

UNIVERSIDADE FEDERAL DO ABC  
Trabalho de Conclusão de Curso | Bacharelado em Química

Ednaldo Dario da Silva

# **Boas Práticas em Planejamento Experimental**

Santo André

NOVEMBRO – 2022

Ednaldo Dario da Silva

## **Boas Práticas em Planejamento Experimental**

Monografia de Trabalho de Conclusão de Curso, apresentado ao Bacharelado em Química da UFABC para obtenção do título de Bacharel em Química

**Orientador:** Prof. Dr. Bruno Guzzo da Silva

**Coorientador:** Me. Renan da Silva Nunes

## RESUMO

Como advento das tecnologias computacionais, a quimiometria, bem como o planejamento experimental, se tornaram áreas de enorme interesse e essencial a qualquer químico. Inúmeros problemas práticos podem ser melhor enfrentados no escopo do planejamento experimental, que oferece grande vantagem no estudo de fenômenos e sistemas, especialmente na construção de modelos empíricos, no seu entendimento e otimização de variáveis. No entanto, apesar de ser uma área de grande interesse, existem poucos materiais que apresentam discussões sobre boas práticas e possíveis problemas práticos enfrentados por aqueles que fazem uso da ferramenta. Assim, este trabalho tem como objetivo apresentar alguns tópicos sobre boas práticas em planejamento experimental, discutindo os problemas envolvidos e apresentando sugestões de abordagens mais adequadas a cada caso. Para isso, realizaram-se pesquisas bibliográficas e posterior análise dos materiais selecionados, a fim de reunir as informações que permitiram discutir cada um dos tópicos apresentados, em conjunto com descrição do problema, sugestão de abordagem e apresentação de caso exemplo. No geral, apesar dos protocolos existentes de uso do planejamento experimental, detalhes práticos e erros comuns são ausentes ou pouco discutidos na literatura. Assim, o presente trabalho representa um material de consulta importante para aqueles que se venturam no planejamento experimental e desejam evitar erros que possam comprometer o estudo, pois reúne tópicos de boas práticas, pouco descritos em materiais da área, capazes de evitar o comprometimento total ou parcial do estudo devido a falhas que poderiam ter sido evitadas.

**Palavras-chave:** planejamento experimental; quimiometria; boas práticas

## ABSTRACT

With the advent of computational technologies, chemometrics, as well as experimental design, have become areas of enormous interest and essential to any chemist. Numerous practical problems can be better addressed in the scope of experimental design, which offers great advantage in the study of phenomena and systems, especially in the construction of empirical models, in their understanding and optimization of variables. However, despite being an area of great interest, there are few materials that present discussions about good practices and possible practical problems faced by who use the tool. Thus, this work aims to present some topics about good practices in experimental design, discussing the problems involved and presenting suggestions for the most appropriate approaches for each case. For this, bibliographic research and subsequent analysis of the selected materials were carried out, in order to gather the information that enable the discussion of each topic presented, together with a description of the problem, suggestion of approach and presentation of example cases. Overall, despite existing protocols for using experimental design, practical details and common errors are absent or at least little discussed in the literature. Thus, the present work represents an important reference material for those who venture into experimental planning and wish to avoid errors that could compromise the study, as it brings together topics of good practices that are little described in materials in the area and that are capable of avoiding total or partial commitment of the study, due to failures that could have been prevented.

**Keywords:** experimental design; chemometrics; good practices.

# SUMÁRIO

1	Introdução.....	7
2	Objetivos.....	9
2.1	Objetivo Geral.....	9
2.2	Objetivos específicos.....	9
3	Revisão Bibliográfica .....	10
3.1	Quimiometria .....	10
3.2	Planejamento de Experimentos.....	13
3.3	Tipos de planejamento .....	15
3.3.1	Planejamentos Fatoriais Completos .....	15
3.3.2	Planejamento Fatorial Fracionado .....	17
3.3.3	Planejamento Plackett – Burman.....	20
3.3.4	Planejamento de Composto Central .....	22
3.3.5	Planejamento Box-Behnken .....	25
3.4	Etapas de aplicação .....	27
3.5	Softwares .....	32
4	Metodologia .....	34
5	Resultados e Discussões.....	35
5.1	Conhecendo o sistema.....	35
5.2	Tempo como variável de entrada .....	37
5.3	Consideração de fatores qualitativos.....	42
5.4	Consideração de respostas qualitativas .....	44
5.5	Variáveis discretas .....	46
5.6	Escolha do tipo de planejamento.....	48
5.7	Faixa Investigada .....	51
5.8	Importância do ponto central .....	53

5.9	Planejamentos Fracionados .....	57
5.10	Planejamentos Plackett-Burman .....	60
5.11	Randomização .....	62
5.12	Análise dos resultados estatísticos .....	65
5.13	Validação Final.....	70
5.14	Uso de Softwares.....	72
5.15	Outros erros frequentes .....	74
6	Conclusão.....	76
7	Referências.....	78

# 1 INTRODUÇÃO

O desenvolvimento computacional e tecnológico possibilitou o aumento da instrumentação nos laboratórios químicos e, conseqüentemente, o aumento da complexidade e do volume de dados gerados a partir de amostras em estudos multivariados. A quimiometria, uma área da química, se propõe a resolver as dificuldades frente a obtenção e análise de grande volume de dados complexos fazendo uso de métodos estatísticos, computação e matemática, em diferentes aplicações que possibilitam a extração de informação científica relevante.

Parte da geração dessas informações relevantes deve-se ao uso de métodos estatísticos para planejar e selecionar procedimentos experimentais, a fim de otimizar a extração de informação e fazer uso da análise estatística de dados. Para isso, é aplicado um conjunto de ferramentas estatísticas e matemáticas na definição de uma estratégia viável e adequada ao estudo multivariado, o chamado planejamento experimental.

O planejamento experimental se insere como parte essencial da quimiometria e é caracterizado como um conjunto de ferramentas matemáticas e estatísticas para definição de um delineamento experimental, que permitem minimizar os custos de execução e fornecer uma estrutura otimizada, na qual os fatores significativos para uma dada resposta são variados sistematicamente, e cuja análise subsequente permitem identificar as condições ótimas para um sistema.

São diversos os tipos de planejamento experimentais que podem ser empregados em uma infinidade de estudos e objetivos, seja para selecionar as variáveis de maior impacto sobre a resposta de um sistema, ou otimizar as respostas do sistema por completo, definindo parâmetros ótimos, modelo estatístico e gerando superfície de resposta. Para cada tipo de planejamento existe uma técnica de construção própria, baseada em princípios matemáticos, estatísticos e geométricos, bem como análises específicas e informações diferentes que podem ser obtidas, a depender do propósito e escopo do estudo.

Em todos os casos, existe uma série de cuidados e boas práticas a serem seguidas para que sejam obtidos resultados confiáveis, com validade estatística e minimização dos custos diversos. Embora sejam ações necessárias e que oferecem

uma perspectiva de aumento de qualidade, muito pouco é descrito na literatura sobre as abordagens mais apropriadas diante de situações que podem comprometer o estudo e seus resultados.

O presente trabalho então se propõe a apresentar tópicos sensíveis referente as boas práticas de uso do planejamento experimental, bem como abordagens e pontos de atenção, para minimização dos erros e melhoria da qualidade do estudo, com apresentação de informações derivadas da prática experimental e casos exemplos.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

Este trabalho objetiva reunir e apresentar conhecimentos acerca de boas práticas na aplicação do planejamento experimental, abordando tópicos usualmente não descritos ou pouco encontrados em literatura da área.

### 2.2 OBJETIVOS ESPECÍFICOS

- Apresentar tópicos de boas práticas na aplicação do planejamento experimental, sendo eles:
  - Conhecimento prévio sobre o sistema;
  - Uso do tempo como variável de entrada;
  - Uso de fatores e respostas qualitativos;
  - Uso de variáveis discretas;
  - Escolha do tipo de planejamento adequado;
  - Definição da faixa de investigação apropriada;
  - Importância do ponto central;
  - Particularidade do uso de planejamentos Fatoriais Fracionados;
  - Particularidade do uso de planejamentos Placket-Burman;
  - Uso da randomização;
  - Análise dos resultados estatísticos;
  - Importância da validação final;
  - Bom uso e escolha de softwares estatísticos.
- Discutir cada um dos tópicos e propor uma abordagem adequada a ser seguida em cada caso;
- Descrever e discutir casos exemplos que auxiliem a compreensão do tópico apresentado;

### 3 REVISÃO BIBLIOGRÁFICA

#### 3.1 QUIMIOMETRIA

O desenvolvimento computacional e tecnológico na química possibilitou o aumento da instrumentação nos laboratórios e, conseqüentemente, o aumento da complexidade e do volume de dados gerados a partir de amostras em estudos multivariados. Mais parâmetros podem ser analisados em curtos espaços de tempo, possibilitando o acesso a uma maior quantidade de informação científica a respeito das amostras e processos. No entanto, esse potencial informacional complexo necessita de ferramentas de planejamento estatístico e métodos computacionais que permitam interpretar adequadamente o grande volume de dados gerados.

A quimiometria, uma área da química, se propõe então a resolver as dificuldades frente a obtenção e análise dos dados complexos gerados nessas situações, fazendo uso de métodos estatísticos, computação e matemática, em diferentes aplicações que possibilitam a extração de informação científica relevante a partir desses conjuntos de dados (BRERETON, 2013; GEMPERLINE, 2006; HOPKE, 2003).

A quimiometria nasceu, não oficialmente, na década de 60, período no qual os sistemas computacionais passavam a fazer parte da vida de alguns poucos cientistas que poderiam usufruir dos aparatos computacionais da época ou da imersão interdisciplinar necessária para fazer uso dos recursos tecnológicos e das ferramentas. Com o maior alcance dos sistemas computacionais, as análises matemáticas e estatísticas de grandes volumes de dados de natureza multivariada passaram a ser incorporadas nas rotinas de cientistas de diversas áreas e na resolução de problemas outrora custosos ou impossíveis de serem abordados (BRERETON, 2013; BROWN, 2017).

A palavra “Quimiometria” passou a existir no mundo somente em 1971, criada pelo químico analítico Bruce Kowalski e pelo químico orgânico Svant Wold, fundadores da International Chemometrics Society (Sociedade Internacional de Quimiometria), consolidando o progresso da quimiometria como um novo campo da química (BRERETON, 2013; MASSART et al., 1997; HOPKE, 2003; KATEMAN,

1990). Embora a alcunha seja atribuída somente a Kowalski e Wold, vale pontuar a contribuição de pesquisadores como William Gosset (pseudônimo Student), químico e estatístico, e George Box, estatístico e estudante de química, que trabalharam em problemáticas semelhantes aos quimiométricos da época e que propuseram soluções até hoje relevantes do escopo estatístico na tratativa de problemas experimentais na química, como o teste t de Student e o design experimental e otimização de Box (BRERETON, 2013; BROWN, 2017; KATEMAN, 1990).

A quimiometria passou a ser reconhecida como disciplina pela comunidade científica apenas na década de 80, culminando na criação de journals especificamente dedicados a área e a publicação de livros acadêmicos. No início, a quimiometria não gozava de completa aceitação, mas com o aumento de estudos relevantes e resolução de problemas relacionados a interpretação de dados complexos em química, o ceticismo e a resistência passaram a diminuir, evidenciado pela inserção gradual dos estudos associados a quimiometria em journals, como feito pela *Analytica Chimica Acta*, primeiro journal a introduzir uma seção dedicadas aos estudos de otimização, métodos estatísticos e matemáticos aplicados em química (HOPKE, 2003). A partir deste ponto, a quimiometria passa então a ser incorporada em estudos de graduação e pós-graduação em química, se torna tema em workshops, conferências e cursos de treinamento em várias universidades pelo mundo, se sobrepondo ao estigma de subdisciplina (BRERETON, 2013; HOPKE, 2003; TEÓFILO; FERREIRA, 2006).

A quimiometria hoje pode ser entendida como uma área da química extremamente importante, sobretudo para a química analítica, que faz uso da matemática e dos métodos estatísticos para planejar e selecionar procedimentos experimentais ótimos, a fim de otimizar a extração de informação relevante a partir da análise estatística de dados, que resultam na elucidação de conhecimento científico acerca do sistema químico em estudo (BRERETON, 2013; HOPKE, 2003; WOLD, 1995).

O cientista quimiométrico parte de uma hipótese construída com base no conhecimento prévio que se tem sobre determinado sistema em estudo e com isso planeja um procedimento experimental que permite obter os dados que serão analisados, para que se extraiam deles informações a fim de validar ou não a hipótese. A construção do procedimento, extração de dados, tratamento e interpretação são

feitos com base em ferramentas estatísticas e matemáticas, que em último revelam como os resultados das medidas experimentais (respostas) se relacionam com os observáveis (variáveis), entregando informações sobre o sistema químico (MASSART et al., 1997).

O planejamento e otimização experimental faz parte de um conjunto de métodos quimiométricos, assim como o processamento de sinais, reconhecimento de padrões e calibração multivariada. As metodologias de análise multivariada mais empregadas no estudo de dados provenientes do uso de instrumentos como os de NMR, NIR e HPLC, que geram uma variedade de dados brutos (espectros, cromatogramas e séries temporais), são a PCA (Principal Component Analysis) e a PLS (Partial Least Squares) (FERREIRA et al., 1999).

A PCA, que tem como base a redução de um grande volume de variáveis a combinações lineares das variáveis originais, ou seja, os dados são transformados de modo que possam ser descritos usando um novo sistema de coordenadas de dimensão menor, com as componentes principais que identificam a direção de maior variância dos dados, ao custo de perda de informações devido à redução de dimensionalidade, embora possibilite um melhor manejo e interpretação dos dados. Na PCA os componentes são ordenados de acordo com a variância, assim a primeira variável composta possui a maior variância e é seguida por aquelas de menor variância em ordem de grandeza.

A PLS, assim como a PCA também promove a redução de dimensionalidade, redução de variáveis preditoras, mas neste caso combina funcionalidade de regressão múltipla, fazendo uso das variáveis compostas, como obtidas na PCA, a fim de explicar a variabilidade máxima na resposta (variável dependente) no contexto de regressão. Neste caso, as variáveis compostas são ordenadas de acordo com a correlação com a variável dependente no modelo, tendo, portanto, as variáveis compostas que melhor explicam os dados. Assim, se o interesse do estudo se baseia em obter um grupo das variáveis explicativas, preditoras iniciais, em um modelo linear que explique a maior variabilidade na resposta, o uso de PLS é mais eficaz (JOLLIFE; CADIMA, 2016; ROSIPAL; KRÄMER, 2006).

## 3.2 PLANEJAMENTO DE EXPERIMENTOS

Nas ciências experimentais como a química, as informações necessárias para que se tenha avanço no conhecimento científico e na resolução de problemas são obtidas por meio de experimentos. Segundo Montgomery (2020), experimentos são definidos como testes nos quais as variáveis de entrada são alteradas intencionalmente a fim de se observar mudanças em um sistema e identificar suas causas, estabelecendo uma correlação entre as variáveis importantes e as respostas. Em resumo, experimentos possibilitam ao pesquisador entender o que ocorre com a resposta quando mudanças nas configurações das variáveis de entrada são efetuadas propositadamente.

A disponibilidade de recursos computacionais, junto ao avanço na instrumentação e consolidação do estudo de vários fatores simultaneamente, evidenciou a necessidade de uma estruturação experimental que permitisse explorar esse potencial informacional. A produção de conhecimento científico pode, a partir dessa nova estruturação experimental, fazer uso do recurso tecnológico com aplicação facilitada de ferramentas matemáticas, estatísticas e computacionais na elaboração, execução e análise de experimento.

Nesse cenário, o experimento tem importância crítica na qualidade da informação extraída, que mais tarde servirá de substrato para as conclusões e decisões a serem tomadas, muitas vezes essas representando vantagem estratégicas no mercado, ou passos importantes para o conhecimento científico sobre a performance de um sistema (produto ou processo).

A qualidade e confiabilidade dos resultados produzidos dependem das condições de construção do experimento, e estes devem considerar as informações desejáveis, conhecimentos prévios, possível interação entre variáveis, recursos, tempo e limitações operacionais. Para isso, deve-se aplicar um conjunto de ferramentas estatísticas e matemáticas na definição de uma estratégia viável e adequada ao estudo multivariado, o chamado planejamento experimental.

O planejamento de experimentos é parte essencial da quimiometria e é caracterizado como um conjunto de ferramentas que fazem uso de conhecimentos matemáticos e estatísticos na definição de um delineamento experimental, que

permitem minimizar os custos de execução (tempo, recursos, pessoal e outros), ao passo que fornece uma estrutura otimizada, na qual os fatores significativos para uma dada resposta são variados sistematicamente, e cuja análise subsequente dos dados gerados permitem identificar as condições ótimas para um sistema (GEMPERLINE, 2006; RODRIGUES; IEMMA, 2015; SANTOS et al., 2019).

O planejamento experimental teve origem com Fischer na década de 30, com tratamento de comparação em agricultura em situações nas quais os experimentos tendem a ser em larga escala, tomando muito tempo para serem executados e sujeitos a variações não controladas (campo e clima). Tais condições promoveram o desenvolvimento das técnicas de randomização, replicação, blocagem e do planejamento fatorial fracionado. Os trabalhos em planejamento experimental seguiram com Box, em aplicações em modelagem e otimização na indústria química, que ocasionou o desenvolvimento do planejamento composto central, que lida com modelos de regressão e análises gráficas. E mais tarde, Taguchi fez sua contribuição com propostas focadas em qualidade e melhoria da produtividade, reduzindo a sensibilidade do sistema a variação pelo estabelecimento das configurações ótimas dos chamados fatores de controle (WU; HAMADA, 2009).

As vantagens oferecidas impactam estudos multivariados na química, seja na indústria ou academia, com objetivos diversos que vão desde avaliação simples de efeitos de variáveis (concentração de reagentes, temperatura, pressão, etc) até estudos complexos de otimização e robustez de processos (maximização de rendimento, configuração ótima de variáveis, redução de custo, diminuição de ruído, etc). Sua aplicação tipicamente resulta em redução do número de ensaios e repetições realizadas, e em paralelo é acompanhado da melhoria na qualidade da informação obtida. Neste aspecto já se tem uma vantagem competitiva, uma vez que a quantidade de ensaios pode inviabilizar a realização de um estudo.

Outro ponto está na possibilidade de se estudar mais de uma variável ao mesmo tempo e otimizar simultaneamente mais de uma resposta, com acesso a informações acerca de interações entre variáveis e ampliação do poder de melhoria do estudo. Por fim, tem-se acesso ao erro experimental, o que permite ter uma medida quantitativa da confiabilidade estatística, repetibilidade dos resultados e uma abordagem mais crítica sobre o resultado (BRERETON, 2003; MONTGOMERY, 2020; RODRIGUES, IEMMA, 2015).

