Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Review

Design of Experiments Applications Review: Food Processing Engineering as an Example Field

Bilel Hadrich

Department of Chemical Engineering, College of Engineering, Imam Mohammad Ibn Saud Islamic University (IMSIU), 11432, Riyadh, Saudi Arabia
E-mail: bmhadrich@imamu.edu.sa

Received: 21 March 2025; **Revised:** 22 April 2025; **Accepted:** 6 May 2025

Abstract: The relationship between food processing engineering and the various properties and qualities of the final product is both evident and significant for industrial companies and customers alike. However, understanding the mechanisms underlying this relationship can be effectively achieved through a powerful and efficient mathematical and statistical methodology known as Design of Experiments (DoE), which aids in optimizing quality while minimizing cost. DoE encompasses various types, each offering distinct advantages. As a versatile tool, it is not only straightforward and practical but also highly insightful. This review begins with an overview of DoE to illustrate its applicability and relevance to the engineering field in general and to food processing engineering in particular. Its crucial roles are also discussed in both process scale-up and ensuring regulatory compliance and standards. Furthermore, this review presents a rundown of published research studies that have employed DoE, highlighting how this methodology has been applied to address various challenges in food processing engineering. These studies-like many others utilizing DoE-focus on investigating the relationship between specific food properties (considered dependent variables or responses) and the independent variables, which represent the tested factors. Finally, the future of DoE in food processing engineering is explored, with particular attention to the new opportunities enabled by emerging technologies such as Artificial Intelligence, Machine Learning, and real-time process monitoring.

***Keywords*:** food processing engineering, food quality, design of experiments, response surface methodology

MSC: 62K05, 62K15, 62K20

1. Introduction

Over the past several years, the application of Design of Experiments (DoE) has witnessed steady growth. This methodology has become increasingly significant across various disciplines, particularly in engineering, where it is applied extensively in its numerous subfields, including mechanical, civil, electrical, environmental, biological, chemical, and process engineering. The versatility and effectiveness of DoE in optimizing processes, improving product quality, and enhancing system performance-while reducing costs-have contributed to its widespread adoption. Food processing engineering is considered one of the most engineering-intensive fields for humanity, characterized by various manufacturing processes that can alter the appearance and properties of food. Indeed, it provides a wide range of products

DOI: https://doi.org/10.37256/cm.6320256818

derived from various natural raw materials for consumers. For example, the process of fermentation transforms raw milk into several final products such as cheese, yogurt, kefir, butter, …, etc. Typically, in such cases, the preferred final products vary from country to country and from culture to culture, even when the raw materials and end products are the same. This engineering challenge depends on numerous parameters that influence food properties specifically and food quality in general. For example, during food processing, several phenomena may occur that can easily and significantly affect the properties and overall appearance of the treated food. As we know, the quality of the final product is a crucial and complex concept which also varies across countries and cultures. For this reason, quality control during processing should not be overlooked by the engineers and the research and development departments within factories. This vital and challenging task can be facilitated through the use of DoE. Numerous studies and reviews highlight the importance of this approach in addressing various problems using different types of DoE. The present review delves into the multifaceted applications of DoE, highlighting its significance, methodologies, and contributions to engineering innovation in general and food processing engineering in particular.

This review begins with a comprehensive description of DoE from multiple perspectives, including its historical development, core concepts, applications, objectives, types, mathematical foundations, regression model fitting, and definitions of key variables, including independent and dependent variables as well as other relevant parameters.

Once implemented and validated, DoE serves as a valuable platform for addressing two critical challenges in food processing engineering: process scale-up and compliance with regulatory standards. Accordingly, this review also discusses how DoE has dealt with these challenges.

The paper then examines key studies in the field of food processing engineering, highlighting the role of DoE in optimizing processing conditions while maintaining food quality. It further explores the future of DoE in food engineering, particularly in the context of emerging technologies such as Artificial Intelligence (AI), Machine Learning (ML), and real-time process monitoring. The review concludes by emphasizing the growing importance and impact of DoE in driving advancements in modern food engineering practices.

2. History of DoE

The first known adopter of DoE methodology was Sir Ronald A. Fisher, a statistician and geneticist. Fisher introduced the foundational principles of DoE in the 1920s while working at the Rothamsted Experimental Station, an agricultural research institution in England. Fisher's pioneering work was initially focused on agriculture, where he used statistical methods to optimize crop yields by systematically testing the effects of different variables (e.g., fertilizers, soil types, and crop varieties). He developed key concepts such as randomization to reduce bias in experiments, replication to ensure results were reliable, and blocking to account for variability in experimental conditions [1]. Fisher systematically integrated statistical reasoning and principles into the design of experimental investigations, pioneering concepts such as factorial design and analysis of variance. His two seminal books, with their most recent editions being Fisher [2–4] significantly influenced the application of statistics, particularly in agriculture and related life sciences [1, 5].

Next, the statistical design applications in industrial settings began in the 1930s, but the industrial era truly gained momentum with the introduction of Response Surface Methodology (RSM) by Box and Wilson in 1951 [1, 6]. They identified and highlighted two key distinctions between industrial and agricultural experiments: (1) immediacy of observing results: the response variable in industrial experiments is often observed almost immediately, and (2) sequentially of experimental design: crucial insights can be obtained from a small number of experimental runs, enabling iterative planning for subsequent experiments. Over the next three decades, RSM and other experimental design techniques gained significant traction, particularly within the chemical and process industries. These methods were predominantly applied in research and development settings, where the need to optimize processes and improve efficiency was paramount. The adoption of these techniques allowed researchers to systematically explore the relationships between variables, identify optimal operating conditions, and enhance product quality. This widespread use also highlighted the versatility of RSM in addressing complex, real-world industrial challenges, further solidifying its role as a critical tool in experimental design. Box played a pivotal role in advancing statistical design. However, its adoption in the manufacturing process level remained limited due to insufficient training in statistical methods for engineers and process specialists as well as the lack of accessible computing tools and user-friendly statistical software. Moreover, during this same period, Kiefer and Wolfowitz [7] and Kiefer [8] introduced a formal framework for selecting designs based on objective optimality criteria, aiming to maximize the precision of model parameter estimates [7, 8]. However, practical implementation was initially limited because of the absence of adequate computing tools. Over the past 25 years, advancements in algorithms and computing technology have greatly enhanced the feasibility and application of optimal design techniques.

The growing focus of Western industries on quality improvement, which emerged in the late 1970s, marked the beginning of the third era of statistical design: the quality era. Taguchi [9–11] popularized statistical design with his robust parameter design methodology. He focused on reducing process sensitivity to uncontrollable factors, minimizing product variation, and optimizing process variables. Despite its widespread adoption, Taguchi's methods faced criticism for statistical shortcomings, leading to extensive peer reviews that identified issues in his experimental strategies. This controversy resulted in the broader adoption of experimental design in industries like automotive, aerospace, and electronics, and spurred the development of alternative approaches that addressed Taguchi's engineering goals more effectively.

Another positive and significant outcome of the Taguchi controversy is the establishment of the fourth era, known as the modern era (1990s-present). This era saw renewed interest in statistical design, driven by advancements in statistical software, algorithms, and formal education. Designed experiments are now integral to engineering curricula and are applied across diverse industries and sectors, including business, government, and nonprofits. Multivariable Testing (MVT), a term for factorial designs, has gained recognition for its success in optimizing complex processes. Statistical experimental design has expanded far beyond its agricultural roots, becoming an essential tool in science, engineering, and industry, with continuous innovation ensuring its relevance in addressing modern challenges.

3. Concept and objectives

The DoE is a powerful statistical methodology widely used to explore and optimize complex systems across various fields of engineering. Moreover, DOE is a systematic approach to planning, conducting, analyzing, and interpreting controlled tests to evaluate the factors that influence a particular outcome. Originating from statistical principles, DoE has transcended its foundational disciplines to become a cornerstone methodology across various engineering fields. The ever-growing complexity of engineering problems demands robust methodologies for optimization, cost-effective problem-solving, and informed decision-making. DoE provides engineers with a structured framework to investigate the relationships between inputs and outputs in systems, leading to enhanced performance, quality improvement, and cost reduction.

