), e validação cruzada (RMSECV). nv: conjunto de amostras do conjunto de validação, y_i e \hat{y}_i valores de referência e aos previstos pelo modelo, respectivamente, nc: número de amostras do conjunto de calibração, ncv: número de amostras do conjunto de validação cruzada. Δx : desvio padrão do sinal de referência, b_k : vetor dos coeficientes de regressão do modelo para a espécie k, $S\hat{E}N$: sensibilidade. DP_{cal} e DP_{val} se referem ao desvio padrão dos valores de referência do conjunto de calibração de validação, respectivamente.

Fonte: Adaptado de (Baqueta, 2019)

4 METODOLOGIA

4.1 Amostragem

Para este estudo, foram adquiridas 63 barras de chocolates, representando as variedades ao leite, meio amargo e amargo, de diversas marcas e lotes, comercializados em supermercados da região de Campo Mourão (Brasil), Maringá (Brasil) e de Paris (França). Foram utilizados chocolates ao leite (tradicional), estes sem saborizantes e adição de produtos extras (como amendoim, nozes, entre outros). A seleção foi baseada em critérios como disponibilidade no mercado, representatividade das diferentes categorias de chocolate e diversidade de marcas.

Todas as análises foram feitas em triplicata totalizando 189 amostras no total. Antes da realização das análises, cada amostra de chocolate foi dividida em três partes iguais e devidamente identificada e armazenadas em condições adequadas de temperatura e umidade, seguindo as recomendações do fabricante.

4.2 Verificação da porcentagem de cacau

A obtenção da porcentagem de cacau foi feita através da leitura da embalagem e rótulo das amostras, como pode se observar na Figura 3.

Figura 2 – Leitura da embalagem e rótulo das amostras



Fonte: Autoria própria

4.3 Aquisição de imagens digitais

As imagens digitais das amostras de chocolate foram adquiridas utilizando um conjunto de dispositivos de captura de imagem, incluindo três tipos de *smartphones* com câmera de alta resolução (Samsung Galaxy S23, Xiaomi mi 11 T Pro, e Iphone 13). Foram capturadas três imagens de cada amostra com cada um dos dispositivos, resultando em um total de 567 imagens por *smartphone* e 1701 imagens no conjunto total.

A aquisição das imagens ocorreu em um ambiente de iluminação controlada, dentro de uma caixa preta com apenas uma pequena abertura na parte superior, onde a câmera do *smartphone* foi fixada. A caixa estava equipada com uma fonte de luz constante e uniforme (anel de luz) na parte superior, e uma base completamente branca e plana na parte inferior. Aproximadamente 10 gramas da amostra foram posicionados na superfície da base, a uma distância fixa de 20 cm em relação ao *smartphone*. Como pode-se observar na Figura 3.

Figura 3 – Caixa para aquisição das imagens digitais.



Fonte: Autoria própria

4.4 Calibração Multivariada

Após a aquisição das imagens dos chocolates, cada imagem passou por um processo de pós-processamento e recorte, o que foi realizado com o auxílio do software Matlab R2007B da The MathWorks Inc., localizada em Natick, EUA.

Inicialmente, as 1701 imagens foram carregadas e cortadas de tamanhos aleatórios, uma a uma, utilizando o Matlab. Em seguida, identificou-se a menor imagem cortada e todas as demais foram ajustadas para terem o mesmo tamanho (100x100x3).

Para desenvolver o modelo de calibração multivariada para determinar a porcentagem de cacau nos chocolates, utilizou-se a técnica de regressão PLS (Partial Least Squares), com o pacote computacional PLS Toolbox 5.2 da Eigenvector Research Inc.

Para o desenvolvimento dos modelos PLS, com todos os dados já preparados o histograma de cor de todas as amostras de chocolate formou uma matriz **X**, enquanto o vetor **y** conteve os valores de referência (porcentagem de cacau).

Para o conjunto de dados foram separados $\frac{3}{4}$ das amostras para compor o conjunto de calibração e $\frac{1}{4}$ para os de validação, ambas foram obtidas pelo algoritmo de Kennard-Stone. O propósito desse algoritmo consiste em escolher as amostras com base em suas distâncias; a primeira amostra é escolhida com base na maior distância entre ela e a amostra média, e as amostras subsequentes são selecionadas com base na maior distância em relação à amostra anterior (Kennard; Stone, 1969).