No geral, alguns problemas experimentais tipicamente fazem uso do planejamento experimental em sua resolução, como em comparações, seleção de variáveis, geração de superfície de resposta, otimização de sistema e melhoria na robustez. O primeiro deles é caracterizado pela comparação de tratamentos, cujo principal objetivo é comparar diversos tratamentos e selecionar o melhor deles, como no caso de seleção de espécies diferentes de plantas para um mesmo propósito ou enzimas para um mesmo substrato, tipos de catalisadores para uma certa reação, etc.

No estudo de variáveis normalmente se objetiva avaliar seus efeitos sobre as respostas e reduzir uma grande quantidade delas a algumas poucas que são de fato significativas, viabilizando o estudo, melhorando os resultados e tendo graus de liberdade suficiente para estimar efeitos de interação. Com as variáveis significativas selecionadas previamente na análise estatística dos resultados, pode-se visualizar graficamente os efeitos principais e de interação sobre uma dada resposta por meio da plotagem de superfície de resposta, que auxilia no processo analítico, uma vez que oferece uma forma visual de análise. Já a otimização do sistema pode ser exemplificada na química em casos de reações que necessitam de maximização do rendimento a partir da condição ótima definida pelos níveis de fatores, como concentração de reagentes e temperatura, que minimizam efeitos antagônicos ou maximizam a resposta.

Por fim, temos a robustez do sistema, que pode ser tratada no sentido de reduzir a sensibilidade do sistema a variações de ruído por meio da seleção das configurações dos fatores de controle durante manipulação de ruído que promovem a ocorrência de variabilidade, ou seja, obtém-se a configuração ótima dos parâmetros sobre os quais se tem controle (RODRIGUES; IEMMA, 2015; WU; HAMADA, 2009).

### 3.3 TIPOS DE PLANEJAMENTO

#### 3.3.1 Planejamentos Fatoriais Completos

Quando se quer estudar um sistema complexo, constituído de várias variáveis que podem influenciar a resposta, deve-se optar por um planejamento experimental que comporte as variáveis estudadas e seus níveis. Para os casos nos quais a

quimiometria usualmente se insere, que são multivariados, as opções mais eficientes a serem consideradas contemplam o tratamento de diferentes níveis de dois ou mais fatores, e são chamados de planejamentos fatoriais.

O planejamento fatorial completo é a classe mais popular de planejamento experimental e tem a vantagem de poder revelar os efeitos de fatores em um sistema. Essa classe foi introduzida por Fischer, propondo que todos os fatores podem variar simultaneamente e ainda assim os efeitos individuais de cada fator e suas interações podem ser estimados com o tratamento matemático adequado (DEMING; MORGAN, 1988; WU; HAMADA, 2009).

No planejamento fatorial completo, todas as combinações de níveis dos fatores são investigadas e em cada ensaio estará presente uma das combinações lineares possíveis dos níveis dos fatores investigados, de modo que cada ensaio realizado seja único e apresente uma combinação diferente de níveis dos fatores.

Assim, no caso de haver todas as combinações lineares possíveis de níveis de cada fator, tem-se a matriz de um planejamento fatorial completo, cujo número de ensaios é denotado pela expressão  $N^k$ , na qual N representa o número de níveis dos fatores e k representa o número de fatores considerados no estudo. Por exemplo, se em um planejamento temos três fatores (A, B e C), cada um com dois níveis (alto e baixo), este será chamado de planejamento fatorial completo  $2^3$ , ou seja, teremos 8 ensaios que combinam todos os fatores e seus níveis, cuja matriz é apresentada na Tabela 1.

Com isso, temos que número de ensaios a ser realizado em um planejamento fatorial depende do número de fatores e de seus níveis, e cresce exponencialmente conforme o número de fatores considerados aumenta. Assim, pode-se facilmente perceber que para planejamentos com muitos fatores, considerando 2 ou mais níveis, o planejamento fatorial completo é praticamente inviabilizado. (BRERETON, 2003; MONTGOMEY, 2020; RODRIGUES; IEMMA, 2015)

**TABELA 1** – Matriz Planejamento Fatorial Completo  $2^3$ 

<b>Ensaio</b>	<b>Fator 1</b>	<b>Fator 2</b>	<b>Fator 3</b>
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

Fonte: Elaboração Própria

O esquema fatorial completo  $2^k$ , que ocorre quando temos  $k$  fatores com apenas dois níveis, é uma condição muito comum quando se quer realizar observação preliminar dos efeitos dos fatores, principalmente quando se tem 4 ou menos fatores envolvidos e sem a necessidade de descrição mais completa da relação entre fatores e respostas. O caso mais simples de um planejamento fatorial completo é caracterizado por  $2^2$ , composto por dois fatores com dois níveis cada, e apresenta uma matriz com 4 tratamentos, ou seja, são necessários apenas 4 ensaios para abarcar todas as combinações possíveis dos níveis de cada fator, como deve ser em um planejamento fatorial completo. Nesse tipo de planejamento, em apenas dois níveis temos que a resposta apresenta comportamento linear com relação aos fatores envolvidos (ANTONY, 2014; BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003).

### 3.3.2 Planejamento Fatorial Fracionado

Dado o caráter exponencial dos planejamentos fatoriais, temos que considerar a possibilidade de se obter um planejamento completo inviável em casos de maior número de fatores ( $>4$ ), uma clara fraqueza desse tipo de planejamento. Como exemplo, tendo um caso no qual se considera 10 fatores, teríamos uma composição 210 com 1024 ensaios. Neste caso, a alternativa mais apropriada seria realizar uma

triagem dos fatores relevantes e prosseguir com um tratamento mais sofisticado somente com aqueles que são significativos, ou realizar a aplicação de um planejamento fatorial fracionado, que considera um número muito menor de ensaios (ANTONY, 2014; BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003).

Em alguns casos, parte dos ensaios executados apresentam resultados de interações entre fatores que podem ser desconsideradas por serem praticamente insignificantes, devido a interações de primeira ordem apresentarem importância maior do que as interações de ordem maior, não havendo, portanto, necessidade de se prosseguir com um planejamento completo (BRERETON, 2003; BARROS NETO; SCARMINIO; BRUNS, 2001).

Pode-se considerar também a necessidade de se identificar quais são os fatores que afetam significativamente a resposta dentre uma quantidade elevada de fatores iniciais. Como a quantidade inicial de fatores pode ser grande, e não se quer excluir fatores potencialmente relevantes sem a certeza de que realmente não os são, a realização de planejamentos fracionados permite considerar todos eles e avaliar os efeitos posteriormente para um screening dos fatores importantes (BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003).

Os planejamentos fatoriais fracionados são apresentados na forma  $2^{k-p}$ , onde  $k$  é o número de fatores e  $p$  representa a fração do esquema completo  $2^k$ . Por exemplo, no caso de se querer a fração  $\frac{1}{2}$  do planejamento completo  $2^4$  com 16 ensaios, 4 fatores em 2 níveis, teremos  $\frac{1}{2} \times 2^4 = 2^{-1} \times 2^4 = 2^{4-1} = 2^3 = 8$  ensaios. Assim, tomando apenas uma fração do planejamento completo, podemos reduzir pela metade o número de ensaios para 4 fatores em 2 níveis (BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003; RODRIGUES; IEMMA, 2015).

A matriz de planejamento fracionado se distingue da matriz para o caso completo e sua aplicação apresenta características próprias como a possibilidade de confundimento entre os efeitos principais e interações, devido a composição de fatores serem constituídas de produtos dos níveis codificados de outros.

Para que se tenha uma fração adequada do planejamento completo, deve-se manter alguns dos ensaios da matriz completa, de forma que se permita uma variação dos níveis dos fatores a cada ensaio. Assim, a matriz obtida apresenta como

características colunas diferentes umas das outras, números iguais da codificação + e – em cada coluna e para cada ensaio no qual o primeiro fator presente nível +, existirá o mesmo número de ensaios para os demais fatores com nível + e -, em colunas ortogonais (BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003).

**TABELA 2** – Matriz de Planejamento Fracionado  $2^{4-1}$

<b>Ensaio</b>	<b>Fator 1</b>	<b>Fator 2</b>	<b>Fator 3</b>	<b>Fator 4</b>
1	-	-	-	-
2	+	-	-	+
3	-	+	-	+
4	+	+	-	-
5	-	-	+	+
6	+	-	+	-
7	-	+	+	-
8	+	+	+	+

Fonte: Elaboração Própria

Se observarmos o exemplo da Tabela 2, que apresenta uma matriz de um planejamento fracionado  $2^{4-1}$  para o caso de quatro fatores com dois níveis, pode-se notar que o produto dos níveis das colunas dos fatores 1, 2 e 3 corresponde ao nível da coluna do fator 4. Esse achado se refere ao fato de que temos a interação entre os fatores 1, 2 e 3 confundidas com o fator 4, o que significa a impossibilidade de se distinguir a interação entre três fatores dos valores do quarto fator. No entanto, como alento, temos que nem todas as interações são significantes e para o propósito geral do planejamento fracionado, screening de variáveis, as interações são informações de importância secundárias, podendo ser exploradas em outro planejamento (BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003).

### 3.3.3 Planejamento Plackett – Burman

Em 1946, Plackett e Burman, publicaram um artigo mostrando que um planejamento fatorial completo poderia ser fracionado de modo a fornecer uma combinação de fatores que são múltiplos de 4, e com dois níveis. O achado surgiu da necessidade da indústria, de melhoria na qualidade de componentes na fabricação de equipamentos no período de guerra.

Essa nova classe de planejamento permite avaliar o efeito de um grande número de fatores de forma eficiente, como o exemplo, pode-se determinar os efeitos de primeira ordem de 12 fatores em apenas 16 ensaios, como apresentado na Tabela 3, onde a primeira linha é sequencialmente rotacionada para a direita, a fim de se obter a linha subsequente.

A matriz de planejamentos da classe Plackett-Burman é chamada de Hadamard, que são matrizes construídas partindo da definição de uma primeira linha e seguida por permutações cíclicas, com linhas mutuamente ortogonais. Essa classe de experimentos não pode ser descrita como cubos, como as demais classes o são, e, portanto, são planejamentos não geométricos (BARROS NETO; SCARMINIO; BRUNS, 2001; MONTGOMERY, 2020; BROWN et al. , 2020).

O planejamento Plackett-Burman pode conter usualmente 12, 16, 20, 24, 28 e 36 experimentos e é geralmente empregado no processo de seleção de variáveis ou análise dos efeitos dos fatores, para o caso de quando se tem mais de 5 fatores envolvidos. Contudo, Rodrigues & lemma (2015) apontam que, para casos como a aplicação industrial, na qual se tem custo e tempo como limitantes, o planejamento Plackett-Burman pode servir como principal planejamento, sem que haja necessidade de aplicação posterior de outro tipo, possibilitando seguir com ensaios de validação na sequência.

O número de experimentos ( $n$ ) excede o número de fatores ( $k$ ) e pode-se estimar todos os principais efeitos de  $k=n-1$  fatores, com variância mínima. Podemos construir uma matriz de experimentos para se estudar  $n-1$  fatores, embora o recomendado seja escolher um número menor ou igual a  $n-4$  fatores (igual ou superior a 5 fatores), ou seja, um número menor de fatores com relação ao número de ensaio

(BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003; BROWN, 2020; RODRIGUES; IEMMA, 2015; MASSART et al., 1997).

**TABELA 3** – Planejamento Plackett-Burman (12 ensaios)

Ensaio	I	1	2	3	4	5	6	7	8	9	10	11
1	+	+	+	-	+	+	+	-	-	-	+	-
2	+	+	-	+	+	+	-	-	-	+	-	+
3	+	-	+	+	+	-	-	-	+	-	+	+
4	+	+	+	+	-	-	-	+	-	+	+	-
5	+	+	+	-	-	-	+	-	+	+	-	+
6	+	+	-	-	-	+	-	+	+	-	+	+
7	+	-	-	-	+	-	+	+	-	+	+	+
8	+	-	-	+	-	+	+	-	+	+	+	-
9	+	-	+	-	+	+	-	+	+	+	-	-
10	+	+	-	+	+	-	+	+	+	-	-	-
11	+	-	+	+	-	+	+	+	-	-	-	-
12	+	-	-	-	-	-	-	-	-	-	-	-

Fonte: Elaboração Própria

Os planejamentos de Plackett-Burman abrangem o estudo de 5,7,11,15,19, 23, etc, fatores e se o número de fatores a serem estudados for menor do que os planejamentos disponíveis neste caso, os últimos fatores, ou colunas finais, são considerados inertes. Por exemplo, quando se tem 10 fatores a serem estudados, usando-se um planejamento experimental para 15 fatores, temos que os últimos fatores que sobram seriam fatores sem qualquer efeito no experimento, podendo ser úteis para estimativa do erro experimental. Assim, as colunas do planejamento que não são usadas no estudo devem ser de ao menos três, e representam fatores inertes ou sem efeito, resultando em graus de liberdade para que seja estimado o erro dos valores obtidos para os efeitos (BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003; DEMING; MORGAN, 1988; MASSART et al., 1997).

A classe de planejamento Plackett-Burman é um exemplo de planejamento não regular, ou seja, não pode ter todos os efeitos estimados de forma independentes, desvantagem que é comum aos planejamentos fatoriais fracionados. Neste caso, os efeitos principais são confundidos com os efeitos das interações e esse confundimento ocorre de forma complexa, o que torna difícil as escolhas dos ensaios adicionais necessários para desconfundir os efeitos, embora os efeitos das interações de ordem maior sejam geralmente de importância secundária (BARROS NETO; SCARMINIO; BRUNS, 2001; DEMING; MORGAN, 1988; MONTGOMERY, 2020).

### 3.3.4 Planejamento de Composto Central

Alguns planejamentos experimentais são muito úteis na entrega de informações mais simplificadas sobre o sistema, como no caso exploratório e na obtenção dos fatores importantes e seus níveis de influência na resposta. Porém, eles não se atêm ao fornecimento de informações sobre erro experimental, termos quadráticos ou detalhes de interações. Assim, quando se propõe a obter um modelo, para maior manejo das condições visando predição de resposta e otimização de um sistema, muitos pesquisadores recorrem ao planejamento composto central (DCC).

O planejamento composto central é um tipo de planejamento de experimentos que permite obter informações mais detalhadas a respeito de um sistema, com ajuste de um modelo de segunda ordem útil para que se encontre as condições ideais dos fatores relevantes, comumente para fins de otimização. Tal planejamento é constituído de três partes sequenciais, a primeira é referente ao planejamento fatorial completo ou fracionado com fatores ( $k$ ) codificado em dois níveis (+ e -) e com  $2^k$  ensaios, uma outra parte axial constituída por pontos em todas as coordenadas nulas e em coordenada  $\alpha$  ( $+\alpha$  ou  $-\alpha$ ) com  $2^k$  ensaios, e a última parte é composta por  $n_c$  ensaios realizados no ponto central com valores iguais a zero (mínimo de 3 ensaios).

A construção em partes permite que sejam alocados ensaios de acordo com a necessidade do experimento, podendo avaliar curvatura caso seja importante, ou deixa-la para um outro momento e fazer uso apenas da parte cúbica do experimento, com a consideração dos pontos centrais. Um exemplo da matriz deste tipo de planejamento, constituída pelas três partes, pode ser vista na Tabela 4 para o caso de três fatores (BARROS NETO; SCARMINIO; BRUNS, 2001; BRERETON, 2003; BROWN, 2020).

**TABELA 4** – Planejamento Composto Central para 3 Fatores

<b>Ensaio</b>	<b>Fator 1</b>	<b>Fator 2</b>	<b>Fator 3</b>
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+
9	<b>-<math>\alpha</math></b>	<b>0</b>	<b>0</b>
10	<b><math>\alpha</math></b>	<b>0</b>	<b>0</b>
11	<b>0</b>	<b>-<math>\alpha</math></b>	<b>0</b>
12	<b>0</b>	<b><math>\alpha</math></b>	<b>0</b>
13	<b>0</b>	<b>0</b>	<b>-<math>\alpha</math></b>
14	<b>0</b>	<b>0</b>	<b><math>\alpha</math></b>
15	<b>0</b>	<b>0</b>	<b>0</b>
16	<b>0</b>	<b>0</b>	<b>0</b>
17	<b>0</b>	<b>0</b>	<b>0</b>

Fonte: Elaboração Própria

No caso exemplo de três fatores com dois níveis, tem-se um planejamento fatorial completo constituído de 8 ensaios, que é capaz de estimar os termos lineares e de interação, tendo todas as combinações possíveis dos níveis dos três fatores. A adição de outros ensaios contendo pontos axiais ( $\alpha$ ) e níveis nulos (ponto central) para todos os fatores neste conjunto, possibilita a obtenção dos termos quadráticos e informações de curvatura, com valor de  $\alpha$  dependendo do número de fatores e das propriedades desejadas para o planejamento. Por fim, a adição de no mínimo três replicatas no ponto central possibilita a estimativa do erro experimental (BRERETON, 2003; BROWN, 2020).

Apesar de na sua construção ser usualmente usado o fatorial completo para um pequeno número de fatores, quando se expande o número de variáveis estudadas, o aumento exponencial dos ensaios necessários prejudica essa escolha. Portanto, não se exclui a possibilidade de usar um esquema fatorial fracionado na primeira parte da matriz a ser construída, embora tenha-se que atentar a utilização de uma resolução

V, que permita o uso com confundimento aceitável (BARROS NETO; SCARMINIO; BRUNS, 2001).

O valor de  $\alpha$ , novidade para este tipo de planejamento, é um parâmetro deste tipo de planejamento que deve ser especificado e tem valor localizado entre 1 e  $\sqrt{k}$ , sendo k o número de fatores estudados. A definição do valor de  $\alpha$  se baseia na distância dos pontos axiais com relação aos pontos centrais e se relaciona com a forma e tamanho do domínio do planejamento composto central, sendo esférico, quando  $\alpha = \sqrt{k}$ , ou cúbico, quando  $\alpha = 1$ , com impacto na rotabilidade do planejamento.

A rotabilidade é uma propriedade que diz respeito a consistência e estabilidade da variância da resposta predita pelo modelo gerado em um dado ponto, devendo ser a mesma para qualquer ponto x que esteja a uma mesma distância do centro, ou seja, que se possa estimar com precisão igual em todas as direções, como em uma esfera ao redor do centro. Essa propriedade é a base de um planejamento de superfície de resposta, que permite gerar modelo e otimizar sistemas, pois permite que se tenha igual precisão de estimativa em todas as direções, dado que a localização do ponto ótimo do sistema e seus valores de coeficientes são desconhecidos a princípio.