3.1 *Dependent and independent variables, and coded variables*

Any target property, quantity, or yield can be defined as a dependent variable, commonly referred to as the response (denoted by y). On the other hand, any process or physicochemical parameter is considered an independent variable, commonly referred to as a factor-with the exception of mixture design, where variables are governed by the fundamental constraint that the component proportions must sum to one. Factors are generally classified into two main types: environmental factors, which can be deliberately manipulated and fixed during experimentation, and noise factors, which are difficult or impossible to control but may still influence the response.

For each environmental factor, the experimental range should be clearly defined in advance to ensure meaningful and reproducible results.

For example, in food processing engineering field, several responses can be considered, such as fermentation yield, food color, moisture content, ash content, sugar content (in drying as well as several other processes), ..., etc. Additionally, there are various factors, such as temperature and initial pH, which are commonly used in various processes, as well as the initial concentration of solvent in the extraction process, oxygen concentration in fermentation process, and others.

All factors have an experimental range defined by the operator. The lower, center, and higher levels can be easily determined within this range. Since the factors may differ in type, dimensions, and units, and their experimental ranges may vary in scale, all DoE designs are constructed using coded values. Specifically, the lower, center, and higher levels of each factor are mathematically converted to -1, 0, and +1, respectively. These coded variables are denoted by x_i for the factor i.

The design matrix is always constructed from a number of combinations between all factors' levels. Types of design matrices are presented in many books, such as those by [1, 12].

3.2 *Mathematical modeling of the response*

When no prior information is available about the function linking the response to the factors, an assumed general evolutionary law is proposed, formulated as follows (Equation (1)):

$$y = f(x_1, x_2, x_3, \ldots\ldots, x_k) + e = \widehat{y} + e \tag{1}$$

This function is highly general, and it is common practice to use a truncated Taylor-Mac Laurin expansion as an approximation. If the derivatives can be treated as constants, this expansion simplifies to a polynomial of varying degrees (Equation (2)):

$$\widehat{y} = a_0 + \sum_{i=1}^{k} a_i \cdot x_i + \sum_{i=1}^{k} \sum_{j>i}^{k} a_{ij} \cdot x_i \cdot x_j \sum_{i=1}^{k} \sum_{j>i}^{k} \sum_{l>j>i}^{k} a_{ijl} \cdot x_i \cdot x_j \cdot x_l + \ldots\ldots + \sum_{i=1}^{k} a_{ii} \cdot x_i^2 \tag{2}$$

where y represents the vector of experimental response values in the standard order of runs (experimental conditions); $\widehat{y}$ represents the vector of predicted response values in the standard order of runs, calculated using Equation (2) after determination of model's coefficients; e: vector of the deviation between the model results and the experimental results; x_i, x_j, x_l represent the coded levels of factors i, j, l, respectively; a_0, a_i, a_{ij}, a_{ijl} and a_{ii} are the coefficients of the model representing the intercept, linear term coefficients, first-order interaction coefficients, second-order interaction coefficients, and quadratic term coefficients, respectively; and k: the number of factors.

The established models are predictive models applicable within the specified field of study, which must always be defined at the outset. They are not based on theoretical physicochemical or mechanical laws. In some rare instances, however, known theoretical physical laws may be used [1, 12].

Each experimental point provides a value for the response. This response is modeled by a polynomial (Equation (2)), the coefficients of which are the unknowns that need to be determined. At the end of DoE, we have a system of n equations (if there are n trials) with p unknowns (if there are p coefficients in the chosen model *a priori*). This system can be written simply in matrix notation (Equation (3)) [1, 12].

$$y = X \cdot a + e \tag{3}$$

where X is the matrix for calculating the coefficients or model matrix, which depends on the experimental points chosen to conduct the plan and the assumed model while a is the vector of coefficients.

This model matrix X is derived from the design matrix, which presents the levels of each factor for each experimental condition. In the design matrix, the number of columns corresponds to the number of factors. However, the model matrix has a number of columns equal to the number of terms in the model, as it is used to determine all the coefficients of the model.

This resolution can only be achieved by using a regression method. It is based on the least-squares optimization criterion. Thus, the estimates of the coefficients (denoted by $\widehat{a}$) are obtained (Equation (4)) [1, 12]:

$$\widehat{a} = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot y \tag{4}$$

where X^T is the transpose of X.

There are many software programs that perform this calculation and directly provide the values of the coefficients.

3.3 *Types of DoE*

The process begins with selecting a design type, which is primarily determined by the goal of the experiment and the number of factors involved. In this step, several types of designs can be classified into four principal categories [1, 12].

3.3.1 *Factorial designs*

This type of design is considered one of the simplest and most comprehensible, both in terms of the experimental design matrix and the characteristic model. In fact, the factors can be represented with at least two levels, and the number of factors, k, must be greater than or equal to 2. There are three main subcategories in this design. For each case, the general model is presented.

3.3.1.1 *Full factorial design 2^k*

The Full Factorial Design 2^k (FFD 2^k) is a widely used experimental design in which all possible combinations of factors and their levels are tested. This approach enables the investigation of both main effects of individual factors and interaction effects among them, making it a powerful tool for exploring the influence of multiple variables on a response in a systematic and cost-effective manner.

The total number of required experiments is determined by the formula $n = 2^k$, where k is the number of factors, each considered at two levels. While this design is comprehensive and provides detailed insights, it becomes less practical when dealing with a large number of factors due to the exponential increase in the number of required experimental trials. Therefore, it is most suitable for studies involving a limited number of factors-typically not exceeding four (2^4 = 16 runs) or five (2^5 = 32 runs)-which allows for straightforward implementation and interpretation.

To further enhance the design, center points-where all factors are set at their mid-levels-can be added. These points are useful for assessing the **linearity** of the system's response. The inclusion of center points also applies to related designs such as Fractional Factorial Design 2^{k-p} (see Section 3.3.1.2) and Plackett-Burman Design (see Section 3.3.1.3), where they serve a similar purpose in evaluating model adequacy and detecting curvature in the response surface.

In this case, the model assumes a linear relationship between the response and the factors. As a result, only linear terms from Equation (2) are retained, while quadratic terms are excluded. The mathematical representation of this model is presented in Equation (5), which reflects the core assumption of linearity in this type of design.

$$\widehat{y} = a_0 + \sum_{i=1}^{k} a_i \cdot x_i + \sum_{i=1}^{k} \sum_{j>i}^{k} a_{ij} \cdot x_i \cdot x_j + \sum_{i=1}^{k} \sum_{j>i}^{k} \sum_{l>j>i}^{k} a_{ijl} \cdot x_i \cdot x_j \cdot x_l + \ldots\ldots \tag{5}$$

3.3.1.2 *Fractional factorial design 2^{k-p}*

When the number of factors increases to five or more-resulting in a rapid growth in the required number of experimental runs-the study can be efficiently conducted using either the Fractional Factorial Design 2^{k-p} (FFD 2^{k-p}) or the Plackett-Burman Design, both of which are designed to reduce the number of runs while retaining meaningful insights.

Specifically, in FFD 2^{k-p}, the number of experimental runs required for a Full Factorial Design ($n = 2^k$) is reduced by a factor of 2^p, where p is the number of fractions used. This involves selecting a representative subset of the full design matrix, significantly reducing the experimental effort while preserving critical information about the study.

Although this approach is cost-effective, it may limit the ability to detect or distinguish certain interaction effects, depending on the resolution of the design (e.g., Resolution III, IV, V, etc.). Fundamentally, it is a streamlined version of the full factorial design aimed at reducing the number of experimental runs. In such a design, some columns originally assigned to specific factors may simultaneously represent interactions between multiple factors, depending on the selected resolution level.

The general notation $n = 2^{k-p}$ can also be applied analogously to calculate the number of runs. The construction of interaction terms follows the specified resolution, ensuring a balance between model complexity and experimental efficiency. The resulting model structure is similar to that of Equation (5), with minor modifications reflected in Equation (6):

$$\widehat{y} = \ell_0 + \sum_{i=1}^{k} \ell_i \cdot x_i + \sum_{i=1}^{k} \sum_{j>i}^{k} \ell_{ij} \cdot x_i \cdot x_j \tag{6}$$

where ℓ_0 is the intercept of the model, which includes two coefficients: the usual intercept, as in a Full Factorial Design, and the coefficient of the highest possible interaction; ℓ_i represents the coefficients of the model which include at least two components (depending on the resolution): a coefficient of main factor (a_i) in the case of a Full Factorial Design, and at least a coefficient of one simple interaction (between two factors, a_{ij}).