Os dados foram centralizados na média e o número de variáveis latentes (VLs) foi escolhido com base no menor valor encontrado para a raiz do erro quadrático médio da validação cruzada (RMSECV) por validação cruzada de blocos contínuos com 10 amostras (Brereton, 2000).

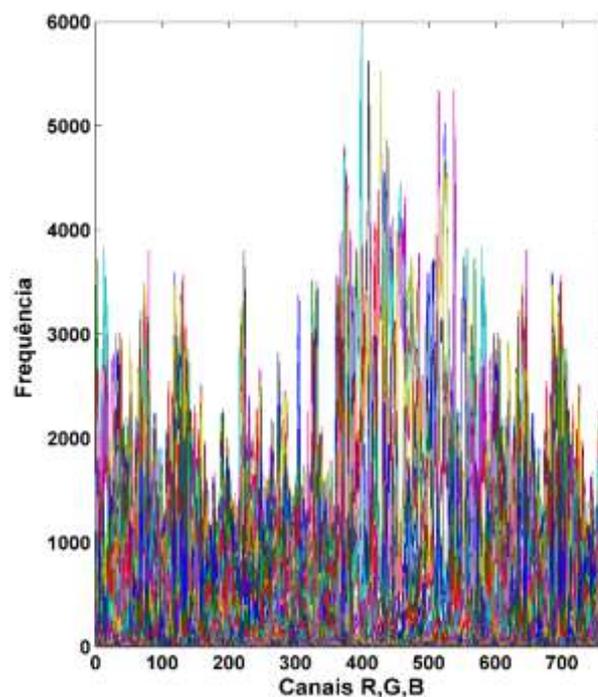
4.5 Validação

A validação dos modelos PLS foram realizadas de acordo com os parâmetros de mérito: exatidão, sensibilidade, inverso da sensibilidade analítica, linearidade, relação de desempenho do desvio, ajuste e limite de detecção e quantificação, como descrito no tópico 2.4.

5 RESULTADOS E DISCUSSÃO

As imagens digitais podem ser organizadas no formato de vetores de dados. A Figura 4 é uma representação detalhada da maneira como os valores dos componentes R, G e B são distribuídos nos conjuntos de dados relacionados às amostras de chocolate. Essa representação gráfica favorece uma análise das características das imagens, destacando a complexidade dos padrões de cores presentes nos chocolates.

Figura 4 – Canais R,G,B das amostras de chocolate



Fonte: Autoria própria

O gráfico proporciona uma visualização detalhada da distribuição das intensidades de cor no canal RGB. Onde o eixo x representa os canais dos canais RGB, e o eixo y representa a frequência de ocorrência de cada valor de intensidade.

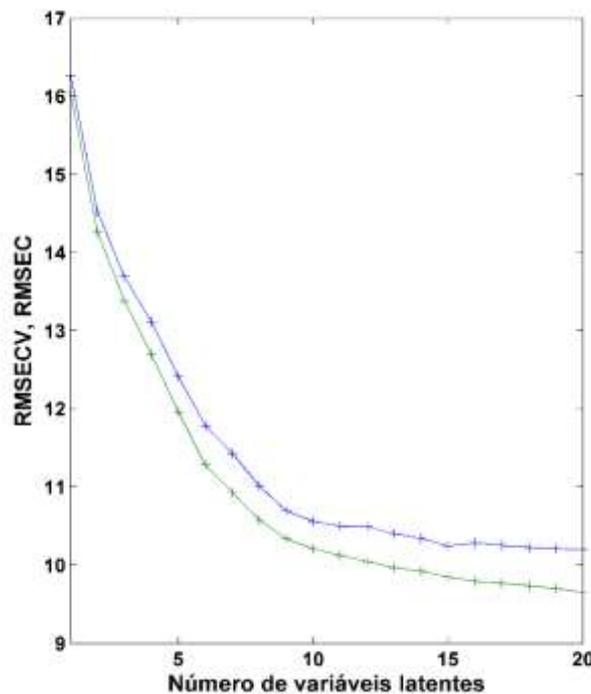
As barras no gráfico representam a frequência de ocorrência de cada valor de intensidade, oferecendo uma percepção clara de como as cores estão distribuídas na imagem. À medida que percorremos o eixo das intensidades, as alturas das barras refletem a intensidade associada a cada valor, destacando áreas onde a intensidade é mais pronunciada ou menos prevalente.