A atribuição da propriedade rotacional ao planejamento composto central é conseguida usualmente ao definir  $\alpha = (2n)^{1/4}$ , sendo n o número de variáveis independentes, resultando em um planejamento composto central rotacional (DCCR) (BROWN, 2020; WU; HAMADA, 2009; MONTGOMERY, 2020).

Com a realização de um planejamento composto central se pode calcular valores de coeficientes empregados em um modelo fazendo uso de regressão, obter a significância dos coeficientes aplicando ANOVA, teste-F, teste-t ou outras operações estatísticas para se extrair informações detalhadas acerca do sistema. Além disso, pode-se com um modelo matemático gerar uma superfície de resposta com o propósito de se otimizar o sistema e obter uma visualização gráfica do ponto ótimo, das coordenadas dos fatores plotados nas quais a resposta apresenta valor interessante e do comportamento geral da resposta de acordo com mudança contínua dos valores dos fatores dentro dos limites empregados (BRERETON, 2003).

### 3.3.5 Planejamento Box-Behnken

Este tipo de planejamento foi proposto por Box e Behnken em 1960 e tem como característica um planejamento de segunda ordem para gerar superfície de resposta, assim como o composto central, embora apresente as desvantagens de usar apenas três níveis dos fatores (pobre experimentalmente) e sempre resulta em quantidade de ensaios maior do que o DCCR.

O planejamento combina uma estrutura fatorial  $2^k$  com um planejamento de bloco incompleto. O resultado é bastante econômico e eficiente por gerar um número enxuto de ensaios, ter propriedade rotacional e constituir uma alternativa ao planejamento composto central (BARROS NETO; SCARMINIO; BRUNS, 2001; WU; HAMADA, 2009; MONTGOMERY, 2020).

**Tabela 5** – Esquema de Construção Box-Behnken

Blocos	Fator 1	Fator 2	Fator 3
1	$\beta$	$\beta$	x
2	$\beta$	x	$\beta$
3	x	$\beta$	$\beta$

Fonte: Elaboração Própria

A construção da matriz do planejamento pode ser exemplificada pelo esquema da Tabela 5, que considera três blocos na composição dos ensaios para três fatores estudados. Cada símbolo  $\beta$ , em cada um dos blocos, são substituídos pela coluna codificada de dois níveis do fator correspondente, extraída da matriz de um planejamento  $2^2$ , e onde não há x é preenchido com uma coluna de zeros. O procedimento é repetido para cada bloco, considerando os fatores que participam do bloco, e ao final são adicionados no mínimo três ensaios no ponto central, resultando em 15 ensaios (WU; HAMADA, 2009).

O planejamento Box-Behnken, exemplificado pela matriz na Tabela 6, apresenta a característica de exigir apenas três níveis para cada fator, sendo

econômico nesse sentido. Além disso, todos os pontos tem distância iguais com relação ao centro  $\sqrt{2}$ , definindo assim uma região esférica com variância da resposta igual em todo o domínio, atendendo ao critério de rotatividade. No entanto, para situações nas quais níveis extremos dos fatores são necessários, ocupando posições nos vértices do cubo, o planejamento não é adequado, pois tem domínio esférico que não abrange os vértices. (WU; HAMADA, 2009; BARROS NETO; SCARMINIO; BRUNS, 2001).

**TABELA 6** – Planejamento Box-Behnken para três fatores

Ensaio	Fator 1	Fator 2	Fator 3
1	-	-	0
2	-	+	0
3	+	-	0
4	+	+	0
5	-	0	-
6	-	0	+
7	+	0	-
8	+	0	+
9	0	-	-
10	0	-	+
11	0	+	-
12	0	+	+
13	0	0	0
14	0	0	0
15	0	0	0

Fonte: Elaboração Própria

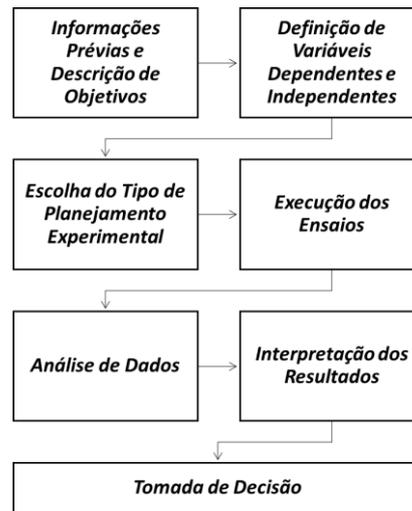
### 3.4 ETAPAS DE APLICAÇÃO

A necessidade de experimentação surge a partir do vislumbre de um problema que precisa ser tratado ou de um sistema que precisa ser explorado a fim de gerar conhecimento. Nas atividades de rotina em laboratórios químicos comumente se tem de encontrar soluções para problemas que surgem como parte da realização de ensaios, aprofundamento do entendimento acerca de um fenômeno, processo, análise, síntese, etc. Em qualquer que seja o problema experimental, ele é geralmente composto de variáveis independentes e variáveis dependentes (respostas), e o posicionamento inicial é fazer uso de conhecimento que permita entender o que ocorre e quais são essas variáveis envolvidas a fim de se otimizar a resposta do sistema.

Para isso, o uso do planejamento experimental se mostra uma abordagem muito útil, na maximização da qualidade da informação obtida acerca de um sistema, com a execução de experimentos de forma menos dispendiosa possível. Sua abordagem permite a construção de um delineamento racionalizado em etapas e que faz uso de ferramentas estatísticas na concepção de uma estratégia experimental para se obter conhecimento sobre um sistema como recurso na tomada de decisão.

Com isso, a depender do objetivo do estudo, pode-se seguir caminhos distintos. Na tratativa que visa obter o modelo de uma resposta em função de coeficientes que representam os fatores mais importantes, pode-se assumir inicialmente um modelo linear para o sistema. Nesse sentido, o uso de um planejamento fatorial completo ou fracionado mostra-se válido, pois é capaz de oferecer um modelo linear. No entanto, caso se verifique que o modelo linear gerado a partir de um planejamento fatorial completo não satisfaz os objetivos, deve-se recorrer a um planejamento capaz de fornecer os termos quadráticos.

Ainda, se o objetivo vai além da entrega do modelo, e passa a ser uma otimização do sistema, os procedimentos são então mais completos e exigem adição de etapas adicionais. Em todos os casos, a aplicação do planejamento experimental segue uma estrutura sequencial como apresentada na Figura 1, que apresenta as etapas de definição dos objetivos, seleção de variáveis, seleção do tipo de planejamento experimental, execução dos ensaios, análise de dados e interpretação dos resultados para tomada de decisão.

**FIGURA 1** – Etapas de aplicação do planejamento experimental

Fonte: *Elaboração Própria*

Para que se obtenha um bom planejamento experimental, deve-se se atentar a uma série de informações prévias acerca do estudo a ser realizado, definindo o escopo sobre o qual se construirá o caminho e etapas do planejamento. Além disso, é necessário que, durante sua construção, se tenha em mente qual o tipo de medida necessária e de que forma é possível obtê-la com maior rigor possível (precisão e exatidão), levando em conta as limitações técnicas, interferências e eventuais problemas que afetam e compõem as condições nas quais os ensaios serão realizados.

Tendo sido as condições prévias pontuadas, pode-se prosseguir com a seleção das variáveis importantes, o domínio experimental e o planejamento mais adequado a fim de prover um uma aproximação matemática do sistema estudado. No final, os dados gerados com auxílio de ferramentas estatísticas são interpretados a fim de se obter informação útil a uma abordagem prática, visando gerar conclusões sobre uma hipótese, obter um modelo, visualizar os efeitos de variáveis ou otimizar o sistema.

O uso do planejamento experimental se mostra uma abordagem muito útil no entendimento de problemas experimentais, mas antes é necessário que se tenha claro o objetivo, para que se possa verificar quais os fatores que apresentam maior potencial de afetar a resposta. Longe de qualquer obviedade, está é uma etapa essencial que não deve ser subestimada.

De início, deve-se definir os problemas do sistema a ser estudado, os objetivos e hipóteses, responsáveis pelas tarefas e reunir os conhecimentos prévios que se tem sobre o problema. Nessa etapa, todas as pessoas envolvidas devem participar e contribuir com suas experiências e conhecimentos que podem ser úteis no planejamento dos experimentos e na definição das hipóteses, servindo como “chute inicial” de como o problema ocorre, o que está envolvido e como as variáveis interagem de forma a ocasionar a resposta observada (BENEDETTI; CAPONIGRO; ARDINI, 2022).

A hipótese pode ser desenvolvida de três formas, por analogia, dedução e intuição. A analogia corresponde a observação de semelhanças entre casos anteriores e o fenômeno presente. Via indução, tem-se a prescrição de conhecimentos na tentativa de elaborar uma hipótese para verificar se a observação do fenômeno pode ser generalizada para além do caso pontual. E a forma dedutiva consiste no uso do entendimento sólido das propriedades do fenômeno, quando já se conhece suas características e ocorrência.

Ainda nesta etapa, tem-se a coleta de informações técnicas do sistema a ser estudado, como os fatores de controle, ruídos, níveis de operação, condições limitantes, ajustes necessários e as possíveis variáveis de resposta. A avaliação de tempo disponível, custos e recursos disponíveis são também consideradas e definem os máximos da estratégia experimental, assim como o planejamento mais adequado e suas replicatas (BENEDETTI; CAPONIGRO; ARDINI, 2022).

Tendo definidos os objetivos e hipóteses, a etapa subsequente seria a de avaliar quais são os fatores capazes de exercer influência significativa na resposta de interesse. Inicialmente, no princípio da concepção do estudo, os fatores de real influência no comportamento do sistema são incertos, e por isso é necessário que se possa definir quais dos fatores são significantes, estatisticamente, no domínio experimental. Nesta fase, o experimentador ou equipe devem selecionar as variáveis dependentes e independentes, sendo as dependentes aquelas que são medidas durante o experimento (respostas), enquanto as independentes são aquelas determinadas antes da realização dos ensaios (fatores).

A seleção dos fatores deve ocorrer com base no grau de influência sobre as respostas escolhidas, capacidade de controle e capacidade de se medir com precisão.

Define-se também nesta etapa, as faixas dos níveis, escala e métodos de medição, levando em consideração os conhecimentos sobre o fenômeno, teóricos e práticos, reunidos na primeira etapa (BENEDETTI; CAPONIGRO; ARDINI, 2022; BROWN, 2020).

O conhecimento prévio, em alguns casos, pode delimitar o número de fatores a serem considerados no estudo, sendo possível neste momento, proceder com planejamentos fatoriais completos, cuja viabilidade pode depender do número de fatores. Quando um número grande de variáveis é jugado relevantes para a resposta, um grande número de ensaios pode ser necessário dentro do planejamento fatorial completo, inviabilizando sua realização. No entanto, é sabido que em geral nem todos os fatores listados são relevantes para a resposta, e, portanto, é necessário que seja realizada a aplicação de um planejamento que permita realizar a seleção de variáveis (screening) (BENEDETTI; CAPONIGRO; ARDINI, 2022; BROWN, 2020).

Planejamentos que visam a seleção de variáveis constituem uma etapa anterior a aplicação de planejamento que visam otimização de sistemas, ajuste de modelo ou geração de superfície de resposta, uma vez que nem todas as variáveis são estatisticamente importantes. Tais planejamentos resultam na obtenção dos efeitos lineares de um número extenso de fatores com a realização de alguns poucos ensaios, como o caso dos planejamentos Plackett-Burman e Fatoriais Fracionados, tipicamente usados em screening, que podem reduzir o número de ensaios e apresentar os efeitos dos fatores para uma seleção prévia (BENEDETTI; CAPONIGRO; ARDINI, 2022; BROWN, 2020).

Com a execução adequada da definição dos objetivos e seleção de variáveis, prossegue-se com a seleção do planejamento e matriz adequados aos objetivos dos experimentos e ao caráter das variáveis e processos de medição. Nesta etapa, considera-se o número de fatores importantes previamente selecionados, o número de níveis destes fatores, os fatores não controláveis do sistema, o número de replicatas, recursos disponíveis, interações possíveis e limitações experimentais. Existem diversos tipos de planejamentos que se adequam a determinadas estratégias, e devem ser selecionados de forma a atender os objetivos da realização dos experimentos, seja de otimização, seleção de variáveis, visualização dos efeitos de fatores ou geração de curva de resposta. Alguns planejamentos podem ser inviáveis, então conhecer bem os tipos e estratégias disponíveis auxiliam na escolha adequada,

a fim de evitar retrabalho, perda de informação e não completude dos ensaios devido à falta de recursos (BENEDETTI; CAPONIGRO; ARDINI, 2022).

O planejamento experimental pode resultar em um número de ensaios bastante elevado, dependendo do número de fatores considerados e seus níveis, assim, a observação do número de ensaios necessários é importante para que se possa considerar uma redução desse número já na escolha do planejamento, tornando-o viável. As replicatas no ponto central são outro ponto de interesse, tendo a exigência de no mínimo três ensaios para uma boa estimativa do erro experimental, embora números maiores possam ser considerados em caso de disponibilidade dentro da quantidade total de ensaios praticáveis. Outro ponto que não deve ser deixado de lado é a propriedade de rotabilidade, necessária para quando se deseja otimizar um sistema, definir um modelo de estimativa de resposta ou superfície de resposta (BENEDETTI; CAPONIGRO; ARDINI, 2022).

Com o experimento já estruturado, com matriz definida, pode-se então proceder com as etapas de execução dos ensaios. A equipe ou experimentador deve realizar os ensaios como planejado, exercendo grande controle das atividades e metodologias aplicadas. Deve-se também assegurar o máximo rigor ao plano, manutenção das condições experimentais, não esgotamento dos recursos e precisão na coleta de resultados e informações de cada ensaio. Experimentos realizados com falhas ou sem o devido controle podem invalidar completamente os resultados e condução do experimento.

Uma vez que os ensaios foram realizados e os dados coletados adequadamente, realiza-se a análise estatística dos dados, adequada ao tipo de planejamento escolhido. Nesta etapa é comum a utilização de softwares matemáticos e estatísticos que fornecem ferramentas apropriadas e suporte ao planejamento e análise de experimentos, com a geração do modelo, descrição de comportamento de variáveis, estimativa de efeitos, interações, modelagem e geração de gráficos. A análise estatística efetuada geralmente é feita via análise de regressão, análise de resíduos, análise de variância (ANOVA), validação cruzada ou via validação externa do modelo com um conjunto de testes. Nesta etapa pode-se ter noção do quão bom é o modelo gerado e sua aplicabilidade, e gera-se todo o recurso necessário para a formulação de conclusões acerca das hipóteses, com alcance dos objetivos

inicialmente definidos e tomada de decisão (BENEDETTI; CAPONIGRO; ARDINI, 2022; BROWN, 2020).

Finalmente, observa-se o que foi obtido a partir dos resultados das análises performadas e extrair dos achados, as conclusões práticas, com decisões de melhoria e resoluções que permitam otimizar um sistema ou solucionar um problema. Além disso, elaboram-recomendações para futuros experimentos e, com uma visão crítica, discute-se se os resultados fazem sentido e satisfazem as questões levantadas inicialmente.

### 3.5 SOFTWARES

Na aplicação do planejamento de experimentos, uma série de pacotes de softwares estatísticos são empregados a fim de facilitar e tornar mais dinâmico e intuitivo as etapas de estruturação da matriz do planejamento, bem como as análises estatísticas e computação dos dados coletados. Não há dúvida de que a utilização de softwares é um marco em estudos do tipo, uma vez que, dependendo do número de variáveis e da extensão dos ensaios e medidas efetuadas, os cálculos podem se tornar trabalhosos e morosos de serem realizados manualmente, ou via planilha eletrônica, ainda que sejam possíveis tais usos.

Dentre os softwares mais utilizados para aplicações em planejamento experimental pode-se citar Minitab, STATISTICA, R, JMP, Design Expert, Protimiza, Matlab, etc. Cada um deles apresenta características próprias, podendo significar facilidade ou dificuldade a depender do usuário e do caso em estudo. Alguns softwares possuem alto nível de automação e exigem operações simplificadas que não explicitam ao usuário qual operação de fato está sendo seguida na efetuação do cálculo, já outros são capazes de oferecer liberdade ampla ao usuário que deve ser responsável por cada etapa da operação de cálculo e análise dos dados, o que pode ocasionar erros, quando o usuário não está bem familiarizado com o software ou com as bases matemáticas e estatísticas do tratamento de dados de planejamento experimental.

Além dos softwares especializados em cálculos e análises estatísticas, existem softwares matemáticos mais amplos, que apesar de possuírem pacotes de

funcionalidades para estatística ou mesmo planejamento experimental, exigem a construção própria dos algoritmos e banco de dados necessários para performar os cálculos e análises. Tais softwares, como Excel e Matlab, podem representar uma enorme dificuldade em razão do tempo necessário para preparar as bases de cálculo como atenção para que erros de programação e fórmula não ocorram.

No geral, cada experimentador, baseado em sua familiaridade, conhecimento e caso tratado pode optar pelo software que for mais apropriado em termos de acesso, disponibilidade, custo, tempo, e minimização de possíveis erros. No entanto, o melhor seria optar por um software que ofereça um balanço entre automação e personalização, que possa oferecer recursos facilitadores ao usuário, sem comprometer as liberdades necessárias para personalização e adequação aos diversos casos de estudo possíveis. Assim, o usuário, ou experimentador, pode realizar operações mais usuais de forma segura e facilitada, além de ter disponível possibilidades de explorar os dados com operações mais complexas e de forma aprofundada.

## 4 METODOLOGIA

A aplicação do planejamento experimental, apesar de ser difundida em algumas áreas da ciência, ainda possui particularidades restritas do campo prático que não são descritas na literatura. Tais particularidades representam tópicos sensíveis que dizem respeito ao uso adequado da ferramenta a fim de obter melhores resultados, com qualidade e validade estatística.

A fim de acessar os tópicos de interesse e descrevê-los junto a propostas de abordagens, realizaram-se pesquisas bibliográficas com análise preliminar dos materiais de interesse sobre os temas em diretórios de pesquisa Google Scholar, PubMed, Scielo e ScienceDirect, além de livros da área de Planejamento Experimental.

A partir do material selecionado, realizaram-se análises e estudos das referências a fim de obter informações sobre a descrição, abordagem e casos exemplo referentes aos tópicos apresentados.

Por último, realizaram-se discussões com base no conhecimento reunido com o estudo do material selecionado, a fim de apresentar descrição, abordagem sugerida e, se possível, caso exemplo referente a cada um dos tópicos apresentados, dentro de boas práticas de aplicação do planejamento experimental. A composição dos textos de discussão se baseou então no conteúdo dos materiais selecionados, lidos e estudados, bem como em experiências práticas descritas ou não em literatura da área.