Generally, this type of design is not used in isolation. Instead, it is employed as an initial step (not the key step) to identify the most important factors for a comprehensive study conducted in a second and final step. For this reason, many studies adopt this approach, using the Fractional Factorial Design in the first step, followed by one of the following designs in the second step: Full Factorial Design or Response Surface Methodology.

3.3.1.3 *Plackett-Burman design*

The Plackett-Burman Design is considered a special type of Fractional Factorial Design, as it follows a Resolution III structure and involves factors at **two levels**. However, unlike traditional fractional factorial design, the number of runs in Plackett-Burman Design does not adhere to the formula $n = 2^{k-p}$. Instead, the number of runs is typically does not exceed the number of factors plus one.

For example, one of the most commonly used Plackett-Burman Design matrices includes up to 11 columns (suitable for testing a maximum of 11 factors) and 12 rows, corresponding to 12 experimental conditions. This design is particularly useful for efficiently exploring high-dimensional experimental spaces by testing subsets of factor combinations under a minimum number of conditions. In other words, it is primarily employed for screening purposes, helping to identify the most influential factors among many.

The model used in Plackett-Burman Design is quite simple and is represented by Equation (7):

$$\widehat{y} = \ell_0 + \sum_{i=1}^{k} \ell_i \cdot x_i \tag{7}$$

Similar to the Fractional Factorial design 2^{k-p}, this design is not used on its own.

3.3.2 *Response surface methodology designs*

Response Surface Methodology (RSM) is a statistical technique designed to efficiently optimize processes involving multiple variables while examining their interactions. Its main advantage is the significant reduction in the amount of data required for evaluation, analysis, and optimization, thereby minimizing the total number of experiments needed while maintaining precision in exploring the response surface. Among RSM designs, Central Composite Design, Box-Behnken Design, and Doehlert Design are the most commonly preferred by researchers. In fact, the Central Composite Design offers users the possibility of using an existing factorial design (Full Factorial Design 2^k or Fractional Factorial Design 2^{k-p}) along with its results and adding a number of points called star points. These star points have levels that depend on α, where $\alpha \geq 1$, depending on the selected design and mathematical criterion. In this case, factors can operate at three levels (-1, 0, and +1; if $\alpha = 1$; where the points configuration is called cubic) or five levels ($-\alpha$, -1, 0, +1, and $+\alpha$; if $\alpha > 1$; here the points configuration is referred to as spherical). However, Box-Behnken Design follows a different logic in configuring the positions of points and their levels, which are fixed at three for all factors (-1, 0, +1). In this case, the configuration is spherical but uses a different point arrangement compared to the Central Composite Design with $\alpha = 1$. Moreover, while the Central Composite Design features two main point configurations-cubic and spherical-and the Box-Behnken Design also exhibits a spherical configuration, the Doehlert Design presents a spherical configuration with a unique logic. It adopts a hexagonal structure, offering the possibility of working within the same design to represent the same study using some factors with three levels and others with five levels. This is considered a distinct configuration compared to other designs.

Regardless of the chosen design, the model used in the response surface methodology includes linear terms, quadratic terms, and simple interaction terms (Equation (8)).

$$\widehat{y} = a_0 + \sum_{i=1}^{k} a_i \cdot x_i + \sum_{i=1}^{k} \sum_{j>i}^{k} a_{ij} \cdot x_i \cdot x_j + \sum_{i=1}^{k} a_{ii} \cdot x_i^2 \tag{8}$$

Generally, the number of factors is not very high because the corresponding number of runs would increase rapidly.

3.3.3 *Mixture design*

Mixture experiments are a type of response surface experiment where the product being studied consists of several components or ingredients. These designs are particularly useful in industrial settings, as many product development processes involve creating mixtures. In such DoE, the response depends on the proportions of each ingredient in the mixture, subject to a fundamental constraint of mixtures: the total amount of all components must sum to a fixed value, typically 1, such that $0 \leq x_i \leq 1$, and $\sum_{i=1}^{k} x_i = 1$. Unlike factorial designs, where the response varies based on the amount of each factor, mixture experiments focus on the relative proportions.

Three commonly used designs for mixture experiments are the simplex centroid design, simplex lattice design, and extreme vertices design. In this context, three types of models are employed: the first-degree model (Equation (9)), the second-degree model (Equation (10)), and the restricted third-degree model (Equation (11))-the latter being a specialized design for mixture plans, all formulated without an intercept [1, 12].

The first-degree model (Equation (9)) is derived from the simplest form of Equation (5), presented without interactions terms. This form represents the basic multi-linear model of the full factorial design 2^k, adapted using the fundamental constraint of mixtures. This transformation is achieved through a change of variables. As a result, each coefficient in the first-degree mixture model is expressed as the sum of the classical intercept a_0 from the FFD 2^k model and the corresponding coefficient a_i associated with the component x_i. A similar approach is applied to the basic multi-quadratic model of the response surface methodology (Equation (8), without intercept) to derive the second-degree model (Equation (10)), and to the third-degree polynomial model (special case for the mixture designs) to obtain the complete cubic model. This model is subsequently reduced to the restricted cubic model, as presented in Equation (11) [1, 12].

$$\widehat{y} = \sum_{i=1}^{k} b_i \cdot x_i \tag{9}$$

$$\widehat{y} = \sum_{i=1}^{k} b_i \cdot x_i + \sum_{i=1}^{k} \sum_{j>i}^{k} b_{ij} \cdot x_i \cdot x_j \tag{10}$$

$$\widehat{y} = \sum_{i=1}^{k} b_i \cdot x_i + \sum_{i=1}^{k} \sum_{j>i}^{k} b_{ij} \cdot x_i \cdot x_j + \sum_{i=1}^{k} \sum_{j>i}^{k} \sum_{l>j>i}^{k} b_{ijl} \cdot x_i \cdot x_j \cdot x_l \tag{11}$$

where b_i, b_{ij}, b_{ijl} are the coefficients of the model, x_i, x_j, and x_l are the fractions of components i, j, and l, respectively; and k represents the number of components.

In this design, studies can be conducted using a combination of factorial and mixture designs, where some factors are defined as component percentages (their levels characterized by $0 \le x_i \le 1$, and $\sum_{i=1}^{k} x_i = 1$), while others are defined as independent factors representing process variables. This distinctive approach is not feasible with other types of designs.

3.3.4 *Taguchi design*

A Taguchi design is an experimental approach aimed at selecting products or processes that perform consistently under real-world operating conditions. It recognizes that not all sources of variability-known as noise factors, can be controlled. To address this, Taguchi design focuses on identifying controllable factors (control factors) that minimize the impact of external variation (noise factors). By intentionally introducing variability through noise factors during experimentation, the method helps to determine the optimal settings for control factors that enhance robustness, making the product or process less sensitive to the noise. The result is more consistent process outputs and products with reliable performance across varying environments. This methodology is known as the Robust Parameter Design (RPD), and it is often used within the broader context of "experimental optimization".

In addition, Taguchi proposed summarizing the data from a crossed array experiment using two key statistics: (1) the average response for each inner array setting across all runs in the outer array and, (2) a composite measure known as the Signal-to-Noise (S/N) ratio, which combines information about both the mean and variance. The S/N ratio is structured such that maximizing it helps to minimize the variability introduced by uncontrollable noise factors. Subsequently, an analysis is carried out to identify the settings of the controllable factors that bring the mean response as close as possible to the desired target, and yield the highest S/N ratio. However, signal-to-noise ratios have certain limitations. They can confound location (mean) and dispersion (variance) effects, and they often fail to achieve the main goal of RPD-minimizing variability transmitted from noise factors.

