ISSN: 2456-1452
Maths 2023; SP-8(5): 30-38

https://www.mathsjournal.com
Received: 08-07-2023
Accepted: 15-08-2023

U Sharin Shana
Research Scholar, Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

M Nirmala Devi
Assistant Professor (Mathematics), Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

D Ramesh
Professor and Head (Bioenergy), Department of Renewable Energy Engineering, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Balaji Kannan
Professor (Soil and Water Conservation) and Head, Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

M Djanaguiraman
Associate Professor, Department of Crop Physiology, AC&RI, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

Corresponding Author:
M Nirmala Devi
Assistant Professor (Mathematics), Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Biochemical process optimization *via* statistical methods: A mini review

U Sharin Shana, M Nirmala Devi, D Ramesh, Balaji Kannan and M Djanaguiraman

DOI: https://doi.org/10.22271/maths.2023.v8.i5Sa.1164

Abstract
Biochemical process optimization is now a crucial topic for research and development. Statistical approaches are currently being used by researchers to more effectively optimize the process, reduce waste and unpredictability, improve product quality, and increase process effectiveness. Current advancements in this field include the use of machine learning techniques and the Design of Experiments (DoE). The significance of statistical approaches as useful instruments for process optimization in biochemical research is highlighted in this work. The Taguchi Method, Response Surface Methodology (RSM), and Artificial Neural Networks (ANN) combined with Genetic Algorithm (GA) are three popular approaches that are focused for further comparison. The study presents an overview of each technique, investigates how it might be applied to optimization, examines its benefits and drawbacks, and identifies its main distinctions.

Keywords: Process optimization, design of experiment, response surface methodology, ANN-GA

1. Introduction
Pharmaceuticals, biotechnology, and environmental engineering are just a few fields where biochemical processes are essential. These procedures modify biomolecules, their production, and their breakdown in order to give the raw ingredients for new biomolecules. They achieve this by utilizing living organisms or biological components [1]. Examples include enzymatic processes for waste treatment and biocatalysis, as well as fermentation processes for the manufacture of medicines, biofuels, and enzymes [2-4]. However, in order to maximize product yield, improve process efficacy, and reduce costs, the biochemical processes must be optimized.

In recent years, statistical approaches have become effective instruments for enhancing biological processes. In addition, statistical methods, which provide a systematic approach to experiment design, response surface modeling, and optimization algorithms, enable researchers to efficiently investigate and enhance complex biochemical systems [5]. These methods make it possible to identify process variables that have a major impact on the entire process, build mathematical models that represent the behaviour of the process, and determine the ideal process parameters [6].

The review paper provides a summary of the statistical techniques now employed to improve the efficiency of biochemical processes, such as design of experiments (DoE), response surface methodology (RSM), and optimization algorithms. The goal of this work is to optimize biological processes using statistical methods in a variety of ways. It also analyses case studies from diverse research works, identifies new trends and developments in the field, and investigates the concepts and uses of critical statistical techniques.

2. Process Optimization
The performance of the systems must be enhanced, and the yield must be increased without the cost. The method used for this purpose is called optimization [7]. There is a parameter change in the general practice of determining the optimal operating conditions while keeping the others constant. This is called the one-variable-at-a-time technique. The main drawback of this method is that it does not account for interactions between the variables, and it ultimately

has to show all of the parameters' effects on the process. Process modification to maximize a collection of parameters while adhering to specific limitations is called process optimization [8]. The most common goals are minimizing cost, maximizing throughout, and/or efficiency. This is one of the primary quantitative tools in industrial decision-making [9]. Process optimization aims to increase one or more process specifications while maintaining all others within their bounds [10]. Process optimization tools often use statistical techniques to find the best solution. However, this notion is only partially accurate. While statistical techniques are essential, it is crucial to deeply understand the process before investing time in optimization [11]. Throughout time, various methodologies have been devised for process optimization.

3. Importance of statistical approach in the optimization process

Optimization involves finding the best conditions for a procedure to achieve optimal results [12]. Traditionally, it was done by changing one factor at a time while keeping others constant, known as the one-variable-at-a-time (OVAT) approach. However, OVAT needs to consider the interactive effects among variables, leading to an incomplete understanding of their impact on the desired outcome [13]. Moreover, conducting multiple experiments using OVAT can be time-consuming and labor-intensive when studying the effects of different independent variables. To overcome these limitations, multivariate statistical techniques, such as response surface methodology (RSM), have been used to optimize analytical procedures [14]. Statistical approaches are essential in optimizing chemical experiments for several reasons, including identifying critical factors, optimizing process parameters, reducing the number of experimental runs, predicting process performance, and understanding process variability. By employing statistical methods, the number of experimental runs can be reduced [15]. Optimization techniques aim to find the optimal values of a set of parameters that maximize or minimize an objective function. These are crucial in statistics for estimation and model fitting, allowing for comparing different choices to determine the "best" option.