## 5 RESULTADOS E DISCUSSÕES

### 5.1 CONHECENDO O SISTEMA

Muitas vezes um estudo, ainda em seu início pode caminhar em direção ao fracasso devido ao não planejamento adequado dos experimentos, a fim de se obter respostas as questões geradas sobre um certo problema de um sistema. O pré-planejamento cuidadoso é parte essencial para um planejamento experimental bem-sucedido, mas essa etapa acaba sendo subestimada, ainda que tenha o caráter essencial de servir de embasamento do estudo e que fornecerá as possibilidades e diretrizes para as etapas subsequentes de seleção de variáveis e escolha do planejamento experimental.

Antes de se iniciar o planejamento experimental, é necessário que se realize uma etapa crítica de definição do que se pretende estudar, das respostas para os problemas levantados, com definição dos limites do trabalho que será feito, e formulação das hipóteses e objetivos da realização dos experimentos. Nessa etapa, é ideal que se escolha o que será feito e também o que não será feito, além de reunir conhecimento sobre o sistema e o problema a fim de se obter alternativas para uma solução realista com maior chance de sucesso, com identificação das fontes de ruído e variabilidade nas condições experimentais, os recursos adicionais necessários e as limitações técnicas.

Os problemas e perguntas levantados sobre o sistema a ser estudado, parte da observação inicial, observação essa que deverá dar origem a uma estrutura lógica relacionada aos fenômenos envolvidos, a fim de indicar quais as possíveis variáveis que afetam o sistema e quais as respostas a serem medidas são importantes para o caso. Assim, a visualização do que será feito se torna mais fácil, com as suposições dos efeitos e quais são os agentes envolvidos no impacto nas respostas importantes no caso.

Muitas dificuldades podem ocorrer nesta fase inicial do planejamento experimental, especialmente relacionado a visualização do problema e na concepção da estrutura lógica sobre os objetivos principal e sobre o que é esperado como resposta. Uma forma de se visualizar o problema, por exemplo, é fazendo uso do

diagrama de causa e efeito. O diagrama desenvolvido por Kaoru Ishikawa como ferramenta de controle da qualidade na década de 60, pode auxiliar no resumo e exercício mental quanto a suposições das relações de causa e efeito de um sistema.

Outro ponto é que nessa abordagem inicial do planejamento experimental, não se pode considerar todos os aspectos de um problema e não se pode controlar todos os fatores e comportamentos que podem afetar o sistema. Caso as tentativas de alcançar o maior número de variáveis e resposta seja feita como proposta inicial, pode-se deparar com uma quantidade enorme de ensaios a depender da escolha do planejamento e conseqüentemente com o despendimento de horas de trabalho e recursos, o que inviabiliza completamente o estudo. Para o caso de muitas variáveis existem opções de planejamento que visam exatamente verificar quais delas tem efeito significativo, podendo reduzir o número de variáveis, mas é interessante que o planejamento seja enxuto e se tenha noção do escopo com ideia das questões que podem ser respondidas por meio de um estudo ou se mais de um estudo é necessário.

O desenvolvimento da estrutura inicial do planejamento experimental é capaz de determinar o potencial de falha, caso os objetivos e as hipóteses forem mal formulados, os dados iniciais não forem levantados e caso não sejam consideradas informações de histórico, literatura e conhecimentos do experimentador. Essa fase é suportada por uma análise crítica de dados reunidos com informações sobre o sistema e o problema. Essas informações são obtidas por pesquisa bibliográfica aprofundada, brainstorming com grupos e pessoas que carregam algum conhecimento acerca do sistema em estudo, diálogo com especialistas na área, experiências particulares do pesquisador, resultados de estudos anteriores que foram insuficientes e leitura dos tópicos que possam ser relevantes para avaliar a viabilidade e praticabilidade do estudo.

A execução de experimentos sem que se tenha clareza do das informações prévias sobre o sistema, de forma vaga e sem propósito claro, muitas vezes resulta na percepção tardia de dificuldades operacionais do experimento, incapacidade de gerar respostas satisfatórias, impossibilidade de gerar conhecimento sobre o sistema, ou mesmo encerramento abrupto do experimento sem uma conclusão. Isso ocorre pois são deixadas de lado informações que serão importantes em outras etapas, como a disponibilidade de recursos, tempo, equipamentos e faixas de estudo que não podem ser medidas ou são inseguras.

Nesse sentido, pode-se pontuar, por exemplo, a escolha dos níveis dos fatores, pois depende das considerações de custo e operacionais, tendo em vista de pode inviabilizar a execução do experimento quando se trata de um “material” caro, inviabilizar a medição quando foge dos limites mínimos ou máximos de detecção do aparato de medição, ou quando a combinação de níveis com outros fatores ocasiona uma situação de risco. Deve-se, portanto, se atentar aos níveis que melhor se adequam a toda a condição experimental ainda nas definições prévia do estudo.

Além dos fatores e seus níveis, deve-se também escolher, ainda nesta etapa, as respostas que são relevantes para a hipótese, a qual o efeito de interesse se relaciona. Pode-se ter mais de uma variável de resposta envolvida em um experimento e para que se possa selecionar a mais relevante, é necessário que se entenda as causas e seus efeitos, ou seja, que não se abra mão do estudo prévio e das discussões, pois o conhecimento da variável de resposta permite que sejam verificados se os mecanismos de medição que permite registrar de forma adequada a resposta, como equipamentos apropriados aos níveis, escala e condições experimentais estão disponíveis.

No geral a recomendação acerca do pré-planejamento é de que se colete todas as informações necessárias sobre o sistema, que se discuta com o pessoal envolvido e que considere a princípio todas as variáveis importantes, para que não se descubra tardiamente a falta de informações importantes ou inviabilize o experimento por razões evitáveis. E, por fim, com base nessas recomendações e com conhecimento suficiente acerca do sistema, efeitos, variáveis e escolha do planejamento experimental adequado, a experimentação então tem caminho bem estruturado e é capaz de fornecer respostas satisfatórias.

## 5.2 TEMPO COMO VARIÁVEL DE ENTRADA

Tempo é usualmente descrito como variável independente, e é intuitivo, em alguns casos, que se use o tempo como fator de entrada no planejamento experimental, uma vez que de fato pode ser um parâmetro determinante para que se tenha a ocorrência de um certo fenômeno, formação de um produto em uma reação, alteração físico-química, ação enzimática, etc.

Especialmente em alguns casos, como quando se objetiva estudar o tempo necessário para o estado de equilíbrio, o uso de um catalisador, estudo de atividade enzimática máxima ou rendimento reacional, existe o questionamento acerca do tempo mínimo para que se obtenha os efeitos desejados. No entanto, a inserção do tempo como variável independente não é sempre recomendada e o entendimento do processo e dos métodos de planejamento e análise utilizados, aliado ao bom senso do experimentador, é o que permite a melhor decisão quanto a inserção ou não do tempo como fator de controle no planejamento experimental.

Apesar de a maioria dos softwares utilizados em planejamento experimental permitirem a inclusão do tempo como variável de entrada em qualquer planejamento feito, ou da impulsividade em se incluir o tempo, devido a obviedade de sua importância para o processo, não se pode decidir sobre isso sem que se tenha o pleno conhecimento do processo a ser estudado e das considerações a serem feitas acerca do tempo como fator de controle.

A decisão do uso do tempo como variável de entrada depende exclusivamente da experiência dos experimentadores, da condição experimental e do processo, assim como do bom senso. Cada caso deve ser analisado a fim de tomar a melhor decisão, muitas vezes para que não se use o tempo como variável de entrada inadvertidamente, ou para que não se perca informações relevantes para o entendimento do processo e dos efeitos dos parâmetros.

Na indústria, pode-se ter dois tipos principais de processos: batelada e contínuo, sendo processo transiente e estacionário, respectivamente. Observando-se o caso de um reator químico em um processo em batelada, temos a situação na qual os reagentes são inseridos no reator em quantidades determinadas de acordo com o que se espera obter de produto, com base no rendimento da reação. E após um certo tempo de reação, tem-se os produtos formados e o estado de equilíbrio do meio. Já no processo contínuo se observa um fluxo contínuo de reagentes para dentro do reator, enquanto se retira o produto formado continuamente.

O processo em batelada pode ser ilustrado pelas curvas de decaimento dos reagentes em função do tempo, ao passo que a curva de produtos aumenta ao longo do tempo de reação, até que se atinja um máximo de produtos formados, limitado pelos reagentes e condições reacionais. Neste exemplo, pode-se observar que temos

um tipo de processo que muda com o tempo, ou seja, trata-se de um processo dinâmico no qual a cinética reacional importante e assume um estado transiente. Para este caso, a tomada de amostras em tempos diferentes resulta em resultados diferentes devido ao estado dinâmico e não se recomenda a utilização do tempo como variável de entrada, mas se avalia a resposta de interesse ao longo do tempo.

No caso de um processo contínuo, estacionário, temos um fator chamado de tempo de residência, que é o tempo no qual uma substância permanece ou passa pelo sistema. Observando-se a mudança do fluxo de reagentes e produtos no sistema, tem-se constante a entrada dos reagentes e a saída do produto. Se nada é modificado nesse tipo processo (contínuo), com a obtenção de amostras do sistema ao longo do tempo, os valores da resposta analisada sobre a amostra tendem a permanecer constantes, pois a mesma condição é mantida e tem-se um estado estacionário. Neste caso a variável de interesse não varia com o tempo e pode-se inserir o tempo de residência como variável de entrada.

Para um processo em batelada, a recomendação é que se avaliem as respostas de interesse ao longo do tempo, tomando alíquotas para análise a cada intervalo de tempo definido, para cada ensaio. Tais intervalos de tempo podem ser definidos com um pré-estudo da cinética tomando a condição do ponto central, embora não seja uma regra, com intervalos pequenos de tempo, permitindo escolher os melhores tempos a serem estudados nos demais ensaios do planejamento experimental, ou pode-se basear em conhecimento prévio coletado na fase de pré-planejamento experimental (Item 5.1). Para todos os casos em batelada, deve-se finalizar o estudo com a validação na condição de parâmetros ótimos, realizando o estudo cinético.

Podem existir casos nos quais o estado é dinâmico, cinético, mas não se pode coletar amostras a cada tempo em cada ensaio por um impedimento da condição experimental, limitações do equipamento utilizado, custo de análise. Por exemplo, quando se utiliza um sistema altamente pressurizado, no qual a abertura do sistema compromete a continuidade do processo, quando em um processo de ação microbiológica necessita-se de uma condição anaeróbica, ou quando o custo para a realização das análises é muito alto. Nesta situação, recomendam-se a utilização de “amostras de sacrifício” que podem ser analisadas em cada intervalo de tempo de forma destrutiva, ou seja, na mesma condição de ensaio tem-se uma quantidade de

amostras suficientes para serem analisadas a cada intervalo de tempo, em uma metodologia destrutiva, sem comprometer estudo.

Em um estudo realizado por Mohammed et al. (2017), no qual estudou-se os efeitos dos parâmetros de síntese de nanotubos de carbono na morfologia das estruturas formadas, um processo em batelada, os pesquisadores aplicaram um delineamento fatorial completo, variando alguns parâmetros como temperatura, fluxo de argônio, fluxo de acetileno (fonte de carbono) e tempo de reação. O planejamento resultou em 16 ensaios com dois níveis para cada fator. As análises resultaram em dois parâmetros significativos, temperatura e fluxo de argônio, sendo o tempo não significativo.

Neste caso o tempo aparece como variável de entrada e é possível observar que nenhuma informação sobre a cinética pôde ser obtida, embora se tenha observado que para cada temperatura estudada se tem regimes cinéticos diferentes, o que é corroborado pela literatura, e que o tempo é importante para a morfologia das estruturas formadas. Apesar disso, as informações sobre cinética da reação, variação da morfologia no tempo ou mesmo informações sobre o ponto de equilíbrio da reação não puderam ser obtidas, devido a escolha inicial de se inserir o tempo como variável de entrada. Assim, não foi possível avaliar o comportamento de forma mais detalhada ao longo do tempo de formação dos nanotubos de carbono, em cada uma das variações dos demais parâmetros, o que seria interessante especialmente diante da significância da variação de temperatura.

Um outro caso exemplo do uso do tempo como variável de entrada é apresentado por Rodrigues & lemma (2015) no livro Planejamento Experimental e Design de Processos. O caso apresentado trata de uma otimização do processo de alcalinização do nibs de cacau em um planejamento fatorial com 18 ensaios. Neste estudo, considerou-se como variáveis de entrada temperatura (60-120°C), tempo e concentração do agente alcalinizante  $K_2CO_3$  (carbonato de potássio – 0-6%), avaliando-se as respostas de pH, cor ( $L^*a^*b$ ) e sensorial (cor e sabor).

Os resultados apresentaram o tempo como não significativo (30-150min), não afetando negativamente os resultados, devido ao tempo encontrar-se no estado estacionário, onde não exerce efeito no processo. No entanto, é ressaltado que isso não significa que, em temperaturas diferentes ou condições diferentes do que foi

avaliado, o tempo continuaria não influenciando o processo. Assim, é denotado que há um risco associado a inserção do tempo como variável de entrada, podendo a faixa escolhida não abranger valores para que se atinjam todas as condições adequadas do processo ou mesmo que seja excedido e resulte em condições danosas ao processo.

Um exemplo positivo, no qual a variável tempo não é inserida como variável independente, mas no qual há coleta de amostras e análises ao longo do tempo, em períodos pré-definidos, é apresentado por Guzzo da Silva et al., 2015 no estudo da cinética de secagem de pimenta-rosa. No estudo, foram realizados 11 experimentos de um planejamento composto central rotacional, considerando duas variáveis, temperatura do ar e velocidade do ar, com cinco níveis. Os ensaios foram analisados nos períodos de 4 e 6 h para obtenção do teor de umidade, com investigação da taxa de secagem e da cinética em cada ensaio.

Os autores apresentam como variável significativa a temperatura do ar para os dois períodos de tempo estudados. As informações cinéticas, obtidas com escolha de não inserir tempo como variável de entrada, forneceram achados importantes do comportamento dos parâmetros ao longo do tempo. Os valores de teor de umidade diminuem conforme ocorre o aumento da taxa de secagem, a umidade de equilíbrio é obtida pela cinética e apresenta redução com o aumento da temperatura e, a cinética permite dizer que a secagem ocorre principalmente por remoção de umidade interna do sólido (GUZZO DA SILVA; FRATTINI FILETI; PEREIRA TARANTO, 2015).

Desse modo, quando se insere o tempo como variável de entrada em processos em batelada, não se obtém informações e entendimento sobre a cinética do processo e, além disso, assume-se um risco inerente de seu uso variável de entrada em processos em batelada, que impede o estudo ao longo do tempo e pode resultar no “mascaramento” dos efeitos de outras variáveis importantes para o processo, podendo chegar à conclusão de que apenas o tempo é significativo ou mesmo promovendo conclusões equivocadas devido a faixa de tempo escolhida.

A curva cinética, que é perdida com a inserção do tempo, é parte do processo a ser seguido e é capaz de fornecer informações importantes do processo como o máximo de uma atividade enzimática, o tempo necessário para se atingir o estado de equilíbrio, rendimento máximo de uma reação, etc. A perda de informação e o

mascamamento de efeitos são muito danosos ao estudo realizado e assim, o uso do tempo como variável de entrada em processos em batelada não se justifica.

### 5.3 CONSIDERAÇÃO DE FATORES QUALITATIVOS

Fatores qualitativos são aqueles não associados a níveis mensuráveis, ou seja, que contém como níveis variáveis qualitativas nominais. Pode-se citar como exemplos de fatores qualitativos, normas, condições climáticas, gênero, etc (MONTGOMERY, 2020).

Variáveis qualitativas, ou categóricas, são aquelas que descrevem propriedades, atributos e qualidades e não são descritas por números. Podem ser classificadas como ordinais (intensidade de cor, preferência de consumidor, avaliação de desempenho), quando seus valores podem ser ordenados e sequenciados, ou nominais (nome comercial/marcas, nome de produtos químicos, etnia de uma pessoa, nome de uma espécie de animal), quando não podem ser colocadas em ordem de magnitude (RODRIGUES; IEMMA, 2015)

Em determinadas situações a realização de experimentos pode incluir como elementos importantes capazes de afetar uma resposta, variáveis qualitativas. Tais variáveis muitas vezes caracterizam tipos de operações unitárias de um processo com uma mesma finalidade (peneiramento, corte, agitação, etc.) ou podem caracterizar tipos de uma matéria-prima, enzimas, catalisadores ou microrganismos.

O uso dos fatores com níveis qualitativos em planejamento experimental é permitido, no entanto, deve-se ter atenção quanto ao uso adequado, especialmente quando representam tipos de microrganismos, enzimas, catalisadores, etc, elementos com a mesma função na promoção da resposta avaliada. Quando se considera uma variável qualitativa nominal que tipifica algo como variável de entrada, seus níveis representam os tipos distintos de um dado item, substância, ser vivo, etc. que podem apresentar comportamento, propriedades e efeitos sobre a resposta distintas um do outro, razão para serem estudados, e tal comportamento depende das condições do meio, que por sua vez, podem ser definidas pelas demais variáveis de entrada.

Nas condições experimentais definidas para cada ensaio, de acordo com as variações das demais variáveis, os níveis de um dado fator qualitativo (diferentes tipos) pode estar sujeito a condições inadequadas e ter sua ação comprometida. Assim, há a dificuldade em distinguir detalhes sobre a atuação da variável qualitativa sobre a resposta estudada.

Além disso, não se tem, mais uma vez, o conhecimento individualizado acerca do comportamento e desempenho de determinada enzima, catalisador, microrganismo, etc. na resposta escolhida ao longo do tempo. Por isso, a depender do caso, a recomendação é de que se estude as diferentes variáveis qualitativas em planejamentos experimentais separados, cada um com suas análises separadas. Isso possibilita a geração de informações específicas e detalhadas acerca de cada um nas diversas condições determinadas nos ensaios, pela variação dos fatores.

Uma abordagem mais adequada no estudo individualizado seria a de selecionar os pontos centrais e nestes pontos realizar um estudo cinético completo para cada uma das variáveis qualitativas (tipos de itens), a fim de selecionar o tipo que apresenta resultados mais promissores e em seguida realizar um planejamento experimental completo com análises estatísticas, visando otimização da resposta, por exemplo.

Como exemplo, tem-se um caso apresentado por Rodrigues & lemma (2015), que buscou avaliar os efeitos de oito variáveis sobre o crescimento celular de duas cepas de microrganismos do gênero *Rhodotorula*, que apresentam potencial na geração de carotenoides. No estudo, os pesquisadores realizaram a avaliação de cada cepa em planejamentos experimentais separados.