3.4 *Experimental, modeling, and analysis protocols*

After defining the experimental range of each factor and verifying their independence-except in the case of mixture designs-the operator selects the appropriate experimental design based on the study's objectives. Once these steps are completed, the corresponding matrix of experiments is generated using specialized statistical software such as STATISTICA, Minitab, Design-Expert, JMP, or similar tools.

Each experimental condition is then conducted in triplicate to ensure the accuracy and reproducibility. Failed experiments during the setup phase can negatively impact the final modeling results and, consequently, the overall performance of the DoE approach. In such cases, the operator should repeat the failed experiments to minimize experimental errors and ensure the highest possible reliability of the DoE results.

The resulting data are then entered into the same statistical software for: analysis, model development, and an initial statistical evaluation of the model's goodness of fit. This evaluation typically includes metrics such as the coefficient of determination (R^2) which reflects how well the model explains the variability in the response; the adjusted R^2 which accounts for the number of predictors used; and p-values which assess the statistical significance of both the overall model and individual model terms. These p-values are obtained through Student's t-test and Analysis of Variance (ANOVA)-

also known as the F-test or Fisher test. Together, these indicators help to determine whether the model is both statistically valid and practically meaningful for prediction and optimization purposes.

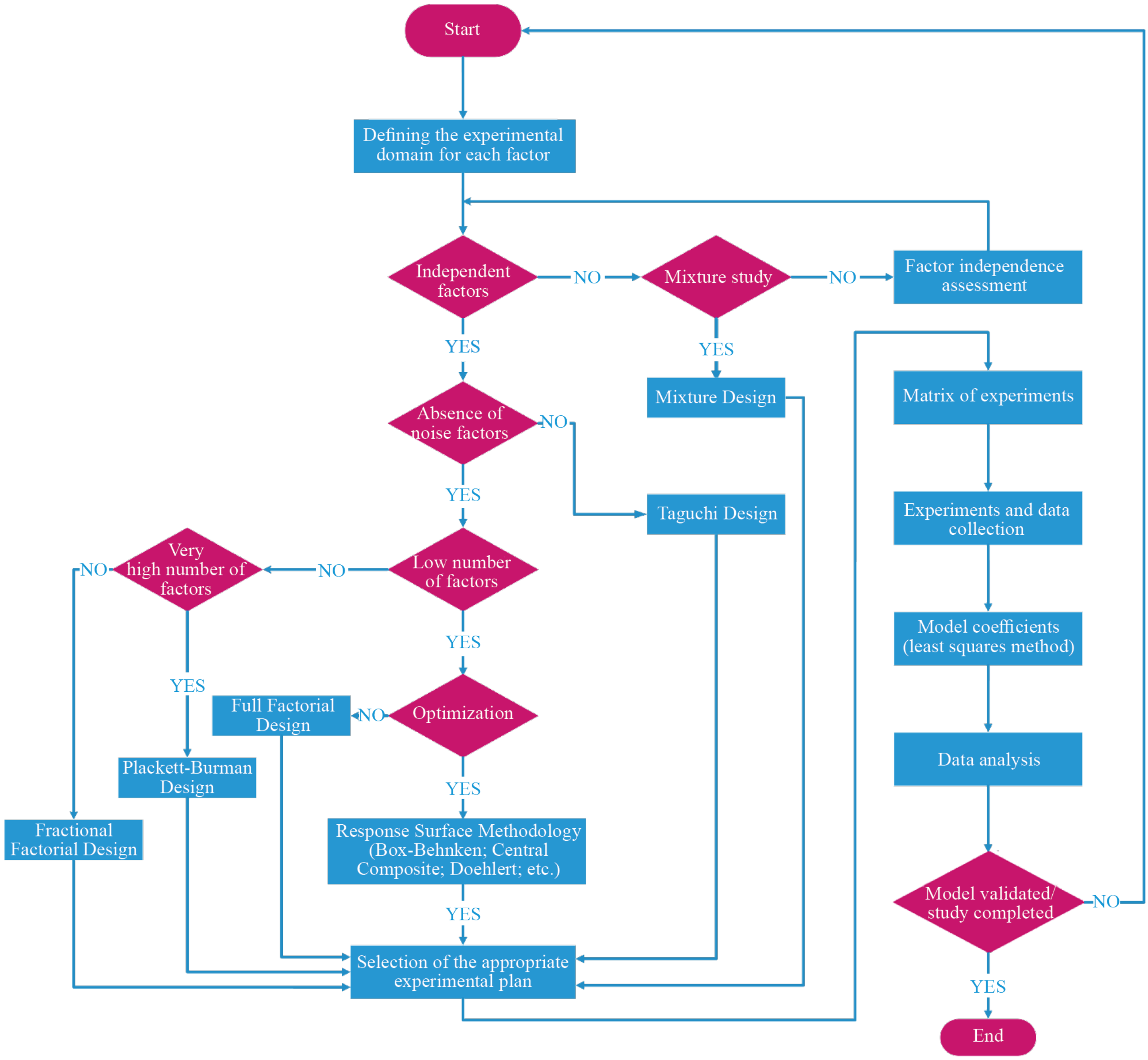

Figure 1. Design of Experiments (DoE) flowchart: guided by study type, objective, and factor type/number

Furthermore, the optimization module within the software is particularly valuable for identifying the optimal value(s) of the response variable(s) and determining the corresponding experimental levels of the tested factors. In most cases, this predicted optimum is validated through additional experiments trials. The discrepancy between the predicted and observed values is typically minimal, reinforcing the reliability and robustness of the developed model.

However, if a significant discrepancy is observed, it may indicate that one or more critical steps or foundational rules in the DoE process were overlooked during the initial setup. In such cases, it is advisable for the operator to re-evaluate the entire process and, if necessary, re-conduct the whole the experiment to ensure the validity of the results.

Furthermore, the software can generate a variety of graphical outputs, including factor effect plots, 2D and 3D predicted surface response graphs (often with experimental values overlaid to visualize differences), and optimization visualizations. Based on these analyses, researchers can decide whether to retain or revise the proposed design and model.

All of these software programs offer comprehensive features to support DoE, though they may differ slightly in terms of interface, output presentation, and user preferences. Such differences may influence the user's choices depending on their specific needs and expectations regarding data visualization and interpretation.

In food engineering, the use of such software is particularly advantageous, as research often requires simultaneous optimization of multiple response variables. This approach is commonly applied with various types of DoE methodologies [13–22]. These tools facilitate the identification of one or more optimal combinations of factor levels that yield the best outcomes across all studied responses.

Figure 1 presents a flowchart illustrating the general procedure for applying DoE to any type of study and objective, regardless of the type or number of factors involved. It serves as a concluding diagram that summarizes all the essential steps to be followed for any DoE approach.

3.5 *Scaling-up challenges of DoE*

In a confirmation experiment, the goal is usually to verify whether the system works as expected based on theory or past experience.

Another common situation for confirmation experiments occurs when scaling up a new manufacturing process to full-scale production. Based on results obtained through DoE at flask-scale or laboratory-scale experiments, the confirmation experiment checks if the same factors and settings used during development are still effective for the full-scale process [1].

DoE offers several techniques that help to minimize variability, identify scale-sensitive factors, and support the development of robust, scalable processes. Table 1 outlines key examples.

Table 1. Examples of DoE techniques for mitigating scale-up challenges

DoE technique	Scale-up benefit
Blocking	Isolates variability due to scale or batch
Randomization	Reduces bias from uncontrolled conditions
Replication	Quantifies variability across scales
Factorial design	Identifies key factors and interactions
RSM	Optimizes process under non-linear conditions
Taguchi	Minimizes the impact of noise factors
Including scale as factor	Directly assesses scale impact
Center points	Detects curvature in response
Mixture design	Optimizes component proportions at different scales

3.6 *The role of DoE in meeting regulatory compliance and standards*

Several peer-reviewed studies highlight the crucial role of DoE in meeting regulatory compliance requirements and standards set by agencies such as the U.S. Food and Drug Administration (FDA) and the European Food Safety Authority (EFSA). For example, Savic et al. discuss the transition from classical DoE to Quality by Design (QbD), emphasizing its alignment with evolving regulatory expectations [23]. Similarly, Okhio-Seaman explores the use of DoE in supporting process validation and quality control in accordance with FDA guidelines [24]. In the same context, Taylor et al. present an integrated modeling approach using DoE to strengthen various stages of process validation, fully aligned with QbD milestones [25]. Additionally, Passerine and Breitkreitz examine the application of DoE within the Analytical QbD framework, underscoring its value in developing robust and regulatorily compliant chromatographic methods [26]. Collectively, these studies demonstrate DoE's pivotal role in enhancing process understanding, risk management, and validation in highly regulated environments.