4. Trends in Statistical methods used for process optimization

Several statistical methods are employed for the process designing and optimization. Figure.1 illustrates a schematic representation of the diverse techniques utilized in process optimization.

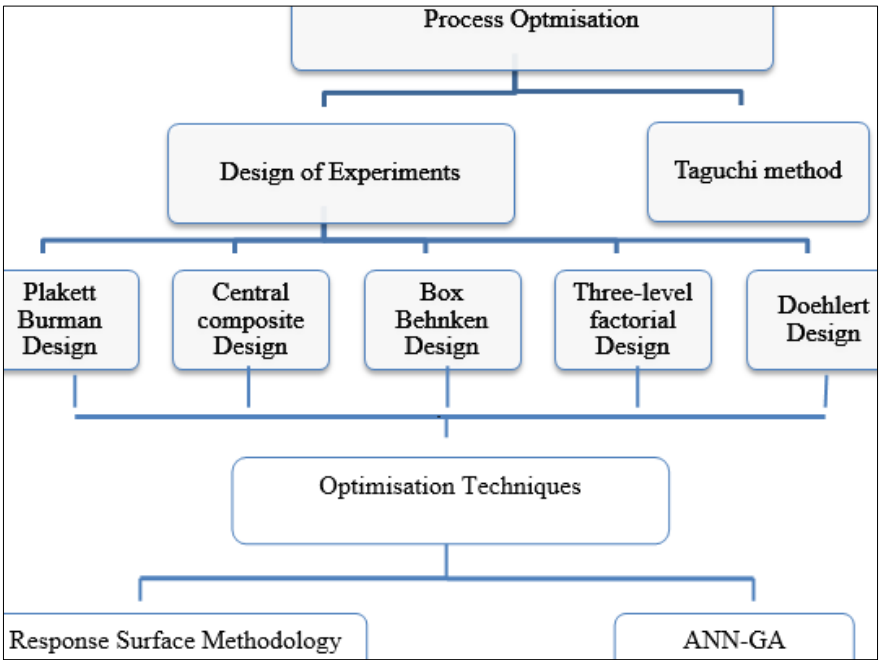

Fig 1: Trends in Statistical Methods for Process Optimization

4.1 Design of Experiment (DOE)

Experimental design is a systematic approach used to achieve specific goals or objectives of the proposed work. Design of experiments (DOE) is a powerful tool for optimization, surpassing the limitations of the classical one-factor-at-a-time (OFAT) method. DOE allows for changing multiple components simultaneously, making the process more efficient [16]. DOE involves strategically planning and executing a series of experiments to gather extensive information about the effects of multiple parameters on the desired output or response. By comparing several factors simultaneously and observing their effects, responses can be determined, ranked, and statistically analyzed [17]. Compared to OFAT, DOE requires fewer experiments, less time, and fewer resources while providing the same amount of information [18]. Properly planned experiments with sufficient sample sizes are crucial to answer research objectives efficiently. These techniques, commonly called as DOE, involve predicting mathematical models and optimizing the process variables based on experimental data [19]. All factor combinations are tested in a complete factorial design, while in a partial factorial analysis, selected combinations are tested based on existing literature or prior knowledge.

4.1.1 Plakett Burman design

Plackett-Burman Design (PBD) is a two-level experimental design that efficiently identifies significant factors in a process while assuming negligible interactions. It allows for screening multiple variables using minimal experiments [20]. The PBD involves actual variables that change concentration and dummy variables that remain constant for error estimation. A trial matrix represents each variable at high (H) and low (L) levels. The effect of each variable is determined using a specific equation [4]. Further, it helps eliminate non-

contributing factors early on and focuses on the significant effects more precisely. The following equation determines the Effect of each variable.

$$E_{X1} = \frac{2(\sum Y_{X1H} - \sum Y_{X1L})}{N} \quad (1)$$

Whereas, E_{X1} = Effect of variable; $Y_{X1\text{-}H}$ = yield from the trials having a high concentration of variable; Y_{X1-L} = yield from the trials having a low concentration of variable and N = total trials conducted. The student's t-test determines the significance level of the Effect of each variable: PBD is a reliable method to assess the importance of variables for a specific output, reducing the number of experiments significantly. It focuses on screening variables directly affecting the desired outcome, disregarding interaction effects [21]. However, PBD's efficiency is limited as it assumes no interactions or only additive effects, potentially enhancing or masking the results of analyzed factors. Nonetheless, PBD serves as a starting point in experimental design, helping to identify non-contributing factors and creating a list for further investigation [22]. It is commonly referred to as a "screening design" due to its ability to distinguish contributing factors for higher yield and has been widely used for screening the variables in many biochemical experiments (Table1)