No estudo, foi utilizado inicialmente um planejamento fatorial fracionado  $2^{8-4}$  com 16 ensaios e 4 repetições no ponto central. Na estratégia apresentada, são inseridas 8 variáveis que impactam o crescimento destas cepas de microrganismos, e foram realizados os 20 ensaios para cada uma das cepas separadamente, ou seja, realizaram-se o planejamento experimental e análises subsequentes para cada cepa, avaliando-se os efeitos das variáveis e medindo o crescimento celular ao longo do tempo de fermentação. Com isso então, obtiveram-se as cinéticas para cada cepa de microrganismo separadamente, as faixas mais adequadas de cada variável, os pontos

de início de crescimento, máximo e início de decaimento de massa celular, e o tempo total para o equilíbrio pós decaimento.

Os dados de tempo de fermentação foram necessários para a seleção da melhor faixa de tempo para que se atinja o máximo de crescimento para cada cepa. Além disso, o estudo cinético possibilitou a comparação entre as cepas, a fim de selecionar aquela que apresenta melhor desempenho para que se prossiga com ela em um segundo planejamento, via delineamento composto central rotacional, permitindo a geração de análises estatísticas para os fatores de efeito positivo e geração de superfície de resposta.

Neste caso, semelhante a tratativa do tempo como variável de entrada, deve-se evitar estudar fatores cujo efeito pode ser descrito cineticamente como variáveis de entrada, para que não se tenha perda da qualidade e quantidade de informação obtida com os ensaios. O processo da ação e efeitos do uso de certas variáveis qualitativas é estudado via cinética, e, portanto, como este é o método definido originalmente para um estudo, deve-se segui-lo experimentalmente quando se deseja abstrair informações dos efeitos e comportamentos de tal variável. Não se pode sobrepor a ferramenta de planejamento experimental sobre o processo ou métodos originais.

## 5.4 CONSIDERAÇÃO DE RESPOSTAS QUALITATIVAS

Respostas qualitativas, variáveis categóricas dependentes, podem ser obtidas com o planejamento experimental, especialmente quando o objetivo e questões experimentais obtidas são direcionados a respostas qualitativas. Tais respostas, ou *outputs*, podem representar condições detectadas pelo aparelho sensorial humano, do pesquisador ou avaliador, sobre a amostra que expressa o resultado do processo em cada ensaio. Como manifestações identificáveis sensorialmente tem-se mudança de cor, separação de fases, mudança de estado físico, alterações organolépticas, etc. Outras medidas podem ser tomadas como respostas qualitativas a partir da interpretação de resultados quantitativos, mas cujo valor pode ser convertido para uma escala que expressa um atributo ou propriedade, a depender a medida utilizada

e do equipamento que fará a medição da resposta, como em sensores qualitativos portáteis que expressam resultados binários (sim ou não).

A consideração de uma resposta qualitativa requer tratamento especial antes que possa ser usada em qualquer tratamento matemático ou estatístico. É necessário representar a variável dependente qualitativa como variável dependente quantitativa por um processo chamado de codificação (*coding*).

Por exemplo, considerando uma variável dependente qualitativa ordinal  $Y_c$ , que representa “intensidade de sabor”, e assume 3 possíveis valores qualitativos (fraco, normal e intenso), temos que:

$$Y_n \begin{cases} 0, se Y_c = fraco \\ 1, se Y_c = normal \\ 2, se Y_c = intenso \end{cases}$$

Assim, para qualquer  $Y_c$  há um valor numérico  $Y_n$  correspondente, e vice versa. Ou seja, é gerada uma escala objetiva numérica, que classifica as observações dentro de *scores*, com classificação ordinal, que representam os achados categóricos. Os valores numéricos podem então ser computados como respostas que se relacionam com observações (*outputs*) qualitativas, sem prejuízo de perda de informação em análises subsequentes (GOURIEROUX, 2000)

Como exemplo, em um estudo realizado por Sjostrom & Wikstrom (2004), buscou-se redução de custo de margarinas, com manutenção ou melhoria na qualidade sensorial. Os autores aplicaram um planejamento experimental fatorial completo para 3 fatores e dois níveis (quantidade de óleo adicionado, velocidade de mexedora e tipo de tratamento térmico), resultando em 8 ensaios e três replicatas no ponto central para cada tipo de tratamento térmico.

Neste estudo, avaliaram-se sensorialmente as amostras produzidas de margarina para cada ensaio, a fim de acessar as observações de dureza, espalhabilidade, brilho, derretimento, aroma de manteiga e sabor residual (off-flavor). Devido ao caráter qualitativo da avaliação sensorial, os autores optaram por utilizar a estratégia de codificação para atribuir a cada valor de intensidade de atributo, um valor número definido em uma escala ordinal. A escala de avaliação apresentou valores de zero a dez (0-10) para a intensidade de cada um dos atributos avaliados, sendo 0 muito fraco e 10 muito forte, com ótimo em 5 (exceto para off-flavor com ótimo em

zero). Assim, é evidenciado o uso de respostas qualitativas em estudos com aplicação de planejamento experimental, em especial nos casos onde há avaliação sensorial, inevitável para o estudo de alimentos (WIKSTRÖM; SJÖSTRÖM, 2004).

## 5.5 VARIÁVEIS DISCRETAS

Variáveis discretas são aquelas que assumem sempre valores numéricos com intervalos bem definidos dentre os valores possíveis. Tal tipo de variável pode ser medida e ordenada, mas apresenta a característica de representar em seus níveis valores contáveis, número de itens, quantidades discretas cujos valores possíveis são determinados por intervalos ou razões fixas.

Por exemplo, pode-se coletar dados sobre o número de esferas em um moinho de esferas, ou da quantidade de colônia de bactérias, e estas são variáveis discretas. Nestes casos, as variáveis colônias ou esferas assumem valores discretos inteiros e positivos, sendo impossível que se chegue à conclusão de que 20,5 esferas são adicionadas ao moinho ou de que 120,63 colônias de bactérias são formadas, pois não se admite valores menores que o intervalo de 1 unidade.

Geralmente variáveis discretas representam a contagem de itens, dentro do conjunto dos números reais, e não podem ser divididos em partes menores que o intervalo definido entre os valores possíveis. Existem casos de intervalos fracionados, nos quais valores com intervalos menores que 1 inteiro são permitidos ou mesmo valores negativos, que são possíveis principalmente quando se contabiliza a diferença entre valores.

No planejamento experimental, é necessário que se tenha cautela com o uso de variáveis discretas. Embora elas surjam como variáveis em alguns estudos, seu uso em planejamento experimental deve prever que, dentro da matriz estruturada para o planejamento, especialmente no caso da inserção de pontos axiais, a variável não codificada pode assumir valores não inteiros, o que é impossível para alguns casos de variáveis discretas. Os pontos axiais codificados como  $\alpha$  tipicamente assumem os valores máximo e mínimo da faixa de estudo da variável independente. Os demais termos da matriz, -1, 1 e 0 são então definidos por interpolação e assim, pode-se obter um valor impraticável, dada a característica discreta da variável.

Em um caso, apresentado por Rodrigues & lemma (2015) pesquisadores buscaram otimizar um método de extração para determinação de aminas em leite cru. Para isso, aplicou-se um planejamento experimental composto central rotacional, com três replicatas no ponto central, avaliando como fatores de entrada o número de extrações sucessivas, tempo de agitação e concentração do ácido utilizado na extração das aminas putrescina, histamina, espermidina e espermina.

No caso, com utilização da matriz de planejamento DCCR que faz uso de pontos axiais, tem-se que os níveis codificados definidos por  $\alpha = \sqrt[4]{2^3} = 1,68$  não são praticáveis para o caso da variável independente  $x_1$  (número de extrações sucessivas). Isso ocorre pois para que a condição  $\pm 1,68$  represente os valores máximo e mínimo do intervalo escolhido para a variável  $x_1$  no estudo, deve-se ter valores de +1 e -1 como números não inteiros (4,19 e 1,81 respectivamente), o que não é possível para a variável em questão, que diz respeito a um número de extrações, uma medida que assume apenas valores inteiros. Desta forma, os pesquisadores optaram por alterar a condição, tomando, para a variável  $x_1$  somente, valores de  $\alpha$  como sendo -2 e 2. Tal condição permite que a variável assumira somente valores inteiros no intervalo estudado.

O caso apresentado demonstra a importância de o pesquisador verificar se as condições estabelecidas pelo planejamento experimental refletem na realidade, condições praticáveis de experimentação. Não se pode seguir as condições ou parâmetros calculados e valores matematicamente obtidos, quando estes não representam medidas e valores possíveis de serem obtidos para as variáveis estudadas. Deve-se considerar e respeitar as condições da variável estudada, como as propriedades do elemento cujo a variável representa e as limitações técnicas impostas por condições ambientais, equipamentos e capacidades práticas do analista. Os pesquisadores no caso em questão fizeram a escolha mais adequada para que a variável pudesse ser inserida e seus níveis definidos da melhor forma.

Assim, o uso de variáveis discretas é possível e muitas vezes indispensável, quando se tem a variável em questão como significativa para os efeitos observados. No entanto, deve-se atentar as limitações e indivisibilidade dos intervalos, pois se uma certa variável discreta permite somente valores inteiros, o planejamento que gera valores fracionados para a variável nos ensaios se torna impraticável.

## 5.6 ESCOLHA DO TIPO DE PLANEJAMENTO

A escolha do tipo de planejamento experimental é uma etapa crítica de toda execução experimental e deve ser realizada com cautela. Antes de tudo, é recomendável que as etapas que antecedem a escolha do tipo de experimentam sejam realizadas com seriedade, a fim de preparar as condições e embasar a escolha mais adequada, evitando perda de tempo e recursos, retrabalho e possível incompletude ou inviabilidade da realização dos experimentos.

A definição do problema a ser estudado, objetivos e hipóteses constituem a fase inicial crítica e vai servir de substrato na definição das variáveis possíveis do sistema a serem consideradas no estudo. A definição de tais variáveis é uma etapa posterior e se baseia nas informações que se obtém acerca do problema e os fatores que possivelmente afetam a resposta. Todas essas informações servem de base para que se possa decidir qual tipo de planejamento mais adequado, uma vez que o número de variáveis e suas características, assim como o problema e objetivo do estudo em questão, podem ser limitantes e determinam as características necessárias para o planejamento.

Nas fases iniciais é importante que sejam considerados e construído conhecimento inicial sobre o problema a partir da participação e contribuição de todos os envolvidos, com adição de experiências, conhecimentos prévios e revisões de literatura que permitam obter insights sobre como o problema ocorre, o que está envolvido, quais são e como interagem as variáveis de forma a ocasionar a resposta observada. Assim, pode-se ter clareza dos objetivos e hipóteses, bem como dos agentes de efeito sobre o sistema.

Com os recursos em mãos, pode-se pensar sobre o melhor tipo de planejamento adequado ao objetivo inicial do estudo. A princípio, quando se obtém uma lista dos fatores que podem provocar efeito sobre a resposta, deve-se optar prioritariamente em se utilizar todos os possíveis fatores no estudo, a fim de evitar descartar variáveis erroneamente e que podem ter efeito significativo sobre o sistema. No entanto, considerar um grande número de fatores tem o poder de inviabilizar a execução do experimento, dado o crescimento exponencial do número de ensaios de acordo com o número de fatores, como é bem ilustrado para o caso de planejamentos

fatoriais completos, cujo número de ensaios é definido por  $2^k$ , sendo  $k$  o número de fatores considerado.

A fim de se tratar da melhor forma os fatores listados, existe a possibilidade de se utilizar um tipo de planejamento experimental que tem como principal objetivo realizar uma varredura sobre eles, reduzindo uma extensa lista de fatores em potencial e suas interações a alguns poucos que tenham efeitos significativos. Essa medida permite que a aplicação posterior de certos tipos de planejamento se torne viável pela redução do número de ensaios a ser realizado.

Os planejamentos de screening, ou seleção de variáveis, são úteis também para o caso de estudos que objetivam saber quais variáveis de um sistema impactam significativamente uma resposta, um estudo de caracterização. Esse tipo de estudo pode fazer bom uso do planejamento fatorial fracionado e do planejamento Plackett-Burman, que permitem com o uso de um número reduzido de ensaios, apontar os efeitos dos fatores e indicar quais deles são significativos, sendo o Plackett-Burman especialmente recomendado para um número de fatores superior a cinco. No entanto, não se pode obter um modelo do sistema com o uso destes tipos de planejamento, pois não temos graus de liberdade suficiente para definir as interações ou equações quadráticas. Para o caso de haver necessidade da construção de um modelo, deve-se optar pelo uso de outro tipo de experimento (Fatorial Completo, DCCR, Box-Behnken) subsequentemente ao planejamento de seleção de variáveis.

A seleção de variáveis tem importância também no que diz respeito a escolha do tipo de planejamento adequado na construção do modelo do sistema. A adição das replicatas no ponto central em planejamentos de seleção de variáveis permite verificar a curvatura e saber se o modelo linear é adequado ou se deve-se optar pelo modelo quadrático.

Os pontos centrais podem apresentar valores muito altos ou muito baixos com relação as demais condições, e com isso o erro padrão obtido é muito alto devido a não adequação dos resultados a um modelo de primeira ordem, o que nos indica haver uma curvatura e adequação ao um modelo quadrático. Na verificação dos efeitos dos fatores, pode-se ter a curvatura como estatisticamente significativa, evidenciando assim a adequação ao modelo de ordem superior. Essa informação então nos permite

escolher entre o uso de um planejamento fatorial completo ou um planejamento composto central com seus pontos axiais, por exemplo.

Para que se possa obter um modelo quadrático e mesmo uma superfície de resposta não se pode utilizar apenas o planejamento fatorial completo, que permite apenas o ajuste de um modelo linear, sendo necessária a inserção de pontos centrais e pontos axiais para que tenha os termos quadráticos e dimensão do erro experimental. A composição derivada do planejamento fatorial completo, com pontos axiais e replicatas no ponto central é o planejamento composto central rotacional. Tal planejamento é capaz de fornecer o modelo quadrático do sistema, assim como possibilitar gerar uma visualização gráfica em superfície de resposta.

Outro ponto importante, além de se conhecer o número de fatores, reduzir esse número para viabilizar o estudo, saber se há curvatura e definir bem o objetivo do estudo, deve-se ter em mente os recursos disponíveis, bem como as limitações técnicas do aparato experimental. Muitas vezes o planejamento define uma estrutura que deve ser seguida na realização dos ensaios, como seus níveis, repetições e combinação de fatores. No entanto, deve-se atentar a disponibilidade dos recursos que correspondem as variáveis, se estarão disponíveis e se são suficientes para a realização de todos os ensaios.

Um certo reagente, por exemplo, pode ter custo muito elevado ou estar disponível em uma quantidade muito baixa para que seja utilizado nos níveis e no número de ensaios definidos, sendo impraticável seu uso da forma que foi estabelecida, necessitando então que seja realizando um reajuste do planejamento para adequar as condições, reduzindo ensaios ou ajustando os valores dos níveis. Situações como esta podem ser evitadas, tendo de início a quantidade dos recursos disponíveis, a fim de se otimizar seu uso e evitar o término abrupto do estudo sem obtenção de resultados conclusivos.

A disponibilidade de equipamento e pessoal também é importante, especialmente quando a execução dos ensaios se dá em planta industrial, na linha de produção. Paradas de linha para realização de ensaios pode significar perdas dramáticas para a indústria, por isso, a otimização dos ensaios é necessária a fim de reduzir as perdas e manter resultados robustos.

No geral, a escolha do planejamento dá-se com base em informações prévias e considerações que levam em conta a perspectiva de tempo e recurso destinado a realização dos ensaios. Não se pode pular etapas na realização do planejamento experimental, e o estudo inicial a fim de definir objetivos, problemas, recursos e conhecimento acerca do sistema define muito da taxa de sucesso do estudo utilizando planejamento experimental, assim como a escolha adequada do tipo capaz de fornecer os resultados esperados e alinhados ao propósito. Um estudo do gênero sem propósito ou clareza, resulta em perda de tempo e recurso, sem resultados conclusivos e evidente mal uso da ferramenta.

## 5.7 FAIXA INVESTIGADA

A definição da faixa investigada para cada variável independente é uma etapa inicial que antecede a escolha do tipo de experimento. As discussões preliminares que visam coletar conhecimento acerca do sistema em estudo, permitem saber ou ao menos ajudam a definir as faixas nas quais os possíveis fatores têm seus efeitos percebidos sem representar a inviabilidade da realização dos ensaios.

Para certas variáveis, valores muito pequenos ou muito altos acabam resultando na nulidade ou excesso dos efeitos ou que se espera observar. Valores em faixas muito estreitas, com intervalos muito pequenos entre um nível e outro, podem resultar em efeitos observados, ou respostas, muito próximos entre si, de modo que a distinção entre o que seria uma medida diferenciável e o erro experimental fica prejudicada. A resposta medida, em certos casos pode não ter sensibilidade suficiente para detectar variações sutis dos efeitos provocados por variações pequenas da variável independente e mesmo variações pequenas, podem resultar em efeitos naturalmente semelhantes e equivalentes, ainda que seja possível mensurar. Já, quando se tem faixas de investigação muito amplas, tem-se o risco de não se obter valores ótimos, que permitam otimizar o sistema, por exemplo. Também, pode muitas vezes ser impraticável dada limitações técnicas que impossibilitam o ajuste fino dentro do limite estabelecido. Os equipamentos empregados, por exemplo, podem não ser capazes de atingir os níveis escolhidos devido ao intervalo discreto de valores permitidos para o ajuste ((baixo, médio e alto; 1, 5, 10), podem apresentar valor de erro que

impossibilita a precisão desejada ou mesmo precisão menor que o necessário (uso de balança semi-analítica, aparatos de baixa sensibilidade).

Além disso, quando se define os limites máximos, também deve-se considerar a capacidade máxima dos equipamentos de medição utilizados, para que sejam capazes de atingir o valor estabelecido sem representar custo elevado ou risco de operação. Especialmente no caso do nível máximo, a variável pode representar um elevado custo, como nos casos de reagentes de elevado preço e equipamentos cuja execução demanda gasto energético elevado.

Outro ponto é a disponibilidade de reagentes para atender as faixas de estudo. Dependendo dos limites máximos estipulados, a quantidade de reagentes a ser utilizada nos ensaios que fazem uso desse nível, pode representar elevado gasto do material. Deve-se avaliar a real necessidade de se estabelecer valores de nível tão elevados e prospectar a disponibilidade de reagente para que não haja interrupção do estudo.