More specifically, within the domain of food processing engineering, DoE serves as a critical tool not only for process optimization but also for ensuring regulatory compliance. For instance, Fidaleo illustrates how the integration of DoE with functional data analysis can enhance the performance of batch processes-such as hazelnut-and-cocoa paste milling-while concurrently satisfying regulatory expectations for process validation [27]. Similarly, Benetti and Benetti highlight the contribution of DoE in aligning production practices with internationally recognized food safety standards, including HACCP and ISO/FSSC 22000, both endorsed by regulatory authorities such as the FDA and EFSA [28].

Beyond regulatory considerations, a substantial body of research has focused on evaluating the nutritional, biological, pharmaceutical, and medicinal attributes of final products. These studies typically employ a diverse array of standardized tests to validate such properties, thereby reinforcing the scientific rigor of product development efforts [29–35].

Collectively, these findings underscore the indispensable role of DoE in advancing food processing systems that are efficient, scientifically robust, safe, and fully compliant with global regulatory frameworks.

4. Importance of utilizing DoE in food processing engineering

A large number of studies are continuously published, each involving different types of DoE depending on the issue being addressed. Table 2 presents twenty-three examples, classified according to the type of DoE, with three to five examples per category, highlighting the most commonly used ones.

Table 2. Examples of studies utilizing DoE in various food processing engineering applications

No.	Process	Factors	Response(s)	Main results	Ref.
		Full factorial design 2^k			
1	Extraction of reducing sugars	2 factors: enzyme and substrate concentrations	Reducing sugars concentrations	Positive influence of substrate concentration*	[36]
2	Extraction of polyphenols	3 factors: temperature; solid-liquid ratio; ultrasonic power	Yield of extraction	Positive influence of temperature*	[37]
3	Atlantic salmon smoltification and post transfer	3 factors: trace minerals; vitamins; amino acids	Atlantic salmon smoltification and post transfer performance	Negative influence of relatively high dietary vitamins supplementation levels, and positive effects of amino acids*	[13]
		Fractional factorial design 2^{k-p}			
4	Enzymatic synthesis of lipophilic piperic acid esters by immobilized	5 factors: molar ratio piperic acid/butanol (A); temperature (B); enzyme quantity (C); dichloromethane (DCM) quantity (D); agitation (E)	Reaction yield (%)	Selection of only two factors out of five. Negative influence of A = BD = CE**; Positive influence of D = AB***	[38]
5	Production of lipase by fermentation	4 factors: corn steep liquor (A); Mozzarella cheese whey (B) and corn oil (C); pH (D)	Concentration of produced lipase and productivity	Selection of only two factors out of four positive and non-significative effect of A = BCD negative effect of B = ACD***	[14]
6	Degumming, neutralization and bleaching on lampante virgin olive oil's quality, using four clays	7 factors: phosphoric acid (A); ratio acid/oil (B); degumming contact time (C); neutralizing agent excess (D); earth clay dosage (E); contact time (F); temperature (G)	Several responses: residual phospholipids; free fatty acid; trans C18 : 2; oxidation stability; Hue angle; Chroma; and lightness	Different results/influences function of the used clay	[15]

Table 2. (cont.)

No.	Process	Factors	Response(s)	Main results	Ref.
		Plackett-Burman design			
7	β-carotene production from agro-industrial by-products	11 factors: sucrose (A); lactose (B); peptone (C); beef extract (D); NaCl (E); $MgSO_4$ (F); KH_2PO_4(G); pH (H); inoculum size (J); agitation (K); dummy 1 (L)	Produced β-carotene	Only three factors affect the response: A, C, H***	[39]
8	Fermentation	9 factors: incubation time (A); inoculum size (B); pH (C); starch (D); peptone (E); yeast extract (F); $MgSO_4$, $7H_2O$ (G); NaCl (H); $CaCl_2$ (J)	Alpha amylase activity (U/mL/min)	Only three factors affect the response: A, C, J***	[40]
9	Fermentation	14 factors: incubation time (A); Temperature (B); pH (C); agitation (D); DO (E); inoculum size (F); inoculum age (G); $MgSO_4$ (H); L-cysteine (J); KH_2PO_4 (K); $CaCl_2$ (L); K_2PHO_4 (M); Corn steep liquor (N); lactose (O)	Lactase activity (U/mL)	Only 6 factors affect the response**	[41]
		Response surface methodology: Central composite design			
10	Enzymatic synthesis of lipophilic piperic acid esters by Immobilized	2 factors: molar ratio piperic acid/butanol (A), dichloromethane (DCM) quantity (B)	Enzymatic reaction conversion yield (%)	Negative influence of A*; And positive influence of B***; Significant impact of B^{2}***; Reaction yield $92.21 \pm 4.53\%$	[38]
11	Production of lipase by fermentation	2 factors: corn steep liquor (A); Mozzarella cheese whey (B)	Concentration of produced lipase and productivity	Maximization of the lipase production and productivity***	[14]
12	β-carotene production from agro-industrial by-products	3 factors: sucrose (A); peptone (C); and pH (H)	β-carotene concentration	Only factors C and H affect the response linearly***; All three factors affect the response quadratically***; Optimization of β-carotene concentration***	[39]
13	Fermentation	3 factors: incubation time (A); pH (C); $CaCl_2$ (J)	Alpha amylase activity (U/mL/min)	Only one factor affects the response: A, both linearly and quadratically***; Optimization of alpha amylase activity***	[40]
14	Production of mixed juice powder by drying	3 factors: foaming agent concentration (A); drying temperature (B); foam thickness (C)	Juice powder solubility	Optimization of the powder properties***	[42]
		Response surface methodology: Box-behnken design			
15	Extraction of essential oils	3 factors: time, power, temperature	Essential oils extraction yield	Positive influences of all factors***; All interactions involving factor power are significative*; The quadratic term of factor; Time is significant*; Extraction yield 3.093 g/100 g d.m.	[43]
16	Fabrication of gluten-free bread's formulation based on chickpea, carob and rice flours	3 factors: water (A); fermentation time (B); flour percentage (C)	Several essential properties of a sponge cake	Influence of each component on the properties of sponge cakes***; Optimization of a sponge cake based on customer requirements***	[16]
17	Production of date paste for sustainable food systems	3 factors: hydration time (A); drying time (B); drying temperature (C)	Four technico-functional properties: firmness; moisture content; water activity; color parameters: L^*; a^*; b^*	Several types of influences were carried out; Optimization of the properties of the final product***	[17]

Table 2. (cont.)

No.	Process	Factors	Response(s)	Main results	Ref.
		Mixture design			
18	Fabrication of gluten-free sponge cake fortified with whey protein concentrate	3 factors: maize (A); rice (B); whey protein concentrate (C)	Several essential properties of a sponge cake	All components influence the properties of sponge cakes***; Optimization of the sponge cake composition based on customer requirements***	[18]
19	Production of extrudate food with Mango by-products (*Mangifera indica*)	3 factors: white corn flour (A), mango peel flour (B) and mango kernel flour (C)	Several essential properties of the final product	Influence of each component on the properties of the final product***; Optimization of the properties of the final product	[19]
20	Green extraction of bioactive compounds from chickpea (*Cicer arietinum* L.) sprouts, using a natural deep eutectic solvent formulation	3 factors: citric acid (A), glycerol (B), and water (C)	3 responses: Total phenolic content, Total flavonoid content, Antioxidant activity	All components influence the responses***; Optimization of extraction***	[20]
		Taguchi design			
21	Acetoin production by fermentation	4 factors: agitation (A), aeration rate (B), time (C) pH (D)	Acetoin concentration	Optimization of the fermentation process***	[44]
22	Microwave-assisted freeze- drying	5 factors: moisture content (A), drying temperature (B), Rotational speed (C), Microwave power density (D), Vacuum pressure (E)	Moisture content, rehydration rate, color, sugar content, energy consumption, texture analysis, sensory analysis	Optimization of the microwave-assisted freeze-drying process***	[21]
23	Development of nutritionally enhanced fish burgers	8 factors: atlantic bonito (A), chickpeas (B), garlic (C), pepper (D) *Spirulina* (E), *F. vesiculosus* (F), xanthan (G), carrageenan (H)	Physicochemical properties, texture, phenolic content, and antioxidant activity (DPPH, ABTS)	Optimization of all responses***	[22]