O próximo passo consistiu na separação dos dados em conjunto de calibração e validação externa. Dessa forma, as amostras do conjunto de validação externa não fizeram parte da etapa de calibração. Para esta etapa utilizou-se o algoritmo kenston aplicado separadamente levando em consideração os diferentes *smartphones* utilizados na coleta das imagens. (Kennard; Stone, 1969). Um total de 1275 amostras foi selecionado para compor o conjunto de calibração (425 amostras de cada um dos *smartphones*), e um total de 426 amostras foi utilizado no conjunto de validação externa (142 amostras de cada um dos *smartphones*).

Posteriormente, na construção do modelo multivariado, tornou-se necessário selecionar o número apropriado de variáveis latentes para compor o modelo PLS, visando correlacionar os sinais extraídos da Figura 4 com a porcentagem de cacau presente nas amostras de chocolate.

A seleção apropriada do número de variáveis latentes foi realizada por meio do emprego do procedimento de validação cruzada (*Cross Validation*) (Brereton, 2000). O número de variáveis latentes foi determinado com base nos resultados da raiz quadrada do erro quadrático médio da validação cruzada (RMSECV). No contexto do modelo de determinação da porcentagem de cacau em chocolates, foi identificado um total de 10 variáveis latentes sendo a quantidade adequada, como pode-se observar na Figura 5.

Figura 5 – Número de variáveis latentes



Fonte: Autoria própria

Essa quantidade foi escolhida devido ao fato de que, no gráfico acima, após o uso de 10 variáveis, observou-se uma estabilização nos dados, mantendo-se em uma tendência linear.

Posteriormente, o modelo elaborado passou por uma avaliação criteriosa a fim de identificar potenciais outliers nos conjuntos de calibração e validação. Entretanto, não foram identificados outliers significativos que demandassem exclusão do conjunto de dados. Outliers são observações que se destacam por se comportarem de forma divergente em relação à maioria dos dados, podendo surgir devido a erros laboratoriais ou instrumentais (Silva *et al.*, 2012). Esse processo de verificação assegurou a robustez e confiabilidade do modelo desenvolvido, fornecendo uma base sólida para a posterior interpretação dos resultados obtidos na determinação da percentagem de cacau em amostras de chocolate.

Portanto, o modelo PLS foi construído correlacionando os histogramas RGB e os valores de referência (percentagem de cacau), utilizando 10 variáveis latentes e com os dados centrados na média. O próximo passo consistiu em validar o modelo construído através da determinação de alguns parâmetros de mérito.

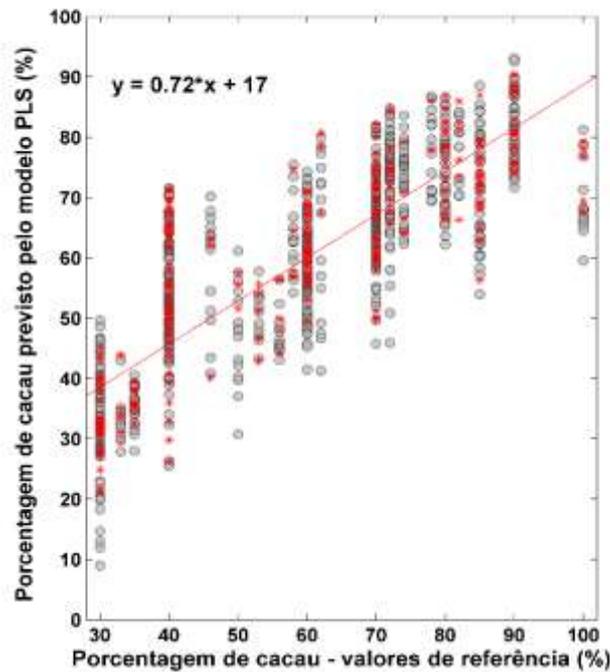
A exatidão do modelo PLS pode ser avaliada considerando o resultado dos erros médios quadráticos da calibração (RMSEC) e da validação externa (RMSEP), conforme Tabela 2, bem como através do gráfico do ajuste (Figura 6).