Na definição dos valores da faixa de investigação, também deve-se levar em conta o critério de segurança. É importante ter conhecimento sobre os possíveis efeitos da diminuição ou elevação dos valores de determinadas variáveis, bem como os impactos dos valores dentro da faixa delimitada em termos de segurança experimental, pois alguns efeitos danosos podem ocorrer em determinados níveis de algumas variáveis, como por exemplo, equipamentos que podem ter sobreaquecimento, comprometimento de resultado, possibilidade de destruição da amostra ou mesmo riscos mais sérios como de mau funcionamento com chance de prejudicar a saúde do operador. Casos semelhantes podem ocorrer com o uso de reagentes químicos, que em determinadas concentrações podem provocar efeitos das reações com desprendimento de gases em volumes elevados, gases tóxicos que não podem ser controlados ou elevação de temperatura em casos exotérmicos, por exemplo.

Além dos aspectos supracitados, deve-se considerar que, após a delimitação da faixa de estudo, os níveis são inseridos dentro do que é requerido pela matriz experimental, de acordo com a codificação gerada. Para o caso de um experimento fatorial completo, temos que a codificação possível é -1, 0 e 1, que descrevem os valores máximo, mínimo e o ponto central. No entanto, quando se tem um

planejamento do tipo Composto Central Rotacional, considera-se os pontos axiais definidos por valores de  $+\alpha$  e  $-\alpha$ , tais pontos assumem valores maiores ou iguais a 1. Assim, os pontos axiais representam idealmente os valores máximos e mínimos da faixa investigada, sendo os demais níveis (-1,0,+1) definidos em função deles.

A inserção de pontos axiais após a construção da matriz de um planejamento fatorial completo, pode fazer surgir certas dificuldades. Neste caso quando a matriz do planejamento fatorial completo é construída, atribui-se à codificação -1 e +1, os valores mínimos e máximos da faixa investigada, respectivamente. No entanto, a inserção posterior dos pontos axiais no composto central rotacional, codificados por  $+\alpha$  e  $-\alpha$ , resulta na atribuição de valores de nível abaixo do valor mínimo e acima do valor máximo da faixa investigada. Isso pode resultar, em alguns casos, na impraticabilidade do estudo, dada a possibilidade de haver valores negativos como nível da variável ou valores que comprometem o que fora planejado considerando as limitações experimentais. Neste caso, pode-se optar pela execução de um delineamento de faces centradas, no qual tem-se  $\alpha=1$ , resultando em valores dentro da faixa utilizada. No entanto, de forma geral, a definição dos níveis e da faixa investigada exige cautela, e que se faça uso de todas as informações prévias possíveis na definição dos níveis dentro da matriz experimental do planejamento escolhido.

## 5.8 IMPORTÂNCIA DO PONTO CENTRAL

Pontos centrais são ensaios realizados nas coordenadas 0, na origem da escala codificadas dos fatores, que representa o centro matemático entre os valores máximo e mínimo definidos pela codificação +1 e -1. Normalmente executa-se ao menos três repetições no ponto central para que se tenha seu potencial de uso e real utilidade.

A inserção de três ou mais replicatas no ponto central permite avaliar a existência de curvatura, ou seja, permite distinguir entre dois modelos possíveis para os dados, modelos de primeira ou de segunda ordem. Com replicatas no ponto central é possível verificar se há o ajuste a um modelo linear ou deve-se considerar termos quadráticos que caracterizam o modelo de segunda ordem e incorporam curvatura na

modelagem da resposta, o que não é possível com a avaliação em apenas dois níveis (+1, -1). Além disso, há o benefício de poder estimar o erro puro para que se teste a falta de ajuste do modelo, que indica o quão bem o modelo escolhido se ajusta aos dados pela obtenção do desvio de cada medida observada com relação à média.

Em um planejamento fatorial, por exemplo, a adição de medidas no ponto central permite estimar o erro experimental, uma grande vantagem, pois a repetição das medidas nos pontos dos fatores em duplicata dobraria o número de medidas de cada ensaio, podendo inviabilizar a execução do estudo. No entanto, ainda que não se tenha repetições das medidas dos pontos dos fatores, as repetições no ponto central possibilitam a obtenção da estimativa de erro.

A estimativa do erro puro permite a avaliação da repetibilidade do processo, ou seja, permite verificar se se tem o controle sobre o processo. Assim, caso haja uma etapa do processo sob a qual não se tem completo controle, como no caso de limitações e falta de acurácia de equipamentos de medida, medidas como o erro puro ou erro padrão, obtidos a partir dos pontos centrais, podem fornecer informações importantes sobre a condição do processo e potencial de redução de variabilidade. Além disso, pode-se identificar a definição de valores dos limites mínimos e máximos inadequados, como quando a resposta nestes limites é aproximadamente igual, sugerindo que a faixa estudada não compreende efeitos significativos das variáveis. No entanto, medidas de ponto central podem evidenciar a existência de curvatura, tendo valores de resposta para o nível 0, maiores ou menores.

Existe, entretanto, situações nas quais não se pode definir o ponto central, como no caso de se utilizar fatores categóricos, ou a impossibilidade de se definir o valor no nível 0, dada restrições de intervalo permitido, como no caso de fatores discretos. Em situações como essa, com o uso de fatores categóricos ou discretos, deve-se definir qual dos níveis (-1 ou +1) representa o padrão, para que este seja considerado como um “pseudo ponto central”, e possa ser utilizado em conjunto com os demais fatores em seus pontos centrais, sem que haja o prejuízo de se abster de seu uso.

Quando se faz uso de um experimento com apenas dois níveis, como em um planejamento fatorial  $2^k$ , assume-se que a relação dos efeitos dos fatores com a resposta é linear, ainda que se tenham termos de interação na equação obtida para o

modelo. A utilização de apenas dois níveis não permite que se ajuste modelos mais complexos e além disso, não se sabe se há de fato uma dependência linear dos fatores com a resposta.

Caso não se tenha uma função linear, conclusões e observações feitas a partir do modelo linear não podem então ser consideradas verdadeiras. O estudo realizado servirá de base para tomada de decisão e predição de respostas baseado no modelo ajustado e assim, é necessário que o ajuste seja adequado a fim de evitar que se obtenha conclusões equivocadas.

Como suporte a isso, pode-se avaliar a falta de ajuste do modelo, que uma vez confirmada nos diz que o modelo é inadequado, destacando que termos de ordem superior envolvendo os fatores já presentes no modelo são necessários para compor a equação que representa a relação dos fatores com a resposta. Essa medida se baseia na diferença entre o valor predito no ponto central e o valor observado neste ponto, com a identificação de desvio da relação linear dos fatores com a resposta. De forma equivalente, os pontos centrais permitem responder se a presunção de linearidade adotada para o planejamento fatorial é satisfeita testando se há de fato a curvatura zero do modelo.

Para um modelo de um planejamento fatorial  $2^2$ , considerando  $n_f$  o número de pontos dos fatores no planejamento. Para testar a hipótese de curvatura considera-se a inserção de  $n_c$  pontos centrais. Sendo  $y_f$  e  $y_c$  as médias das respostas para os pontos de fatores, se não há curvatura presente é esperado que  $y_f$  se aproxime muito do valor de  $y_c$ , assim, a diferença  $y_f - y_c$  se aproxima de zero. No entanto, quando a diferença  $y_f - y_c$  é grande e significativa, temos que a curvatura é presente. Vale ressaltar que, essa diferença dificilmente será igual a zero devido a existência de erro aleatório e por isso, deve-se verificar se o valor obtido é significativo pelo cálculo da relação sinal-ruído, que faz uso da diferença entre as médias, seu desvio padrão e a comparação do valor obtido para a relação sinal ruído com valores críticos tabelados da distribuição de t de student. Caso o valor de t calculado seja superior ao t tabelado, temos que o valor é significativo.

Uma outra forma de visualizar essa medida é com a verificação do parâmetro  $\delta$ , que representa a discrepância entre os valores esperado no ponto central ( $\beta_0$ ) e o valor esperado da resposta no ponto central  $E(y_{pc})$ , a fim de testar se a curvatura é

nula. Para isso, temos que  $\beta_1$  e  $\beta_2$  representam os efeitos do primeiro e do segundo fator, respectivamente,  $\beta_{12}$  representa o efeito da interação entre os fatores na equação do modelo para o caso  $2^2$ :

$$E(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \delta$$

De acordo com a equação, no ponto central temos que  $\delta = E(y_{pc}) - \beta_0$ , sendo  $x_1$  e  $x_2$  iguais a zero. Assim, quando o termo  $\delta$  é não nulo e significativo, tem-se a existência de uma discrepância entre o valor esperado e a resposta medida, indicando a necessidade do ajuste de um modelo com termos quadráticos.

Muito comumente, na literatura, não se inclui pontos centrais no estudo com aplicação de planejamento Plackett-Burman, no entanto, em situação apresentada exemplificada em um estudo de caso por Rodrigues & Lemma (2015) tem-se que os pontos centrais, quando utilizados, podem apresentar valores muito altos ou muito baixos do que as demais ensaios realizados, elevando o erro padrão. Essa observação é um indicativo de que os resultados não se adequam a um modelo de primeira ordem, sendo necessário a avaliação da curvatura. No exemplo, é encontrado valores de resposta muito maiores para os pontos centrais, com relação aos demais ensaios, e como consequência, na análise dos efeitos das variáveis, tem-se elevados valores de erro padrão e elevados p-valores, mascarando os efeitos das variáveis. No entanto, quando a análise é performada considerando a checagem da curvatura, tem-se a redução do erro padrão e p-valores muito menores são obtidos para as variáveis, com curvatura significativa. A inserção dos pontos centrais permitiu avaliar não somente os efeitos das variáveis, como também checar a adequação dos resultados ao modelo de primeira ordem, permitindo que decisões mais acertadas acerca das próximas etapas sejam tomadas.

Em casos especiais, como em simulações computacionais com aplicação de planejamento experimental, temos que as medidas de respostas simuladas apresentam valores idênticos a cada replicata e, portanto, tende a apresentar variabilidade igual a zero. Desta forma, não se faz uso de mais de um único ensaio no ponto central. Este ensaio realizado no ponto central resultará em um valor de resposta que permitirá verificar a existência de curvatura potencial e checagem da adequação da faixa de estudo cujo efeito das variáveis sejam significativos sobre a resposta.

## 5.9 PLANEJAMENTOS FRACIONADOS

Experimentos fracionados são particularmente interessantes quando se tem uma quantidade elevada de variáveis independentes ( $k \geq 4$ ) inicialmente consideradas para o estudo e que podem significar a inviabilidade de sua realização, dado o elevado custo em recursos, tempo e pessoal. Diante de um número elevado de variáveis ou fatores é conveniente iniciar o estudo empregando um planejamento fatorial a fim de reduzir o número de variáveis a serem consideradas, de acordo com seus efeitos, de modo que sejam selecionadas neste momento somente aquelas que apresentarem efeitos significativos sobre a resposta. Assim, representa um ganho de tempo e recursos considerável, uma vez que o número de ensaios a ser realizados cresce exponencialmente de acordo com o número de variáveis, como no caso de um planejamento fatorial completo  $2^k$ , no qual  $k$  é o número de variáveis.

Apesar de reduzir o tamanho do experimento, resultando em ganho de tempo e recursos, existe um *tradeoff* a ser considerado, pois informações importantes são perdidas devido a eliminação de parte das combinações possíveis dos níveis dos fatores de interesse (menor número de ensaios). Essa condição impacta nos resultados obtidos com efeitos de confundimento, ou seja, alguns efeitos não podem ser separados e, em especial interações, são confundidas com outras interações ou mesmo com os efeitos principais.

Normalmente a seleção do planejamento fracionado mais adequado leva em conta não somente a redução do número de ensaios a uma quantidade executável, mas deve considerar quais informações estarão comprometidas. A depender da resolução do planejamento fracionado, determinadas interações estarão confundidas e para cada caso deve-se saber se os efeitos de interesse não sofrem os efeitos de confundimento, especialmente se não são afetados os efeitos principais ou de primeira ordem, tipicamente de interesse nos estudos. Assim, o sacrifício feito deve ser o de estimar interações, geralmente de ordem superior, que sejam dispensáveis.

Caso se tenha um planejamento fatorial  $2^k$  com quatro variáveis (A,B,C e D) e cada uma delas com dois níveis, e a fim de realizar uma seleção de variáveis e reduzir o número de ensaios iniciais com esse objetivo, pode-se optar por um planejamento fracionado  $2^{4-1}$ , por exemplo. Pode-se obter uma matriz experimental composta por 8

ensaios e com potencial teórico de se estimar 14 efeitos (principais e de interações), como mostrado na Tabela 7.

**TABELA 7** – Matriz Planejamento  $2^{4-1}$  com Interações

Ensaio	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
1	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+
2	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+
3	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
4	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
5	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
6	+	-	+	-	-	+	-	-	+	-	-	+	-	+	+
7	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Fonte: Elaboração Própria

A coluna que apresenta o efeito de ABCD, tem apenas sinais positivos, ou seja, não se pode estimar os efeitos de ordem quatro. Ainda que se tenha 14 efeitos possíveis de serem estimados, o planejamento conta com somente 8 ensaios, não sendo matematicamente possível que se tenha todos os efeitos isolados mensurados, pois existem apenas 8 graus de liberdade para isso.

Ainda observando a matriz, pode-se notar que o padrão da coluna AB e da coluna CD são idênticos. Então se a interação AB é estimada, a interação CD também será estimada simultaneamente, tendo seus efeitos misturados e inseparáveis, ou confundidos. Neste caso, cada interação de segunda ordem é confundida com outra interação de segunda ordem. Observa-se também que a coluna A tem padrão idêntico a coluna BCD e, neste caso, também se tem efeito de confundimento do efeito principal e uma interação de terceira ordem. O exemplo em questão ilustra a problemática envolvida no uso de planejamentos fatoriais fracionados e os riscos que se corre com os efeitos inerentes de confundimento, o que limita muito seu uso, restringindo a avaliação de efeitos.

O uso deste tipo de planejamento, fracionado, deve ser visto como uma etapa inicial e exploratória dos efeitos e significância destes para a resposta estudada. Não existe a possibilidade de fazer uso deste tipo de planejamento para atender todas as expectativas quanto ao estudo realizado e, assim, não se pode usá-lo para ajustar um modelo ou gerar superfície de resposta. Claro que, se o objetivo do estudo for o de

identificar as variáveis que exercem efeito significativo na resposta do sistema, pode-se validar o achado em um experimento e, a princípio, se exime da necessidade imediata de se prosseguir com modelo e otimização.

Um exemplo interessante é apresentado por Rodrigues & lemma (2015), onde a utilização de planejamento fatorial fracionado foi importante na seleção das variáveis significativas antes do uso de um planejamento composto central com geração de superfície de resposta. No caso apresentado, o objetivo era estudar a síntese de oligossacarídeos com o uso da enzima inulinase em um meio orgânico, estudando as variáveis temperatura, pH, percentual de sacarose, atividade enzimática, razão solvente-tampão e concentração de polietilenoglicol (PEG; estabilizante). As respostas observadas foram rendimento de frutooligosacarídeos e a produtividade do processo.

A consideração de todas as variáveis e respostas resultaria em 80 ensaios em um planejamento composto central rotacional, algo impraticável dado o dispêndio de tempo e recursos. Optou-se então pelo uso preliminar de um planejamento para seleção de variáveis, sendo dois planejamentos possíveis, um planejamento fatorial fracionado  $2^{6-1}$  (36 ensaios) e outro  $2^{6-2}$  (20 ensaios).

Os autores selecionaram então o planejamento  $2^{6-2}$ , que apresenta condições ideais para o objetivo de seu uso, pois os efeitos principais não são confundidos com os efeitos de primeira ordem, preservando a qualidade e confiança da informação obtida. Neste caso não é possível obter informações das interações entre variáveis, devido aos efeitos de confundimento, mas ainda assim o planejamento cumpre seu propósito de seleção de variáveis pelos efeitos principais. Foram selecionadas 4 variáveis das 6 iniciais. Em sequência utilizou-se um outro planejamento fatorial fracionado,  $2^{4-1}$  diminuindo o número de variáveis para apenas duas cujo efeitos são significativos, e por fim, o planejamento CCR com superfície de resposta.

O exemplo apresentado demonstra o uso do planejamento fracionado na seleção de variáveis, que auxilia a execução do estudo com muitas variáveis e também o zelo na escolha do planejamento adequado ao propósito do estudo e suas etapas. Os planejamentos fracionados destinam-se a visualização e dimensionamento prévio dos efeitos das variáveis inicialmente prospectadas para o estudo, a fim de selecionar as variáveis com maiores efeitos significativos.

Um aspecto importante dos planejamentos fracionados, é a inerente característica de confundimento dos efeitos dos fatores, que tendem a estar misturados, e assim, não se pode distinguir todos os efeitos e não se pode construir um modelo válido que considere os efeitos de interação ou de segunda ordem. Assim, o uso de planejamentos fracionados se limita a seleção de variáveis e verificação dos efeitos principais ou de interações primeira ordem.

Embora sua aplicação seja mais restrita, pode-se encontrar na literatura muitos exemplos de estudos que fazem uso de planejamentos fracionados e que, baseado nestes, obtêm superfícies de resposta. O modelo de superfície de resposta envolve os efeitos principais e interações entre múltiplos fatores quando não há indícios de curvatura, no entanto, em certas circunstâncias o modelo completo somente pode ser descrito quando se insere termos quadráticos ou mesmo cúbicos na equação que descreve o comportamento do processo. Porém, quando se tem um planejamento fracionado temos os efeitos de confundimento que não permite distinguir os efeitos de interação ou de ordem superior e, portanto, não seria possível ou correto prosseguir com superfícies de resposta ou ajuste de modelo, por exemplo.

Outro ponto importante é a verificação de curvatura, que pode ser realizada utilizando-se um planejamento fatorial fracionado, com a inserção de pontos centrais, que permite, como pontuado no item 5.8, avaliar se há a necessidade da consideração dos termos quadráticos na composição da equação que modela a resposta do sistema em função das variáveis estudadas. A verificação da curvatura é uma etapa importante que pode ser realizada durante a seleção de variáveis. A não realização desta etapa pode resultar na escolha inadequada do planejamento experimental subsequente, baseado na suposição prévia de relação linear, com a não consideração de potencial curvatura e uma equação quadrática.

## 5.10 PLANEJAMENTOS PLACKETT-BURMAN

Planejamentos Plackett-Burman fazem parte de uma classe de planejamentos saturados fracionários de resolução III que fazem uso de um número definido de ensaios múltiplos de 4 (8,12, 16, 20 ,etc.), no estudo de um número de fatores que

não podem, idealmente, ultrapassar uma quantidade de  $n-4$ , sendo  $n$  o número de ensaios da matriz selecionada (garantindo ao menos 4 variáveis inertes). Assim como os planejamentos fatoriais fracionados, esta classe geralmente se destina a estudos de seleção de variáveis, quando se busca identificar as variáveis capazes de impactar significativamente uma resposta ou diminuir o número de fatores a serem utilizados nos planejamentos subsequentes, com um número mínimo de ensaios necessários.