* Significant influence: $p < 0.05$; ** Very significant influence: $p < 0.01$; *** Highly significant influence: $p < 0.001$

4.1 *Full factorial designs 2^k*

As mentioned earlier, this design is selected to study the linear effects of a small number of factors and their interactions in a simple and cost-effective manner, as demonstrated in studies 1, 2, and 3 of Table 2 [13, 36, 37], where the number of studied factors does not exceed three. The simplicity of this design allows for a direct observation of these effects, making interpretations straightforward and clear. This enables the operator to make well-informed decisions, such as selecting the experimental conditions that maximize or minimize the response(s). These advantages are not typically present in related designs such as Fractional Factorial Designs 2^{k-p} and Plackett-Burman Designs.

4.2 *Fractional factorial design 2^{k-p} and Plackett-Burman design*

Studies 4 to 6 in Table 2 illustrate the application of Fractional Factorial Designs 2^{k-p}, demonstrating the use of this design with more than three factors-ranging from four to seven in these cases [14, 15, 38]. In contrast, studies 7 to 9 showcase applications of the Plackett-Burman Design, which is particularly suitable for experiments involving a large number of factors, ranging from nine to fourteen in these examples [39–41].

Both designs are primarily employed for screening of factors. After the screening stage, researchers may either conclude the study-if the objective has been achieved, as demonstrated in study 6 using a Fractional Factorial Design 2^{k-p} [15] and in study 9 using the Plackett-Burman Design [41]-or proceed with a different design approach, depending on the study's objectives. To investigate linear relationships between responses and factors, a Full Factorial Design 2^k may be appropriate. Alternatively, to explore quadratic relationships, researchers can apply Response Surface Methodology (RSM) using suitable designs. This is illustrated in studies 10 to 13 [14, 38–40].

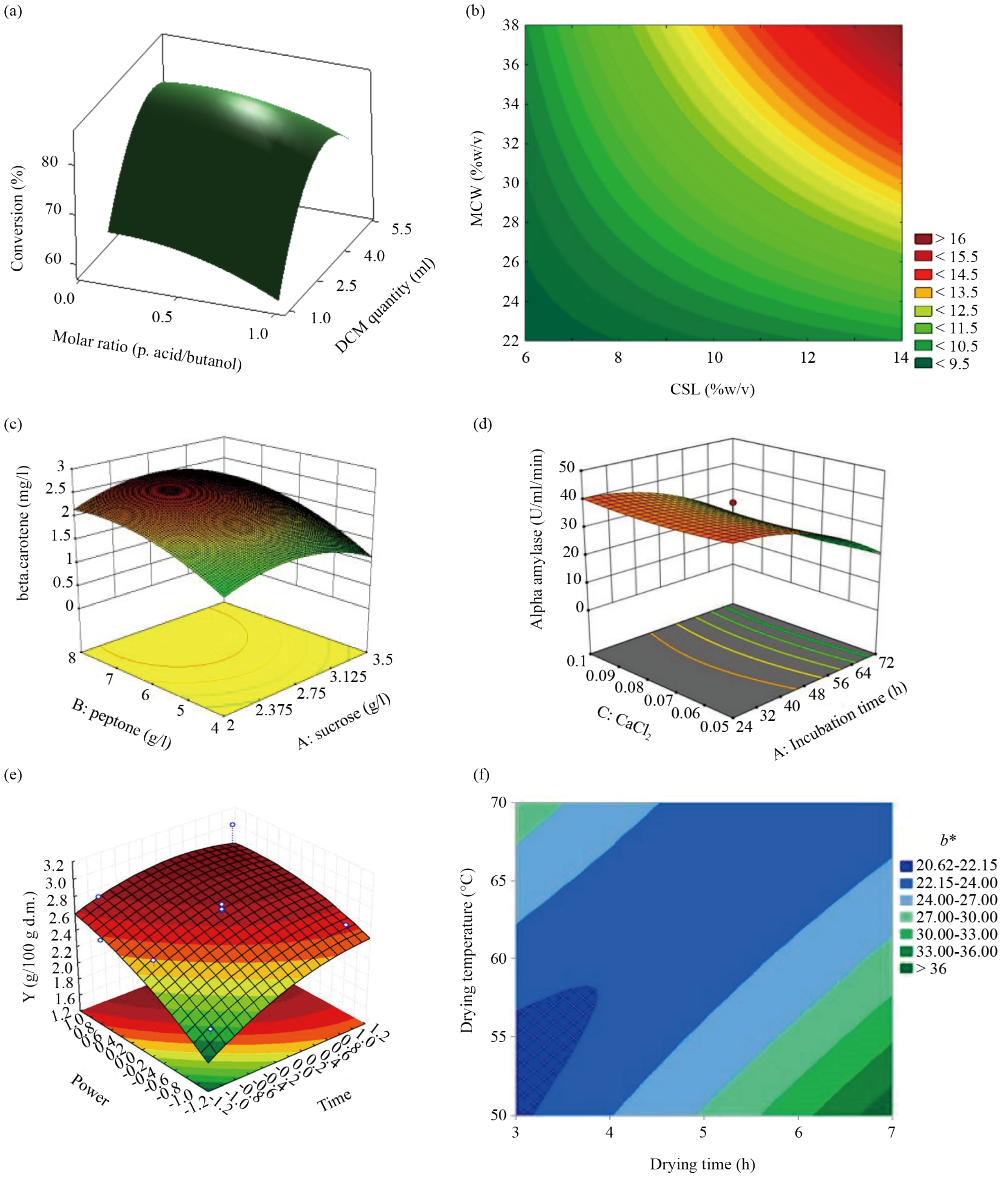

Figure 2. Examples of 3D response surface representations: (a) Enzymatic reaction conversion yield [38]; (b) Concentration of produced lipase and productivity [14]; (c) β-carotene concentration [39]; (d) Alpha amylase activity [40]; (e) Essential oils extraction yield [43]; (f) b^* color parameter of Tunisian date paste [17]

4.3 *Response surface methodology*

Many studies in the literature employ the Response Surface Methodology (RSM) with its various designs. The two main designs frequently used are the Central Composite Design and the Box-Behnken Design. Both designs are applied not only after screening studies, as discussed in Section 4.2 and demonstrated by studies 10 to 13 in Table 2 [14, 38–40], but also directly, as shown in studies 14 to 17 [16, 17, 42, 43].

This type of design is generally recognized for its capacity to optimize responses and its distinctive quadratic model, regardless of the number of factors tested.

One of the most powerful tools in Response Surface Methodology (RSM) is the three-dimensional visualization of response surfaces as a function of two varying factors, while keeping all other factors constant. Figure 2 presents examples of such representations based on the studies listed in Table 2.

4.4 *Mixture designs*

Generally, mixture design is considered a distinctive design compared to others because the factors used are defined as fractions of a mixture, unlike in other designs, as shown in [18–20] (Table 2). In fact, with this design, we can produce a non-existent product that can either serve as an interesting direct product for customers (as demonstrated in [18, 19], respectively) or be used to improve another process (as demonstrated in [20]). This concept could open many new possibilities in scientific research.

4.5 *Taguchi design*

The Taguchi design is often chosen when the primary objective is to identify the key factors and their corresponding levels-acting as controlling factors-among a large number of tested variables, especially in presence of noise factors, and without the need to establish a precise relationship between the studied response(s) and the factors. Particularly appealing to researchers, his approach serving as a practical tool for experimentally optimizing not only the responses but also the entire process, as demonstrated in Studies 21 to 23 [21, 22, 44].