Tabela 2 – Parâmetros de mérito do modelo PLS

Parâmetros de mérito		Modelo PLS
Exatidão	RMSEC ^a	10,24
	RMSEP ^a	10,52
Coeficiente de correlação		0,75
	RPD _{cal}	1,87
	RPD _{val}	1,98
Sensibilidade analítica ^{-1 a}		2,64
Limite de detecção ^a		8,72
Limite de quantificação ^a		26,43

^a (% de cacau). Fonte: Autoria própria

Figura 6 - Ajuste do modelo multivariado para determinação da porcentagem de cacau das amostras de calibração



Fonte: Autoria própria

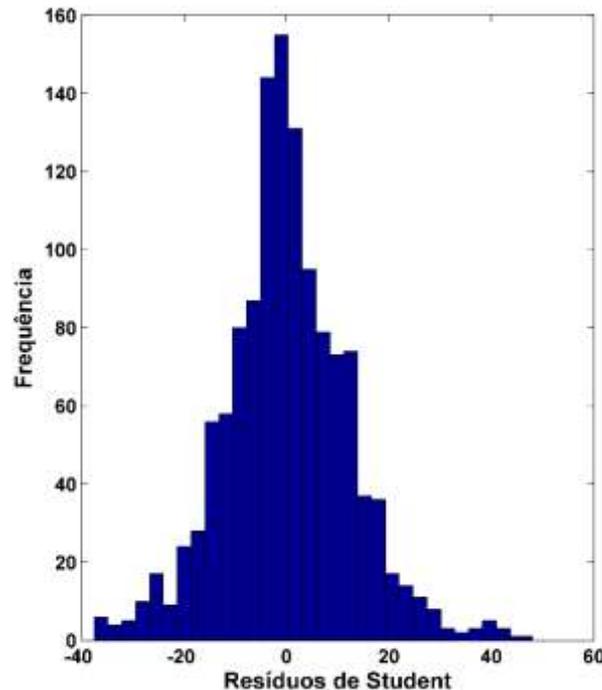
Os resultados de RMSEC e RMSEP indicam que o número de variáveis latentes utilizados em cada um dos modelos estão corretas, não havendo sobreajuste ou subajuste. O gráfico de ajuste mostra, a comparação entre os resultados derivados dos valores de referência do rótulo e as estimativas proporcionadas pelo modelo. Por meio do coeficiente de correlação, é possível analisar a qualidade do ajuste do modelo. Os valores desse coeficiente variam de 0 a 1, onde 0 indica ausência de correlação entre os valores do método de referência e os previstos pelo modelo, enquanto 1 representa uma relação linear perfeita entre eles (Porep et al., 2015).

No contexto deste estudo, o coeficiente de correlação obtido foi de 0,75, evidenciando um valor considerado apropriado, especialmente em contextos nos quais existe uma considerável variabilidade no modelo, como neste caso em que três diferentes *smartphones* foram empregadas na coleta das imagens. Esse valor, indica uma relação significativa entre as variáveis independentes e dependentes, ressaltando a eficácia do modelo na previsão dos valores de referência do rótulo.

A avaliação da linearidade dos modelos de PLS envolveu a análise do gráfico de resíduos de Student, conforme representado na Figura 7. O gráfico dos resíduos Student revela uma notável semelhança com uma distribuição gaussiana, indicando que os resíduos possuem

uma distribuição aleatória. Esse comportamento sugere que os dados estão bem ajustados ao modelo linear para PLS.

Figura 7 - Resíduos de Student para o modelo multivariado na determinação da porcentagem de cacau.



Fonte: Autoria própria

A partir desses resultados, é possível ver que os dados demonstram uma propensão significativa a se alinhar a um modelo linear, uma vez que a distribuição dos resíduos assemelha-se a uma distribuição normal. Essa coerência entre o modelo e os resíduos fortalece a confiabilidade da análise e sugere uma adequada representação do fenômeno em estudo.

Os resultados de RPD (Tabela 2) podem ser considerados satisfatórios uma vez que se encontram na faixa entre 2,4 e 1,5 (Botelho; Mendes; Sena, 2013).