Tais planejamentos são geralmente não geométricos e apresentam resolução III com complexo efeito de confundimento que impacta diretamente nas informações possíveis de serem obtidas nos estudos. Planejamentos dessa classe que não sejam obtidos por uma potência de dois ( $2^n$ ) possuem cada um dos efeitos principais confundidos com as interações, que quando possuem efeitos grandes, podem distorcer os efeitos estimados. Assim, como efeito prático, apenas os efeitos principais devem ser considerados no estudo que faz uso deste planejamento e assume-se as interações como desprezíveis e de menor efeito, não considerando-as em análise ou conclusões, dado efeito de confundimento.

A avaliação prévia, com dimensionamento hipotético dos efeitos de interação, se faz necessária para que não se comprometa o objetivo da aplicação deste tipo de planejamento, que é a seleção de variáveis com base nos efeitos principais. O confundimento complexo dificulta a separação dos efeitos principais e de interação e, portanto, os efeitos de interação não podem ser considerados nesta fase de aplicação do planejamento PB e devem ser desprezíveis. Assim, mais uma vez é ressaltada a importância do conhecimento prévio do problema, hipóteses iniciais e do objetivo claro do estudo e de suas etapas.

Um detalhe importante do planejamento experimental Plackett-Burman diz respeito as colunas vazias ou variáveis inertes na matriz do planejamento. Idealmente recorre-se ao uso de ao menos 4 variáveis inertes, ou colunas vazias na matriz PB, para que se tenham graus de liberdade a fim de se obter apropriadamente a estimativa do erro experimental. Assim, não se pode selecionar uma matriz PB com 12 ensaios para que sejam avaliados números de fatores maiores do que 8, ou em outras palavras, deve-se ter ao menos 4 colunas a mais do que número de variáveis estudadas.

Devido ao fato de não se atribuir variáveis às colunas ditas vazias, o resultado da variação entre os valores obtidos para estas colunas tende a zero na situação hipotética de não haver fontes de erro. Assim, caso haja qualquer outra fonte de ruído no processo impactando a resposta medida, perceber ia-se uma tendência de afastamento da variância zero e ele poderá ser considerado na estimativa do erro experimental. Além disso, como a matriz PB apresenta apenas dois níveis dos fatores (+ e -), os resultados não são suficientes para se obter um modelo matemático adequado, tendo em vista que o modelo de resposta exibirá variação linear como uma linha entre as duas respostas medidas. Assim, similar aos planejamentos fatoriais fracionados, é necessário que sejam feitas medidas no ponto central, como no mínimo três repetições a fim de permitir a avaliação da repetibilidade do processo e obter informações sobre o controle que se tem sobre ele, resultando em dados que permitem minimizar as fontes de ruído e verificar potencial curvatura do modelo.

Outro ponto de atenção trata-se do uso de planejamentos fracionados como planejamentos finais capazes de fornecer uma superfície de resposta confiável, excedendo sua finalidade como planejamentos de seleção de variáveis. Alguns autores realizam o uso inadequado dos planejamentos fracionados, incluindo o planejamento Plackett-Burman, o que nos leva a questionamento a respeito da validade das informações obtidas na análise.

## 5.11 RANDOMIZAÇÃO

A randomização, ou aleatorização, tem a finalidade equilibrar os efeitos de ruídos ou efeitos provocados por fatores interferentes externos, não controlados, que são capazes de provocar efeitos nas variáveis resposta. Com a randomização, ou sorteio, da ordem de realização dos ensaios experimentais, os fatores interferentes e efeitos de ruído são distribuídos igualmente pelos ensaios, tendo assim igual probabilidade de afetar cada uma das respostas.

Com isso tem-se redução do impacto dos efeitos, ou sua “diluição”, com diminuição do erro experimental, eliminação do viés de seleção e possibilidade de

inferências estatísticas válidas, com testes de significância. O teste de significância tem em sua base a randomização e permite dizer se uma significância pequena atribuída a uma diferença observada não é efeito apenas do acaso, excluindo explicações de aleatoriedade e mantendo apenas aquelas relacionadas ao tratamento imposto.

A randomização é um dos princípios básicos do Planejamento Experimental e a experimentação por si é sustentada pela ideia de que as amostras estudadas são parte representativa de um todo e que os resultados estatísticos obtidos por meio destas podem ser generalizados. Ela é uma metodologia que se baseia na afirmativa de que não é possível eliminar todas as fontes de variação externas em um experimento e haverão, portanto, diferenças, ainda que sutis, entre as medidas efetuadas. A randomização é, então, tida como uma maneira de dispersar e equilibrar a variabilidade externa no conjunto de medidas e eliminar vieses nos resultados originados a partir das diferenças não controláveis nas condições experimentais.

Quando se aloca os ensaios de uma maneira na qual a escolha pessoal está voluntária ou involuntariamente envolvida, temos o viés de seleção. É muito difícil, e talvez improvável, que critérios pessoais não estejam envolvidos quando uma seleção é feita sem randomização.

Mesmo em estudos bem construídos e monitorados, as medidas e as condições de ensaio estão sujeitas a variações aleatórias dos valores, chamadas de ruídos ou mesmo sofrer interferências de fatores ocultos não controlados. Temperatura ou umidade do ambiente, por exemplo, são fatores que afetam as condições experimentais e estão a princípio fora do controle, podendo acelerar reações químicas, ocasionar diminuição de viscosidade ou adicionar água ao meio, que são alterações importantes em processos químicos. Podem existir também flutuações aleatórias de diversas origens que podem ocorrer em medidas instrumentais. Assim, ruídos ou fatores externos não controlados podem exercer efeitos de variação nos resultados e pode haver o comprometimento do estudo e perda da confiabilidade, pois não se pode dizer se o resultado obtido se refere puramente aos efeitos do tratamento ou vem de outras fontes.

Além disso, quando a alocação dos ensaios se dá a partir de uma escolha pessoal voluntária ou involuntária, temos o viés de seleção. É muito difícil, e talvez

improvável, que critérios pessoais não estejam envolvidos quando uma seleção é feita sem aleatorização, e este tipo de viés pode comprometer seriamente o estudo.

A randomização a princípio não tem custo e é facilmente aplicada ao planejamento experimental, apesar de oferecer excelentes ganhos quanto a robustez e confiabilidade dos resultados obtidos. Sendo uma medida simples que resulta em melhoria do resultado, é inegável que sua aplicação se faz necessária e completamente acessível em estudos com planejamento experimental, embora possam existir situações nas quais seu uso é dificultado.

Algumas situações nas quais a ordem originada da randomização completa dos ensaios é impraticável ou indesejada podem existir e isso é geralmente devido às limitações de processos ou equipamentos, especialmente no caso industrial. Alguns fatores podem ser difíceis de terem seus níveis alterados em determinados processos, podem ter elevados custos, tempo e esforço operacional, o que pode desfavorecer o respeito a ordem de execução de ensaios.

No entanto, tais questões não impedem a randomização, mas restringe a ordem obtida a uma sequência praticável e, portanto, é preferível que se trabalhe com uma randomização adequada as limitações que se tem a se abster de seu uso. Neste caso pode-se fixar os ensaios nos quais se tenham os fatores com restrições significativas das mudanças de nível e randomizar os demais ensaios nos quais os fatores não apresentam maiores restrições.

Um exemplo que pode ser seguido é o seguinte. Suponha que existam quatro fatores, A, B, C e D. Destes quatro, o fator B tem mudança de nível dificultada por questões operacionais e de maquinário, enquanto os demais não apresentam tais restrições. Pode-se neste caso empregar uma randomização dos ensaios em grupos distintos nos quais os critérios de agrupamento são os níveis do fator B, ou seja, são separados blocos nos quais o nível de B é (+1), (-1) e (0), por exemplo. Assim, fixando o valor do fator B em cada grupo, pode-se randomizar os ensaios nos grupos sem que seja alterado o nível do fator com restrições.

Porém, cabe cautela a atenção às advertências desta prática. A separação da randomização em grupos permite que efeitos de variáveis ocultas, não controladas, como aumento da efetividade de determinado processo com o tempo, exiba efeitos sobre o fator fixado (no caso exemplo, B), sendo necessário que se realizem replicatas

do fator B para acessar seu efeito de forma mais adequada. No geral, é recomendado que se verifique a real necessidade do uso deste método, uma vez que se pode realizar ajustes que permitam a mudança facilitada dos níveis dos fatores com restrições, descartá-los ou mantê-los constante, mediante verificação de significância em um planejamento de seleção de variáveis e avaliação prévia que antecede o planejamento em si.

Embora a randomização seja uma prática de fácil execução e praticamente sem custo, na literatura pode existir estudos que não fizeram uso da randomização e conseqüentemente tem em seus resultados experimentais a possibilidade de interferências significativas de efeitos de ruído, fatores ocultos não controlados e outros vieses. A eliminação das fontes de ruído e total controle das variáveis envolvidas, incluindo as ocultas, é praticamente impossível na experimentação real e, portanto, se abster de métodos que visam considerar a existência de tais efeitos e promover a distribuição de forma igual por todos os ensaios a fim de possibilitar a adequada inferência estatística é colocar intencionalmente em cheque a validade e confiabilidade dos resultados obtidos.

Randomização é um dos princípios básicos do planejamento experimental, e tem sua importância na execução adequada do estudo, com impacto nas análises e respostas obtidas, bem como na análise estatística subsequente. É um método que deve ser inserido no planejamento experimental como forma de atenuar os efeitos de ruídos, variáveis ocultas e vieses sobre os resultados obtidos, permitindo que se tenha confiabilidade dos resultados e validade nas análises estatísticas efetuadas.

## 5.12 ANÁLISE DOS RESULTADOS ESTATÍSTICOS

Como parte da realização do estudo, a cada ensaio são realizadas medidas das respostas, caracterizadas a princípio como um conjunto de valores numéricos medidos e que podem fornecer algumas informações. Para cada ensaio são obtidos um conjunto diferente desses números medidos, referente a cada uma das respostas avaliadas no estudo. Tais números diferem para cada ensaio, seja por conta das variabilidades provenientes de ruídos, variáveis ocultas, erro experimental ou diferenças devido aos efeitos das variáveis independentes. Todo esse conjunto de

dados numéricos e suas diferenças servirão como substrato para que sejam extraídas informações significativas do comportamento das respostas e da influência das variáveis independentes sobre este comportamento. Para isso, a análise estatística é parte fundamental do estudo e trata-se de uma ferramenta matemática capaz de prover essas informações significativas necessárias para que se obtenha conclusões válidas e auxílio na tomada de decisões.

A análise estatística é uma das etapas da execução do planejamento experimental, mas não é uma última instância ou definitiva. Após grande esforço e tempo destinados as discussões iniciais, escolha do planejamento e execução dos ensaios, a etapa estatística é parte do desfecho do trabalho, mas não representa por si só uma conclusão absoluta e segura. A análise estatística se baseia nos resultados brutos obtidos e as conclusões abstraídas das análises se baseiam no conhecimento prévio do experimentador, nas informações buscadas na literatura e no bom senso, que permitem avaliar se a análise faz sentido e responde aos questionamentos de forma adequada.

O planejamento de experimentos demanda tempo e recursos consideráveis e não se pode permitir que seja feito mau uso dos dados obtidos, supondo a obtenção adequada destes, na geração de conclusões equivocadas ou na abstenção de se obter conclusões, que por conseguinte poderão servir de recurso na tomada de decisões importantes, especialmente na indústria. Assim, é de suma importância que haja a discussão, observação e que se recorre a conhecimentos anteriores, presentes e da literatura, a fim de sustentar, corroborar ou fazer entender os achados que se tem com os dados brutos, anteriores ao uso das ferramentas estatísticas.

A avaliação constante do que tem sido feito na execução do planejamento experimental não é menos importante do que aquela realizada previamente a concepção do estudo e deve ser realizada com cautela e apreço aos trabalhos realizados. Muita informação pode ser obtida e isso pode significar uma maior facilidade na interpretação dos dados estatísticos subsequentes, sem que haja desentendimentos e fatalismo amparado nos resultados estatísticos, que porventura, possam parecer um ultimato na validação ou desqualificação completa dos dados.

Tendo dito isso, alguns pontos podem ser discutidos sobre aspectos das análises que podem gerar interpretações equivocadas ou induzir a erros nos cálculos

efetuados e conseqüentemente nas conclusões obtidas com o estudo. São particularidades que podem ocorrer e que precisam de uma observação simples, mas com certo cuidado para que não se abra mão daquilo que foi obtido ou dos cuidados nas interpretações.

Existem casos de estudos nos quais a análise estatística de variância (ANOVA) pode apresentar resultados inesperados e gerar dúvidas sobre a qualidade e validade dos cálculos ou resultados obtidos, especialmente quando se debruça sobre a estatística sem se atentar aos dados brutos obtidos, em uma análise igualmente importante.

Um exemplo notório é a obtenção da totalidade dos fatores avaliados como sendo não significativos após a análise de variância. Essa condição pode evidenciar informações sobre a robustez do processo, pois um processo robusto apresenta variabilidade na resposta muito baixa, com certo grau de insensibilidade aos efeitos de ruídos e variáveis não controladas. Assim, em um processo robusto, é esperado que variações nos fatores avaliados nos processos resultem em respostas com baixa variabilidade, ou seja, sem impactos significativos.

Esse tipo de situação destacam a importância da avaliação prévia dos resultados obtidos, com uma observação atenta das correlações aparentes e das variações apresentadas. Os resultados podem apresentar a robustez do processo quando as respostas avaliadas se enquadram dentro dos limites estipulados como adequados e ideais. Quando isso ocorre, para todos os ensaios realizados, é um indicativo de que as variações nos níveis dos fatores promovidas nos ensaios não foram capazes de alterar as respostas de forma a impactar os resultados esperados do processo.

Um exemplo interessante, apresentado por Rodrigues & Lemma (2015), relata um estudo que objetivava formular um pão sem açúcar adicionado, gordura e emulsificantes, com o uso de amido modificado de mandioca, enzima amiloglicosidase e fosfolipase como variáveis independentes. Para isso, usou-se um planejamento DCCR com respostas de cor, volume, umidade e textura em diferentes tempos. Os pesquisadores encontraram que, dentro da faixa estudada, nenhum dos fatores eram estatisticamente significativos, não sendo possível estabelecer um modelo. O comentário indicado no caso recorre as medidas no ponto central para

evidenciar que, independentemente dos níveis das variáveis independentes estudadas, os resultados obtidos foram muito próximos e, assim, é possível concluir que com adição ou não dos ingredientes estudados como fatores, não se obtém alterações significativas nas variáveis resposta. Ao final do estudo então optou-se por selecionar a opção mais econômica, uma vez que para qualquer dosagem dentro da faixa estudada, os resultados não diferem. É importante destacar que para todas as avaliações recorreu-se aos conhecimentos prévios e da literatura que são capazes de explicar e corroborar os achados, tornando possível a verificação da validade dos dados obtidos quanto a sua coerência.

Assim, a avaliação dos resultados previamente a análise estatística, a busca constante por informações prévias e da literatura que corroborem os achados, assim como o apelo ao bom senso, são parte fundamental para que se tenha uma análise estatística completa, sem olhar única e exclusivamente para os dados de ANOVA. As análises são complementares e auxiliam nas conclusões e tomadas de decisões, evidenciando aspectos pouco considerados como no caso exemplificado de robustez de processo e possibilidade de se fazer uso dos valores dentro de uma faixa de estudo em sua totalidade, quando há a não significância dos efeitos das variáveis independentes sobre as respostas.

Ainda sobre análise de variância, quando se obtém os resultados para regressão e resíduo é necessário realizar uma etapa de verificação da validade e qualidade do modelo, que consiste em avaliar se o valor de F calculado é maior do que o valor de F tabelado, com conseqüente p-valor menor do que o valor de alfa (significância) para o caso. Quando esse critério é satisfeito, temos que a hipótese nula é rejeitada pelo teste F dentro da significância estabelecida. Essa verificação nos diz se o modelo estatístico é válido e se poderá ser utilizado nas etapas subsequentes, embora trata-se apenas de uma das etapas de verificação. A não realização desta verificação coloca em risco as conclusões a serem obtidas, pois pode ser adotado um modelo não válido e que, portanto, não é capaz de gerar resultados corretos, com validade estatística.

Outra verificação que deve ser feita diz respeito aos resíduos, diferença do modelo com relação ao dado real, que pode ser decomposto em Falta de Ajuste e Erro puro, sendo a falta de ajuste a distância entre o valor de resposta médio e o valor de resposta estimado, e o erro puro a distância entre o valor real e o valor médio. O

resíduo é composto então da soma destas duas diferenças entre valores de respostas. O resíduo deve apresentar valor de F calculado menor do que o F tabelado, de forma que o p-valor seja maior do que alfa e, portanto, não apresente significância estatística, aceitando a hipótese nula. Quando a hipótese nula não é rejeitada no teste F, tem-se que o modelo não é adequado para descrever o comportamento da resposta em relação a variável independente.

No entanto, pode existir uma situação peculiar na qual o Erro Puro apresenta valor de Soma dos Quadrados (SQ) muito menor do que SQ de Falta de Ajuste. Neste caso pode-se obter F calculado maior do que o F tabelado e assim, a significância da Falta de Ajuste e invalidação do modelo, mesmo tendo regressão válida e coeficiente de determinação adequado. Isso ocorre especialmente em casos de processos bem controlados, no qual a variabilidade encontrada nas medidas em replicatas é muito baixa, os valores obtidos são muito próximos. Como o F calculado é obtido pela razão entre os quadrados médios  $QM_{\text{falta de ajuste}}$  e  $QM_{\text{erro puro}}$ , sendo  $QM_{\text{erro puro}}$  muito menor do que  $QM_{\text{falta de ajuste}}$ , tem-se que  $F_{\text{calc}}$  tende a atingir valores elevados e, assim, apresente significância estatística como um falso positivo.

Por fim, tendo validado pela verificação do teste F da regressão e da falta de ajuste, pode-se verificar o coeficiente de determinação  $R^2$ , que representa o percentual de dados reais que podem ser explicados pelo modelo gerado. Este deve ser suficiente, a depender de cada processo e, portanto, deve-se observar se o processo em questão permite que o  $R^2$  em obtido seja aceito.

No geral, o coeficiente de determinação é uma medida que diz o quão bem o modelo se ajusta aos dados, mas não deve ser avaliado sozinho, é necessário que a regressão e os resíduos sejam verificados, em conjunto ao conhecimento prévio, para que se tenha uma visualização adequada de seu significado. Ele, apesar de representar uma ideia da correlação dos dados com o modelo, não é capaz de dizer nada sobre a validade estatística dessa correlação, e pode-se obter valores elevados de  $R^2$  para um modelo inadequado aos dados. Podem ocorrer, inclusive, situações nas quais valores baixos de  $R^2$  são aceitáveis, dada a dificuldade de se gerar um modelo preditivo acurado, como, por exemplo, no caso de estudos que lidam com seres humanos, que apresentam padrões comportamentais difíceis de se prever. Assim, a verificação do  $R^2$  é feita acompanhada das demais, e não tem significado de validação do modelo sendo avaliado sozinho.