Furthermore, numerous studies in the literature-particularly in food processing engineering-have adopted the Taguchi design. This preference is largely due to the field's inherent exposure to various sources of noise, such as the heating-cooling dynamics of control systems, Maillard reactions, and other unpredictable process variables. These challenges make Taguchi design more suitable than other screening methods like Fractional Factorial Design (2^{k-p}) and Plackett-Burman Design, which are less robust to noise.

In addition, scale-up studies frequently employ the Taguchi method to assess process performance under variable conditions, as exemplified by the work of [44].

5. Integration of DoE with sensory evaluation

The integration of DoE with sensory evaluation is largely specific to food processing engineering due to the nature of sensory testing itself. When the goal of a study is to develop a food product intended for direct consumer use, sensory testing is commonly employed to validate the results through consumer acceptance. This approach is typically influenced by the nature of the final product rather than by the choice of experimental design. For instance, study 14 [42], study 16 [16], study 18 [18], and study 23 [22] all incorporated sensory evaluation following the application of different DoE techniques-namely, the central composite design, Box-Behnken design, mixture design, and Taguchi design, respectively.

6. Future of DoE in food processing engineering

The application of Design of Experiments (DoE) in food processing engineering holds significant potential for enhancing both laboratory research and industrial-scale production. DoE facilitates efficient experiment planning and execution, and its integration with advanced technologies-such as Artificial Intelligence (AI), Machine Learning (ML), big data analysis, and real-time process monitoring-can significantly optimize food production and enhance product quality [45]. AI-enhanced DoE methods, including adaptive designs and active learning, enable smarter experimental planning based on evolving data. Real-time sensor systems and Internet of Things (IoT) devices support continuous monitoring and closed-loop control, thereby allowing dynamic optimization of process parameters. Hybrid modeling approaches that combine mechanistic understanding with ML provide deeper insights into complex food systems, while transfer learning allows for efficient adaptation across different applications. Historical big data can be mined to inform new strategies, and consumer-focused applications can benefit from DoE integrated with sensory analysis and ML for optimizing preferences and supporting personalized nutrition. These innovations also contribute to sustainability through cleaner formulations and reduced resource use, marking a major shift toward intelligent and responsive experimental designs in food processing engineering.

7. Conclusions

This review is intended to assist researchers who are working with or seeking to understand the concept of Design of Experiments (DoE). The paper provides comprehensive and detailed information on the most commonly used types of DoE, illustrated with several real-world examples from the food processing industry. Although food processing represents a principal domain for the application of DoE, particularly in facilitating process scale-up and ensuring adherence to regulatory requirements and quality standards, it is not the sole area in which DoE demonstrates significant utility. Yet, these applications are particularly notable due to the complex interplay of multiple factors that influence the outcomes. The coupling of the DoE concept with emerging technologies-such as artificial intelligence, machine learning, big data analytics, and real-time process monitoring-holds great promise for significantly advancing both process optimization and food quality in food engineering.

Author contributions

B.H. wrote and thoroughly reviewed the entire manuscript.

Conflict of interest

The author declares no competing financial interest.

References

[1] Montgomery DC. *Design and Analysis of Experiments*. 8th ed. Hoboken, NJ, USA: John Wiley & Sons; 2013.

[2] Fisher RA, Yates F. *Statistical Tables for Biological, Agricultural, and Medical Research*. 4th ed. Edinburgh, UK: Oliver and Boyd; 1953.

[3] Fisher RA. *Statistical Methods for Research Workers*. 13th ed. Edinburgh, UK: Oliver and Boyd; 1958.

[4] Fisher RA. *The Design of Experiments*. 8th ed. New York, NY, USA: Hafner Publishing Company; 1966.

[5] Box JF. *R.A. Fisher: the life of a scientist*. New York, NY, USA: Wiley; 1978.

[6] Box GEP, Wilson KG. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1951; 13(1): 1-38. Available from: https://doi.org/10.1111/j.2517-6161.1951.tb00067.x.

[7] Kiefer J, Wolfowitz J. Optimum designs in regression problems. *Annals of Mathematical Statistics*. 1959; 30(2): 271-294. Available from: https://doi.org/10.1214/aoms/1177706252.

[8] Kiefer J. Optimum designs in regression problems II. *Annals of Mathematical Statistics*. 1961; 32(1): 298-325. Available from: https://doi.org/10.1214/aoms/1177705160.

[9] Taguchi G, Wu Y. *Introduction to Off-Line Quality Control*. Nagoya, Japan: Central Japan Quality Control Association; 1980.

[10] Taguchi G. *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. White Plains, NY, USA: UNIPUB; 1987.

[11] Taguchi G. *Introduction to Quality Engineering*. White Plains, NY, USA: UNIPUB; 1991.

[12] Goupy J, Creighton L. *Introduction Aux Plans D'expériences [Introduction to Experimental Designs]*. 3rd ed. Paris, France: Dunod; 2006.

[13] Kousoulaki K, Krasnov A, Ytteborg E, Sweetman J, Pedersen ME, Høst V, et al. A full factorial design to investigate interactions of variable essential amino acids, trace minerals and vitamins on Atlantic salmon smoltification and post transfer performance. *Aquaculture Reports*. 2021; 20: 100704. Available from: https://doi.org/10.1016/j.aqrep.2021.100704.

[14] Ramos PR, de Oliveira AL, Ramos GVC, Kamimura ES. Esterification process in supercritical carbon dioxide catalyzed by Geotrichum candidum lipase produced with Mozzarella cheese whey and corn steep liquor. *Processes*. 2024; 12(1): 2086. Available from: https://doi.org/10.3390/pr12102086.

[15] Marrakchi F, Kriaa K, Hadrich B, Kechaou N. Experimental investigation of processing parameters and effects of degumming, neutralization and bleaching on lampante virgin olive oil's quality. *Food and Bioproducts Processing*. 2015; 94: 124-135. Available from: https://doi.org/10.1016/j.fbp.2015.02.002.

[16] Ammar I, Sebii H, Aloui T, Attia H, Hadrich B, Felfoul I. Optimization of a novel, gluten-free bread's formulation based on chickpea, carob and rice flours using response surface design. *Heliyon*. 2022; 8(12): e12164. Available from: https://doi.org/10.1016/j.heliyon.2022.e12164.

[17] Ben Amara S, Lakoud A, Mahmoudi I, Ben Tekaya I, Amri A, Snoussi A, et al. Optimization of the industrial production process of Tunisian date paste for sustainable food systems. *Processes*. 2024; 12(10): 2083. Available from: https://doi.org/10.3390/pr12102083.

[18] Ammar I, Gharsallah H, Ben Brahim A, Attia H, Ayadi MA, Hadrich B, et al. Optimization of gluten-free sponge cake fortified with whey protein concentrate using mixture design methodology. *Food Chemistry*. 2021; 343: 128457. Available from: https://doi.org/10.1016/j.foodchem.2020.128457.

[19] Medina-Rendon EA, Guatemala-Morales GM, Padilla-Camberos E, Corona-González RI, Arriola-Guevara E, García-Fajardo JA. Production of extrudate food with mango by-products (*Mangifera indica*): analysis of physical, chemical, and sensorial properties. *Processes*. 2021; 9(9): 1660. Available from: https://doi.org/10.3390/pr9091660.

[20] Khalid W, Koraqi H, Benmebarek IE, Moreno A, Alsulami T, Mugabi R, et al. Optimization of UAE-NADES green extraction of bioactive compounds from chickpea (*Cicer arietinum* L.) sprouts using simplex lattice mixture design methodology. *Ultrasonics Sonochemistry*. 2025; 112: 107186. Available from: https://doi.org/10.1016/j.ultsonch.2024.107186.

[21] Chen BL, Lin GS, Amani M, Yan WM. Microwave-assisted freeze drying of pineapple: kinetic, product quality, and energy consumption. *Case Studies in Thermal Engineering*. 2023; 41: 102682. https://doi.org/10.1016/j.csite.2022.102682.

[22] Solinho J, Gonçalves S, Machado S, Pereira-Pinto R, Vázquez M, Pinheiro R. Development of nutritionally enhanced fish burgers: integrating Atlantic bonito (*Sarda sarda*) with seaweed and hydrocolloids for sustainable food innovation. *LWT-Food Science and Technology*. 2025; 215: 117247. Available from: https://doi.org/10.1016/j.lwt.2024.117247.