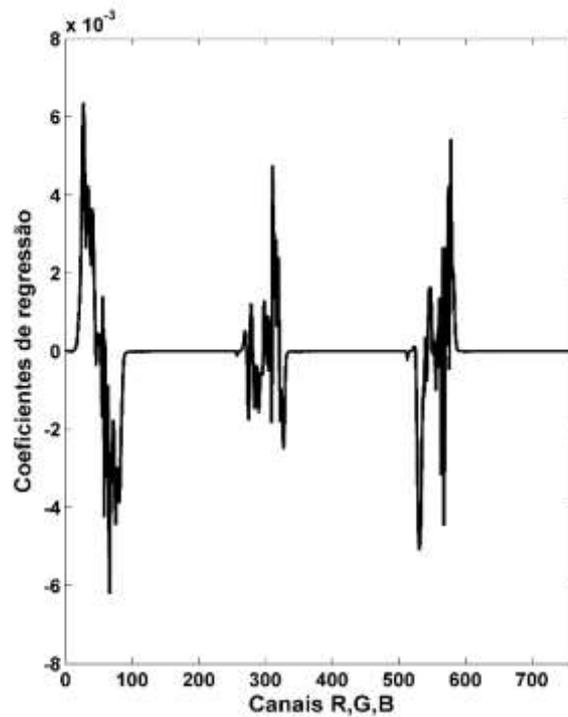
O inverso da sensibilidade analítica (Tabela 2) permite estabelecer a menor diferença de porcentagem de cacau entre as amostras que pode ser distinguida pelo método. Sendo assim, é possível fazer a distinção de amostras com diferença de porcentagem de cacau da ordem de 2,6%.

Os limites de detecção e quantificação do modelo multivariado mostram a menor porcentagem de cacau que pode ser detectada e quantificada, respectivamente. Os resultados alcançados são coerentes com as porcentagens consideradas nesse estudo (30 a 100%). Portanto, os resultados indicam que o modelo PLS é capaz de detectar a porcentagem de cacau

de 8,72%, porém somente consegue quantificar as amostras com porcentagem de cacau superior a 26,43%.

Finalmente, para identificar as variáveis mais significativas nos canais R, G, e B que contribuem para o desenvolvimento do modelo, o vetor de coeficiente de regressão foi considerado (Figura 8). Em geral, as porções iniciais dos canais R (0 a 255), G (256 a 512) e B (513 a 768) foram mias importantes para o modelo na quantificação da porcentagem de cacau.

Figura 8 – Coeficiente de regressão obtidos para os modelos PLS



Fonte: Autoria própria

6 CONCLUSÃO

O presente estudo propôs um método alternativo para determinar a porcentagem de cacau em amostras de chocolates, integrando técnicas de análise de imagem digital e calibração multivariada. A calibração multivariada, empregando a ferramenta de regressão PLS, revelou-se eficaz na correlação entre os histogramas de cor das amostras de chocolate e a porcentagem de cacau. A escolha adequada de 10 variáveis latentes foi determinada por meio de validação cruzada, assegurando a exatidão do modelo.

A abordagem mostrou-se eficaz, proporcionando resultados adequados de parâmetros de mérito e contribuindo para a área de análise de alimentos.

A integração de técnicas de imagem digital e calibração multivariada proporciona uma metodologia eficaz, econômica e ambientalmente sustentável, superando desafios associados a métodos tradicionais. A validação rigorosa do modelo confirma sua confiabilidade, destacando o potencial dessa abordagem para aplicações práticas na indústria alimentícia, e abre perspectivas para futuras aplicações e pesquisas inovadoras nesse campo.

REFERÊNCIAS

ASTM - Annual Book of ASTM Standards. **Standard Practices for Infrared Multivariate Quantitative Analysis** - E1655-00, ASTM International, West Conshohocken: Pennsylvania, USA, 2000.

ASSAD, E. D. *et al.* Identificação de impurezas e misturas em pó de café por meio de comportamento espectral e análise de imagens digitais. **Pesquisa Agropecuária Brasileira**, v. 37, p. 211-216, 2002. Disponível em: <https://www.scielo.br/j/pab/a/TTggqNzNcTkHMq8F9z6WbQz/abstract/?lang=pt#>. Acesso em: 15 abr. 2023.

BAQUETA, M. R.; COQUEIRO, A.; VALDERRAMA, P. Brazilian coffee blends: a simple and fast method by near-infrared spectroscopy for the determination of the sensory attributes elicited in professional coffee cupping. **Journal of Food Science**, v. 84, p. 1247-1255, 2019. Disponível em: <https://ift.onlinelibrary.wiley.com/doi/full/10.1111/1750-3841.14617>. Acesso em: 16 maio 2023.

BAQUETA, M. R. *et al.* Quality control parameters in the roasted coffee industry: a proposal by using microNIR spectroscopy and multivariate calibration. **Food Analytical Methods**, v. 13, p. 50-60, 2020. Disponível em: <https://link.springer.com/article/10.1007/s12161-019-01503-w>. Acesso em: 16 maio 2023.