Todos os fatores analisados na Análise de Variância devem ser verificados e satisfeitos nos critérios de validade estatística para que se tenha um modelo estatisticamente válido. Somente assim o modelo pode ser utilizado, prosseguindo com superfície de resposta e obtenção com sua validação.

Se o modelo não for válido estatisticamente, ou seja, algum dos critérios não for satisfeito, deve-se observar as condições do processo estudado e da realização dos ensaios. Deve-se observar situações como a presença de um processo que tenha um erro puro muito elevado com alta variabilidade intrínseca do processo de baixo controle, ou a faixa de estudo selecionada pode não ser adequada para que se observe efeitos das variáveis independentes nas respostas (como em processos robustos ou uso de faixa de resultados ótimos) ou mesmo uma falta de ajuste muito elevada. Ainda que o modelo não seja válido, tem-se uma quantidade grande de dados que podem ser observados e discutidos, tomando-se uma série de conclusões importantes que não dependem dos resultados estatísticos, fazendo correlações com conhecimentos prévios, literatura e entendimentos que já se tem acerca do sistema e que possam revelar soluções ou, no mínimo, ampliar o que se sabe com novos saberes acerca do processo estudado, suas variáveis e respostas.

### 5.13 VALIDAÇÃO FINAL

Após a realização das análises dos dados, brutos e estatísticos, com um modelo válido em mãos e os dados auxiliares de superfície de resposta, é possível obter a condição ótima, uma combinação de valores das variáveis independentes que são capazes de otimizar a resposta desejada dentro dos valores tidos como ótimos, seja para maximização ou minimização da resposta. Com os parâmetros em mãos a etapa subsequente é a validação final, a realização de um experimento dentro das condições avaliadas no estudo de planejamento experimental, com o uso dos parâmetros ótimos das variáveis independentes e com mensuração das mesmas respostas que fazem parte do estudo.

A realização da validação visa verificar experimentalmente se os valores preditos pelo modelo para as respostas nos valores ótimos de cada variável independente, são compatíveis com os valores reais obtidos com o experimento,

fornecendo uma medida de desvio e atestando ou não a adequação do modelo nesta verificação final. É esperado então que os valores reais obtidos estejam de acordo como os valores preditos, com proximidade adequada de acordo com os limites de predição do modelo.

Para isso, a fim de que se tenham medidas adequadas com dimensionamento da variabilidade, é necessário que se realize o experimento de validação nas mesmas condições experimentais, que se faça uso de valores das variáveis independentes dentro da faixa estudada e que se realizem ao menos duas medidas, para que seja possível dimensionar o erro, permitindo que os dados obtidos sejam comparados com o teórico.

Um modelo ajustado dentro de uma faixa determinada para os fatores, deve fornecer, se válido, capacidade preditiva das respostas para cada possível valor dos fatores, sempre dentro da faixa estudada. Não se pode esperar que seja possível a obtenção segura e estatisticamente válida de valores de resposta predito para variáveis independentes em faixas estranhas com relação as iniciais, pois o comportamento pode não ser o mesmo a partir de determinados níveis, esse tipo de extrapolação pode não ser adequada.

Em alguns casos, especialmente tratando-se de processos e substâncias químicas, certos níveis de operação são definidos em razões de segurança ou performance adequada, em caso de processos, e isso é definido no início da concepção do planejamento experimental, exatamente para que não se tenha o uso de faixas que comprometam os critérios de segurança operacional, irrealis ou muito amplas. Para isso realizam-se estudos de seleção de variáveis antes mesmo da escolha do planejamento adequado, com o objetivo de definir as variáveis significativas e também realizar possíveis ajuste das faixas de estudo.

Outro ponto diz respeito ao comportamento químico das substâncias que tem relação com concentração no meio e interações com outras substâncias ou condições experimentais como temperatura e pressão. Acima de determinado limite estipulado pela faixa de estudo, podem ocorrer interações e deslocamento de equilíbrio químico capaz de alterar drasticamente o comportamento do processo ou fenômeno estudado, que passa então a não ser mais compatível com o modelo gerado. Assim, é

necessário respeito aos parâmetros de contorno inicialmente impostos no estudo de planejamento experimental e limitações do modelo gerado.

Caso tudo seja feito de acordo, seguindo o modelo bem construído e válido, respeitando as condições sobre as quais ele foi construído e tendo em mente, mais uma vez, o bom senso do experimentador, é possível obter resultados satisfatórios e adequados com correspondência entre os valores preditos e os obtidos na prática real. Essa é uma etapa essencial que firma e corrobora os achados teóricos obtidos com a geração do modelo, capaz de finalizar questionamentos acerca da capacidade preditiva e de otimização da função gerada, bem como sua validade estatística e validade em experimentação real. Um modelo é continuamente validado com seu uso, assim, etapa de validação final atribui segurança ao modelo, no uso ao qual se destina, permitindo que ele seja constantemente usado e validado.

#### 5.14 USO DE SOFTWARES

Existe uma quantidade grande softwares estatísticos e matemáticos que permitem a construção de planejamentos experimentais, bem como análises de seus resultados com geração de modelo de regressão ou superfície de respostas. O uso de softwares computacionais facilitou o uso do planejamento experimental, dada a velocidade com que os cálculos e análises são performados. Os softwares mais comuns usados atualmente incluem SAS, JMP, Minitab, R, Design Expert, Protimiza e até mesmo Microsoft Excel.

A maioria dos softwares disponíveis não são focados em planejamento experimental, tendo uma série de pacotes para diversos usos, e oferecem uma grande liberdade ao usuário, não estabelecendo restrições quanto a aceitabilidade matemática ou estatística da operação efetuada com relação a metodologia de planejamento experimental, ou se o valor inserido faz sentido. Esse tipo de liberdade representa uma vantagem quando se tem certo domínio e entendimento das operações e processos estatísticos envolvidos na construção e cálculos, onde se pode testar novas possibilidades e realizar pesquisas mais aprofundadas sobre planejamento experimental, que fogem do uso padrão ou rotineiro da metodologia.

Qual o usuário do software não tem conhecimento suficiente para realizar estudos com maior profundidade e fazer uso da liberdade disponibilizada, acabam sujeitos a realizar operações que resultam na obtenção de resultados e valores equivocados que são posteriormente utilizados para fundamentar conclusões e decisões também equivocadas. Também existem casos de softwares muito simplistas que impedem análises mais robustas e detalhadas, sem muito recurso para que o experimentador possa analisar conjunto de dados mais complexos ou em estruturas de planejamentos personalizadas.

Assim, é necessário estar alerta e ciente de que certas operações referentes ao uso do planejamento experimental e sua análise de resultados deve ser bem compreendidas antes da aplicação real, para que se tenha controle do uso das ferramentas, ciente das limitações sobre o que pode ser feitos em softwares que dão grande liberdade ao usuário ou restringem demais as funcionalidades. Conhecer bem o planejamento experimental e as análises permitem ao usuário maior segurança na utilização de softwares diversos, especialmente quando se sabe utilizar bem o software e conhece as situações nas quais ele é mais indicado, se tem conhecimento de quais operações e qual a metodologia adequada para cada caso, o que pode ou não ser feito, quais valores são permitidos, quais as premissas e o que esperar do resultado final, sempre recorrendo ao bom senso e conhecimentos prévios.

Na maior parte dos softwares a construção é pouco intuitiva e ainda que se tenha uma construção guiada, não deve-se seguir receitas prontas sem total conhecimento das etapas que estarão sendo seguidas, para cada caso deve-se ter atenção aos detalhes, pois a mínima ação impulsiva pode terminar em resultados enganosos e custosos, pois tratando-se de planejamento de experimentos, tem-se uma quantidade grande de ensaios com recursos envolvidos, sejam reagentes, pessoal, maquinário, com impacto econômico e de tempo.

Ainda que se tenha atenção e conhecimento acerca da metodologia de Planejamento Experimental e das análises estatísticas, ainda assim se está sujeito ao uso de softwares ruins, que apresentam grandes falhas de cálculos automatizados por algoritmos mau construídos e que apresentam erros importantes ou mesmo softwares que se tornam inadequados pois exigem do usuário tempo e conhecimento matemático ou computacional complexo na construção de algoritmos particulares. Softwares ruins também podem apresentar variabilidade grande com relação aos

cálculos obtidos por outras fontes confiáveis, com operações incorretas e que nem sempre estão disponíveis ao usuário, por fazer parte de um algoritmo fechado. Neste caso também é importante conhecimento sobre a operação efetuada, para que se tenha noção do que se esperar, qualquer valor que fuja da expectativa deve ser tratado com desconfiança.

Antes de se escolher um software para uso no Planejamento Experimental, vale a pena buscar referências na literatura e com usuários frequentes da metodologia no ambiente acadêmico ou industrial, para que se tenha acesso a softwares confiáveis e tenham sido testados na prática com sucesso, por usuários qualificados. A escolha de software não deve ser negligenciada, pois pode colocar em risco todo o trabalho realizado, além de ocasionar a tomada de decisões equivocadas, quase sempre com resultados de impacto financeiro e de tempo muito importantes.

### 5.15 OUTROS ERROS FREQUENTES

Em qualquer tarefa dentro do planejamento experimental e sua execução se está sujeito a erros, especialmente em episódios de desatenção ou negligência. A fim de evitar boa parte dos erros a dica mais simples é estar atento, ter bom senso e ter uma atitude crítica na execução das tarefas. Dentre tantas possibilidades de erros e ações que comprometem os resultados e execução do planejamento experimental, a mais simples delas se relaciona com a inserção de dados errados nas planilhas e softwares utilizados.

Uma falha simples de desatenção como essa pode ocasionar a obtenção de resultados equivocados, comprometendo todo o estudo. É algo que pode ser evitado com a atenção constante e a verificação dos dados antes de efetuar qualquer operação. A atitude de parar, pensar e planejar cada ação dentro do planejamento experimental faz sentido e ajuda a evitar erros e o comprometimento do estudo.

A inserção de dados incorretos, a resposta de um determinado ensaio inseridas nos espaços correspondentes a um outro ensaio, por exemplo, pode ser evitada e corrigida facilmente quando percebida a tempo. No entanto, a depender do erro de inserção de dado e da etapa de execução do planejamento, pode significar

comprometimento de uma etapa crítica do trabalho, como por exemplo quando o erro diz respeito aos níveis de determinada variável independente na matriz do planejamento. Tal erro, se arrastado ao longo do experimento, pode ocasionar riscos de segurança, ou a completa invalidação do estudo.

Outra fonte comum de erros diz respeito a precisão de equipamentos utilizados, que podem não corresponder a expectativa de resultados dos experimentadores, não podendo medir variações da resposta correspondente aos efeitos provocados pelos níveis das variáveis independentes, ou seja, não há uma adequação da faixa estudada com a capacidade de medição dos equipamentos. Os equipamentos de medição devem ser selecionados logo no início dos preparativos para o planejamento experimental e deve-se verificar se ele é capaz de detectar variações na resposta de acordo com a variação dos níveis das variáveis independentes. Além disso, o valor obtido por um equipamento deve ser inserido integralmente nas planilhas de resultados, sem arredondamentos desnecessários e que possam comprometer a capacidade de distinguir valores ou detectar adequadamente as variações que ocorrem, ainda que sutis. Nesse sentido também é válido considerar o erro dos equipamentos de medição, para que não se considerem medidas dentro da escala de erro, e sua calibração e uso adequado, a fim de minimizar variabilidade nos resultados decorrente de ruídos e variáveis ocultas.

Por fim, tendo em vista as inúmeras possibilidades de erros, especialmente decorrente de desatenção, negligência ou não atribuição da importância devida a certos detalhes, é necessário que antes de se aventurar no planejamento experimental, seja feito um preparo cuidadoso das orientações iniciais, do comprometimento da equipe envolvida e do experimentador, com elevado critério de qualidade na execução de cada uma das etapas, incluindo as etapas iniciais de discussão e estudo prévio. Todo cuidado e atenção são importantes para garantir a execução com minimização dos erros humanos, minimização das fontes de variabilidade e funcionamento adequado dos equipamentos.

Com tudo rigorosamente ajustado, bem estruturado e bem executado, no pior dos casos uma quantidade de dados corretos coletados estará disponível para serem avaliados e várias conclusões ou ideias para novos estudos poderão ser geradas, sem comprometimento total do estudo e dos esforços aplicados.

## 6 CONCLUSÃO

O trabalho se propôs a apresentar e discutir tópicos de boas práticas na aplicação do planejamento experimental e, portanto, apresentaram-se cada um dos tópicos selecionados com uma abordagem especial voltada a casos da química, em ambiente industrial ou laboratorial. Discutiram-se os tópicos selecionados, trazendo considerações importantes e críticas úteis para experimentadores que buscam usar a metodologia em sua rotina ou em casos pontuais. Descreveram-se as situações de modo a tornar claro como elas impactam seriamente os resultados, com amparo de casos exemplos de uma abordagem adequada.

As informações reunidas, apesar de sua importância evidente, não são encontradas facilmente ou abordadas com clareza na literatura da área, estando, portanto, pouco acessível, em especial para aqueles que possuem pouca familiaridade ou contato inicial com o planejamento experimental. Desta forma, destaca-se a importância da existência de fontes de consulta com descrições das situações de impacto, com clareza a respeito de certos cuidados não tão óbvios e que apresente abordagens ideais em cada caso apresentado, a fim de que sejam evitados os erros que possam comprometer, parcial ou totalmente, os resultados obtidos e o trabalho realizado.

O conteúdo apresentado neste trabalho diz respeito a um conhecimento de boas práticas da aplicação do planejamento experimental essencial a qualquer um que se aventure nessa área, em especial aos químicos quimiométricos iniciantes, pois trata-se, em parte, de um conhecimento prático dificilmente descrito na literatura e que está, em muito, restrito a uma transmissão vocalizada em grupos de pesquisa e nas práticas da indústria.

Assim, de forma geral, o conteúdo apresentado tem sua utilidade clara nos estudos ou aplicações do planejamento experimental, principalmente por transmitir de difícil localização e de caráter crítico no uso adequado da metodologia, além de evitar que o uso dela seja rejeitado devido a experiências falhas potencialmente relacionadas aos tópicos relatados neste trabalho.

Por fim, o presente trabalho é fruto de um estudo amplo dos potenciais fontes de erros pouco descritas do uso do planejamento experimental, reforça os cuidados

já apresentados na literatura de forma simples e direta, e abre caminho para que sejam construídos mais materiais de suporte aos experimentadores, de modo que se tenha maior controle da experimentação e maior qualidade nos resultados obtidos.

## 7 REFERÊNCIAS

ANTONY, J. **Design of Experiments for Engineers and Scientists**. 2nd Edition, p.63-112. Elsevier Insights, 2014.

BARROS NETO, B. DE.; SCARMINIO, I. SPACINO.; BRUNS, R. EDWARD. **Como fazer experimentos pesquisa e desenvolvimeno na ciência e na indústria**. Editora da Unicamp, 2001.

BENEDETTI, B.; CAPONIGRO, V.; ARDINI, F. **Experimental Design Step by Step: A Practical Guide for Beginners**. Critical Reviews in Analytical Chemistry. Taylor and Francis Ltd., 2022.

BRERETON, R. G. **The evolution of chemometrics**. Analytical Methods, Issue 16, n. 5, 2013.

BRERETON, R. G. **Chemometrics : data analysis for the laboratory and chemical plant**. Chichester: Wiley, 2003.

BROWN, S. D. **The chemometrics revolution re-examined**. Journal of Chemometrics, v. 31, n. 1, 2017.

BROWN, S.; TAULER T.; WALCZAK, B. **Comprehensive Chemometrics: Chemical and Biochemical Data Analysis**. 2nd Edition. Elsevier, 2020.

DEMING, S. N.; MORGAN S. L. **Experimental design: A chemometric approach**. 'Data Handling in Science and Technology', 3rd Edition. Elsevier, 1988.

FERREIRA, M. M. C. et al. **Quimiometria I: calibração multivariada, um tutorial**. Química Nova, v. 22, n. 5, 1999.

GEMPERLINE, PAUL. **Practical guide to chemometrics**. 2nd Edition. Taylor & Francis, 2006.

GOURIEROUX, C. **Econometrics of Qualitative Dependent Variables**. 1st Edition. Cambridge University Press, 2000.

GUZZO DA SILVA, B.; FRATTINI FILETI, A. M.; PEREIRA TARANTO, O. **Drying of Brazilian Pepper-Tree Fruits ( *Schinus terebinthifolius* Raddi): Development of Classical Models and Artificial Neural Network Approach**. Chemical Engineering Communications, v. 202, n. 8, 2015.

HOPKE, P. K. **The evolution of chemometrics**. Analytica Chimica Acta, v. 500. Elsevier, 2003.

JOLLIFE, I. T.; CADIMA, J. **Principal component analysis: A review and recent developments**. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. Royal Society of London, 2016.

KATEMAN, G. **Evolutions in Chemometrics: Plenary Lecture**. Analyst, Issue 5, 1990.

MASSART, B. *et al.* **Handbook of Chemometrics and Qualimetrics: Part A**. 'Data Handling in Science and Technology', 1st Edition. Elsevier, 1997.

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 10th Edition. Wiley, 2020.

RODRIGUES, M. I.; IEMMA, A. F. **Experimental Design Process Optimization**. 1st Edition. Taylor & Francis, 2014.

ROSIPAL, R.; KRÄMER, N. **Overview and Recent Advances in Partial Least Squares**. In: Craig Saunders C.; Grobelnik M.; Gunn J.; Shawe-Taylor J. Subspace, Latent Structure and Feature Selection: Statistical and Optimization Perspectives Workshop. Springer-Verlag, 2006.

SANTOS, M. C. *et al.* **Chemometrics in analytical chemistry - An overview of applications from 2014 to 2018**. Eclética Química. Atlantis Livros, 2019.

TEÓFILO, R. F.; FERREIRA, M. M. C. **Quimiometria II: planilhas eletrônicas para cálculos de planejamentos experimentais, um tutorial**. Química Nova, v. 29, n. 2, 2006.

WIKSTRÖM, M.; SJÖSTRÖM, M. **Commercial spread optimization using experimental design and sensory data**. Journal of the American Oil Chemists' Society, v. 81, n. 7, 2004.

WOLD, S. **Chemometrics; what do we mean with it and what do we want from it**. Chemometric and Intelligent Laboratory Systems, 1995.

WU, C.F.; HAMADA, M. **Experiments : planning, analysis, and optimization**. 2nd Edition. Wiley, 2009.