[23] Savic IM, Marinkovic VD, Tasic L, Krajnovic D, Savic IM. From experimental design to quality by design in pharmaceutical legislation. *Accreditation and Quality Assurance*. 2012; 17: 627-633. Available from: https://doi.org/10.1007/s00769-012-0926-y.

[24] Okhio-Seaman L. Validation, quality, and regulatory considerations in continuous biomanufacturing. In: *Continuous Biomanufacturing: Innovative Technologies and Methods*. Weinheim, Germany: Wiley-VCH; 2017. p.549-567. Available from: https://doi.org/10.1002/9783527699902.ch20.

[25] Taylor C, Marschall L, Kunzelmann M, Richter M, Rudolph F, Vajda J, et al. Integrated process model applications linking bioprocess development to quality by design milestones. *Bioengineering*. 2021; 8(11): 156. Available from: https://doi.org/10.3390/bioengineering8110156.

[26] Passerine BFG, Breitkreitz MC. Important aspects of the design of experiments and data treatment in the analytical quality by design framework for chromatographic method development. *Molecules*. 2024; 29(24): 6057. Available from: https://doi.org/10.3390/molecules29246057.

[27] Fidaleo M. Functional data analysis and design of experiments as efficient tools to determine the dynamical design space of food and biotechnological batch processes. *Food and Bioprocess Technology*. 2020; 13: 1035-1047. Available from: https://doi.org/10.1007/s11947-020-02449-2.

[28] Benetti AA, Benetti C. Regulatory aspects of quality by design. In: *Introduction to Quality by Design (QbD): From Theory to Practice*. Singapore: Springer; 2024. p.35-57. Available from: https://doi.org/10.1007/978-981-99-8034-5_2.

[29] Kallel I, Tarhouni N, Elaguel A, Gargouri B, Hadrich B, Allouche N, et al. Exploration of the antiproliferative and antioxidant effects and the molecular docking study EGFR and VEGFR2 of essential oil from *Citrus aurantium* peels. *Journal of Essential Oil Bearing Plants*. 2023; 26(5): 1130-1150. Available from: https://doi.org/10.1080/0972060X.2023.2282575.

[30] Kallel I, Tarhouni N, Elaguel A, Mekrazi S, Khabir A, Hadrich B, et al. The phytochemical and pharmacological properties of Citrus sinensis 'Maltese Half-Blood' essential oil peels extracted and optimized by response-surface methodology. *Current Pharmaceutical Biotechnology*. 2023; 24(15): 1938-1951. Available from: https://doi.org/10.2174/1389201024666230405120520.

[31] Kallel I, Bayoudh A, Gargouri B, Khannous L, Elaguel A, Tharhouni N, et al. Modeling of antiproliferative effects of *Salvia officinalis* L. essential oil optimized using Box-Behnken design. *Journal of Plant Biochemistry and Biotechnology*. 2022; 32: 239-252. Available from: https://doi.org/10.1007/s13562-022-00799-w.

[32] Kallel I, Hadrich B, Gargouri B, Chaabane A, Lassoued S, Gdoura R, et al. Ezeddine optimization of cinnamon (*Cinnamomum zeylanicum* Blume) essential oil extraction: evaluation of antioxidant and antiproliferative effects. *Evidence-Based Complementary and Alternative Medicine*. 2019; 2019: 6498347. Available from: https://doi.org/10.1155/2019/6498347.

[33] Elaguel A, Kallel I, Gargouri B, Ben Amor I, Hadrich B, Ben Messaoud E, et al. Lawsonia inermis essential oil: extraction optimization by RSM, antioxidant activity, lipid peroxidation and antiproliferative effects. *Lipids in Health and Disease*. 2019; 18: 196. Available from: https://doi.org/10.1186/s12944-019-1141-1.

[34] Ghorbel A, Wedel S, Kallel I, Cavinato M, Sakavitsi ME, Fakhfakh J, et al. Extraction yield optimization of oleaster (*Olea europaea* var. *sylvestris*) fruits using response surface methodology, LC/MS profiling and evaluation of its effects on antioxidant activity and autophagy in HFF cells. *Journal of Food Measurement and Characterization*. 2021; 15: 4946-4959. Available from: https://doi.org/10.1007/s11694-021-01058-7.

[35] Ben-Amor I, Gargouri B, Attia H, Tlili K, Kallel I, Musarra-Pizzo M, et al. In vitro anti-Epstein Barr virus activity of *Olea europaea* L. leaf extracts. *Plants*. 2021; 10(11): 2445. Available from: https://doi.org/10.3390/plants10112445.

[36] Vieira VS, Peron-Schlosser B, Conde MBM, Canan C, Colla LM, Colla E. Valorization of residual fractions from defatted rice bran protein extraction: a carbohydrate-rich source for bioprocess applications. *Processes*. 2024; 12: 2348. Available from: https://doi.org/10.3390/pr12112348.

[37] Hadrich B, Dimitrov K, Kriaa K. Modelling investigation and parameters study of polyphenols extraction from carob (*Ceratonia siliqua* L.) using experimental factorial design. *Journal of Food Processing and Preservation*. 2017; 41(2): e12769. Available from: https://doi.org/10.1111/jfpp.12769.

[38] Mahfoudhi A, Ben Mabrouk S, Hadrich B, Mhadhbi M, Abderrazek H, Alghamdi O, et al. Efficient green enzymatic synthesis of lipophilic piperic acid esters by immobilized Rhizopus oryzae lipase: optimization and antioxidant activities. *Catalysis Letters*. 2024; 154(1): 4283-4301. Available from: https://doi.org/10.1007/s10562-024-04650-w.

[39] Abdelhafez AA, Husseiny SM, Ali AA, Sanad HM. Optimization of β-carotene production from agro-industrial by-products by *Serratia marcescens* ATCC 27117 using Plackett-Burman design and central composite design. *Annals of Agricultural Sciences*. 2016; 61(1): 87-96. Available from: https://doi.org/10.1016/j.aoas.2016.01.005.

[40] Abdullah R, Ahmad S, Nisar K, Kaleem A, Iqtedar M. Response surface methodology as an approach for optimization of alpha amylase production by using bacterial consortium under submerged fermentation. *Kuwait Journal of Science*. 2024; 51(3): 100220. Available from: https://doi.org/10.1016/j.kjs.2024.100220.

[41] Karlapudi AP, Krupanidhi S, Reddy ER, Indira M, Bobby MN, Venkateswarulu TC. Plackett-Burman design for screening of process components and their effects on production of lactase by newly isolated *Bacillus* sp. VUVD101

strain from dairy effluent. *Beni-Suef University Journal of Basic and Applied Sciences*. 2018; 7(4): 543-546. Available from: https://doi.org/10.1016/j.bjbas.2018.06.006.

[42] Silva JM, da Silva Crozatti TT, Figueiredo Saqueti BH, Rovigatti Chiavelli LU, Matioli G, Santos OO. Optimization of foam mat drying using central composite design to produce mixed juice powder: a process and characterization study. *Food and Bioproducts Processing*. 2024; 146: 58-68. Available from: https://doi.org/10.1016/j.fbp.2024.04.009.

[43] Balti MA, Hadrich B, Kriaa K, Kechaou N. Lab-scale extraction of essential oils from Tunisian lemongrass (*Cymbopogon flexuosus*). *Chemical Engineering and Processing-Process Intensification*. 2018; 124(1): 164-173. Available from: https://doi.org/10.1016/j.cep.2017.12.012.

[44] Taiwo AE, Madzimbamuto TN, Ojumu TV. Optimization of process variables for acetoin production in a bioreactor using Taguchi orthogonal array design. *Heliyon*. 2023; 6(10): e05103. Available from: https://doi.org/10.1016/j.heliyon.2020.e05103.

[45] Kovacević S, Banjac MK, Podunavac-Kuzmanović S. Artificial intelligence and experimental design: the flywheel of innovating food processing engineering. *Processes*. 2025; 13(3): 846. Available from: https://doi.org/10.3390/pr13030846.