BELTRAME, K. K. *et al.* Soil Organic Carbon Determination Using NIRS: Evaluation of Dichromate Oxidation and Dry Combustion Analysis as Reference Methods in Multivariate Calibration. **Journal of the Brazilian Chemical Society**, v. 27, p. 1527-1532, 2016. Disponível em: <https://www.scielo.br/j/jbchs/a/DbfKMP9dTcVwMVBGyCTLDmK/abstract/?lang=en>. Acesso em: 14 maio 2023.

BOTELHO, B. G.; MENDES, A. A. P.; SENA, M. M. Implementação de um Método Robusto para o Controle Fiscal de Umidade em Queijo de Minas Artesanal. Abordagem Metrológica Multivariada. **Química Nova**, v. 36, p. 1416-1422, 2013. Disponível em: <https://www.scielo.br/j/qn/a/Ckwdj8fNdDGrcbkTkXHtfwg/?lang=pt>. Acesso em: 15 maio 2023.

BOTELHO, B. G.; ASSIS, L. P.; SENA, M. M. Development and analytical validation of a simple multivariate calibration method using digital scanner images for sunset yellow determination in soft beverages. **Food Chemistry**, v. 159, p. 185-180, 2014. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0308814614004397>. Acesso em: 26 abr. 2023.

BOTELHO, B. G.; DANTAS, K.C.; SENA, M.; M. Determination of allura red dye in hard candies by using digital images obtained with a mobile phone and N-PLS. **Chemometrics and Intelligent Laboratory Systems**, v. 167, p. 44-49, 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169743916304956>. Acesso em: 30 mar. 2023.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. Legislação. **Resolução RDC n. 264, de 22 de setembro de 2005**. 2005. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2005/rdc0264_22_09_2005.html. Acesso em: 01 abr. 2023.

BRASIL. ANVISA - Agência Nacional de Vigilância Sanitária. **Resolução da diretoria colegiada - RDC nº 166, de 24 de julho de 2017**, 2017. Disponível em: http://antigo.anvisa.gov.br/documents/10181/2721567/RDC_166_2017_COMP.pdf/d5fb92b3-6c6b-4130-8670-4e3263763401. Acesso em: 20 maio 2023.

BRASIL. ABICAB - Associação Brasileira da Indústria de Chocolates, Amendoim e Balas. **Pesquisas e estatísticas: chocolate: produção, consumo aparente, exportação e importação**. 2023. Disponível em: <https://www.abicab.org.br/>. Acesso em: 01 abr. 2023.

BRERETON, R. G. Introduction to Multivariate Calibration in Analytical Chemistry. **The Analyst**, Brissol, Reino Unido, v. 125, p. 2125-2154, 2000.

CARRILHA, F.; GUINÉ, R. **Avaliação da cor de peras secadas por diferentes métodos**, 2010. Disponível em: https://repositorio.ipv.pt/bitstream/10400.19/1356/1/2010_Viseu_Acta_Poster%20Fatima%20cor.pdf. Acesso em: 20 maio 2023.

CHEVALLEY, J. Chocolate flow properties. *In: Industrial chocolate manufacture and use*. Editado por S. T. Beckett, 2. ed. Inglaterra: Clays Ltd., p. 139-155, 1994.

CURBANI, L.; GELINSKI, J. M. L. N.; BORGES, E. M. Determination of ethanol in beers using a flatbed scanner and automated digital image analysis. **Food Analytical Methods**, v. 13, p. 249-259, 2020. Disponível em: <https://link.springer.com/article/10.1007/s12161-019-01611-7>. Acesso em: 16 maio 2023.

DAMASCENO, D. *et al.* Análise multivariada de imagens na química: um experimento para determinação do PH águas potáveis. **Química nova**, v. 38, p. 836-841, 2015. Disponível em: <https://www.scielo.br/j/qn/a/xHSLSCQxyvCzK54HFqYm45M/?lang=pt#>. Acesso em: 20 maio 2023.

FULGÊNCIO, A. C. C. *et al.* Development of a Simple and Rapid Method for Color Determination in Beers Using Digital Images. **Food Analytical Methods**. v.13, p. 303-312, 2020. Disponível em: <https://link.springer.com/article/10.1007/s12161-019-01634-0>. Acesso em: 19 maio 2023.

GRUDPAN, K. *et al.* Applications of everyday IT and communications devices in modern analytical chemistry: A review. **Talanta**, v. 136, p. 84-94, 2015. Disponível em: <https://www.sciencedirect.com/science/article/pii/S003991401401025X>. Acesso em: 01 abr. 2023.

JARQUE, C. M.; BERA, A. K. A test for normality of observations and regression residuals. **International Statistical Review/Revue Internationale de Statistique**, p. 163-172, 1987. Disponível em: <https://www.jstor.org/stable/1403192>. Acesso em 10 nov 2023.

KENNARD, R. W., STONE, L. A. Computer Aided Design of Experiments. **Techometrics**, v. 11, p. 137-148, 1969. Disponível em: <https://www.tandfonline.com/doi/epdf/10.1080/00401706.1969.10490666?needAccess=true>. Acesso em: 23 abril 2023.

LEITE, L. R. C. **Estudo de Competitividade do Cacau e Chocolate no Brasil: Desafios na Produção e Comércio Global**. Brasília: Ministério da Indústria, Comércio Exterior e Serviços, p.128, 2018.

LOPEZ-MOLINERO A.; JIMENEZ-LAMANA J. The chromatic variance matrix in digital image-based methods for the potential evaluation as quantitative procedures. **Chemom Intell Lab Syst.** v. 157 p. 29-34, 2016. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169743916301435?via%3Dihub>. Acesso em: 15 mar. 2023.

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

MARTIN, A. V. Chocolate confectionery. *In*: MAN, C. M. D.; JONES, A. A. **Shelf life evaluation of foods**. London, New York: Blackie Academic, 1994.

MERCADO DO CACAU. **Conheça as maiores empresas de chocolate do mundo**. Mercado do Cacau, 2018. Disponível em: <https://shre.ink/Mercado-do-Cacau>. Acesso em: 23 maio 2023.

MIGOT, I. *et al.* Avaliação de cor de diferentes marcas e sistemas de resina composta por espectrofotometria. **Journal of Dentistry & Public Health (inactive/archive only)**, v. 8, p. 68-75, 2017. Disponível em: <https://journals.bahiana.edu.br/index.php/odontologia/article/view/1278>. Acesso em 06 maio 2023.

MORDOR INTELLIGENCE. **Mercado de Chocolate - Crescimento, Tendências E Previsões (2023 - 2028)**. 2023. Disponível em: <https://www.mordorintelligence.com/pt/industry-reports/chocolate-market>. Acesso em: 21 maio 2023.

MORGANO, M. A. et al. Determinação de açúcar total em café cru por espectroscopia no infravermelho próximo e regressão por mínimos quadrados parciais. **Química Nova**, v. 30, pág. 346-350, 2007. Disponível em: <https://www.scielo.br/j/qn/a/W5h5zDVgQQmynhXdRmVjcFC/?format=pdf&lang=pt>. Acesso em: 18 nov 2023.

NACANO, E. T. **Avaliação do comportamento reológico e colorimétrico de chocolates**. 2013. Trabalho de Conclusão de Curso – Faculdade de Engenharia de Alimentos, Universidade Tecnológica Federal do Paraná – UTFPR. Campo Mourão – PR. 2013. Disponível em: <http://riut.utfpr.edu.br/jspui/handle/1/6785>. Acesso em: 22 maio 2023.

OSTROWSKA-LIGEZA, E.; MARZEC, A.; GÓRSKA, A. *et al.* A comparative study of thermal and textural properties of milk, white and dark chocolates. **Thermochimica acta**, v. 671, p. 60-69, 2019. Disponível

em:<https://www.sciencedirect.com/science/article/pii/S0040603118305793>. Acesso em: 03 abr. 2023.

OTTO, M.; **Chemometrics**. 1. ed. Weinheim: Wiley, Alemanha, 1999.

Porep, J. U.; Kammerer, D. R.; Carle, R. On-line application of near infrared (NIR) spectroscopy in food production. **Trends in Food Science & Technology**, v. 46, p. 211-230, 2015. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0924224415002174>. Acesso em: 15 de novembro de 2023.

ROSA, T. R. *et al.* Redgim como aplicativo de smartphone para aplicações quimiométricas por meio de análise de imagens: um uso em pls. **Química nova**, v. 45, p. 550-559, 2022.

Disponível em: <https://www.scielo.br/j/qn/a/FFSvvhMzLkXNyvNFR8FwzPj/#>. Acesso em: 20 maio 2023.

SANTANA, F. *et al.* **Avaliação da cor dos frutos de diferentes genótipos de bananeiras por colorímetro digital**, p. 21., 2010. Disponível em:

<https://www.alice.cnptia.embrapa.br/handle/doc/873833>. Acesso em: 29 maio 2023.

SANTOS, D. A. *et al.* Omega-3 and ômega-6 determination in Nile tilapia's fillet based on MicroNIR spectroscopy and multivariate calibration. **Journal of Brazilian Chemical Society**, v. 31, p. 1883-1890, 2020. Disponível em:

<https://www.scielo.br/j/jbchs/a/NS9pzKzMVnVtJsn4DpYcqVJ/abstract/?lang=en>. Acesso em: 14 maio 2023.

SANTOS, I. A. *et al.* NIR and MIR spectroscopy for quick detection of the adulteration of cocoa content in chocolates. **Food Chemistry**, 2021. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0308814621000972>. Acesso em: 08 maio 2023.

SANTOS, D. A.; LIMA, K. P.; MARÇO, P. H.; VALDERRAMA, P. Vitamin C Determination by Ultraviolet Spectroscopy and Multiproduct Calibration. **Journal of the Brazilian Chemical Society**, v. 27, p. 1912-1917, 2016. Disponível em:

<https://www.scielo.br/j/jbchs/a/9MnSnRdtGnVPztWhFXSZcfp/abstract/?lang=en>. Acesso em: 15 maio 2023.

SANTOS, V. J. **Avaliação das diferentes fases de lactação e caracterização nutricional do leite humano, na forma crua e pasteurizada, e emprego da espectroscopia na região do infravermelho próximo e ferramentas quimiométricas**. 2021. Tese (Doutorado- Programa de Pós-Graduação em Química), Universidade Estadual de Maringá. Maringá - PR, 2021.

Acesso em: 16 maio 2023.

SILVA, M. A.; FERREIRA, M. H.; BRAGA, J. W. B.; SENA, M. M. Development and validation of a multivariate calibration method for determination of amoxicillin in suspension formulations by near infrared spectroscopy. **Talanta**, v. 89, n., p. 342-51, 2012. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0039914011011076>. Acesso em 17 de novembro de 2023.

SOUZA, A. S. L. **Avaliação da estabilidade térmica e oxidativa de chocolates amargos.** 2010. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos), Universidade Federal da Paraíba. João Pessoa – PB. 2010. Disponível em: https://repositorio.ufpb.br/jspui/handle/tede/4087?locale=pt_BR. Acesso em: 10 abr. 2023.

THORNTON, P. The US Chocolate Market-Current and Future. **The Manufacturing Confectioner**, p.35, 2007. Disponível em: <https://www.gomc.com/firstpage/200709035.pdf>. Acesso em: 15 abr. 2023.

VALDERRAMA, P. **Avaliação de figuras de merito em calibração multivariada na determinação de parametros de controle de qualidade em industria alcoleira por espectroscopia no infravermelho proximo.** 2005. Dissertação (Mestrado do departamento de química analítica) - Universidade Estadual de Campinas, Instituto de Química, Campinas - SP, 2005. Disponível em: <https://hdl.handle.net/20.500.12733/1601228>. Acesso em: 29 mai. 2023.

VALDERRAMA, P., BRAGA, J. W. B., POPPI, R. J. Validation of Multivariate Calibration Models in the Determination of Sugar Cane Quality Parameters by Near Infrared Spectroscopy. **Journal of the Brazilian Chemical Society**, v. 18, p. 259-266, 2007. Disponível em: <https://www.scielo.br/j/jbchs/a/YvHdLDHWk6vnMBPhg4PKJvm/abstract/?lang=en>. Acesso em: 12 maio 2023.

VÁSQUEZ, R. H. **Comparação da medida de cor de pescado entre sistema de visão computacional e colorímetro convencional no frescor do pescado.** 2015. Dissertação (Mestrado- Programa de Pós-Graduação em Engenharia e Ciência de Alimentos), Universidade Federal do Rio Grande do Sul - UFRGS. Rio Grande do Sul – RS. 2015. Acesso em: 16 maio 2023.

VITTORAZZI, B. V. *et al.* Classificando cédulas brasileiras (R\$) usando análise de imagem por smartphone. **Química Nova**, v. 43, p. 447–454, 2020. Disponível em: <https://www.scielo.br/j/qn/a/vvHrTwKJGKfts7L4LJp5X4N/?lang=pt#>. Acesso em: 02 maio 2023.