Table 1: Plakett Burman Screening Design approaches for Biochemical experiments

SI. No	Name of Process	Dependent variable	Screened Independent variables	References
1	Hydro distillation process	Essential oil yield (%)	a) Processing time (min) b) Ratio Plant material /water c) Division of the plant material d) Moisture e) Individuality	[17]
2	Protease production	Protease activity (u/ml)	a) Glucose(g/L) b) Corn starch(g/L) c) Yeast extract(g/L) d) Corn steep liquor(g/L) e) Ammonium phosphate(g/L) f) Magnesium sulfate(g/L) g) Inoculum size(% (v/v) h) Incubation period(hours)	[15]
3	Laccase Production	Laccase yield (U/mL)	a) Incubation temperature(°C) b) Incubation period(hour) c) Agitation rate(rpm) d) Yeast extract (g/L) e) $MgSO_4.7H_2O$(mM) f) $(NH_4)_2SO_4$(mM) g) $CuSO_4$(mM) h) Trace elements (mL/L)	[20]
4	lipase production	Lipase production (U mL^{-1})	a) Glucose (g/L) b) Sesame oil (mL/L) c) Peptone (g/L) d) NaCl (g/L) e) $MnSO_4.H_2O$ (g/L)	[23]
5	Ethanol Production	Ethanol conc. (g/L)	a) $FeSO_4·7H_2O$ (g/L) b) $CaCl_2·2H_2O$ (g/L) c) $MnCl_2·4H_2O$ (g/L) d) $ZnSO_4·7H_2O$ (g/L) e) $MgSO_4·7H_2O$(g/L) f) $CoCl_2$ (g/L) g) Fermentation time (h) h) KH_2PO_4 (g/L) i) Inoculum level (%) j) $(NH_4)_2SO_4$ (g/L) k) NaCl (g/L) l) pH m) Temperature (◦C) n) $CuSO_4$ (g/L) o) TRS (%)	[24]

4.1.2 Taguchi Method

Dr. Genichi Taguchi developed the Taguchi method, which focuses on adjusting control factors to optimize system response in the presence of uncontrollable noise factors. It utilizes orthogonal arrays to explore the parameter space with minimal experiments, reducing time and cost [25]. The Taguchi method involves system strategy, parameter designing, and tolerance design. It helps identify significant factors with fewer experiments, saving time and resources [7]. ANOVA is used to analyze data from the Taguchi DOE and optimize performance characteristics [26]. Unlike PBD, the Taguchi method analyzes main effects and two-factor interactions while assuming higher-order interactions to be negligible. It focuses on addressing noise factors, which can affect quality. The Taguchi method has the advantages of saving experimental time, reducing costs, and improving quality. It has been successfully applied in optimizing various processes, and some recent examples are listed in the table.2

Table 2: Taguchi optimization approaches for Biochemical experiments

SI. No	Name of Process	Dependent variable	Independent variable	Design	References
1	Downstream Processes for Prodigiosin Extraction	Prodigiosin (mg/L)	a. The ratio of solid to liquid(g/mL) b. Sonication of duty cycle (%) c. Acoustic intensity (w/cm^2) d. Media pH	Taguchi OA design of L_9 (3^4)	[27]
2	Bioremediation of crude oil-contaminated water	Degraded crude oil & fish growth	a. Temperature, b. Inoculums conc. c. Crude oil conc. d. Time, e. NH_4Cl conc. f. K_3PO_4 conc.	L_{25}Taguchi orthogonal array	[28]
3	Acetoin production in a bioreactor	Acetoin conc. (g/L)	a. Agitation(rpm) b. Aeration(slpm) c. pH d. Fermentation time(days)	Taguchi L_9 orthogonal array (OA) design	[16]
4	Extraction yield of phycobiliproteins (PBPs) from *Oscillatoria sp*.	PC extraction (mg/g) PE extraction (mg/g) APC extraction (mg/ g)	a. Solid-to-liquid ratio (g/ml) b. Duty cycle (%) c. Electrical acoustic intensity (w/cm^2) d. pH	Taguchi L_9 (3^4)orthogonal array (OA) design	[29]
5	Production of Astaxanthin by using Fruit Waste Extract	Astaxanthin production(mg/g)	a. Temperature (ºC) b. Agitation (rpm) c. pH	Taguchi orthogonal array L_9	[30]
6	Biodegradation of crude oil using bacteria	Crude oil biodegradation %	a. Temperature, b. Salinity c. pH d. NH_4Cl conc. e. $FeSO_4.7H_2O$conc	Taguchi experimental design L_{16} (4^5)	[31]
7	Inulinase production from low-cost substrates	Inulinase enzyme Activity (U/gds)	a. Incubation Temp(◦C) b. Initial consent of KH_2PO_4 (%) c. Initial concent. of Ca^{2+} (mM) d. Initial pH	Orthogonal array of L9 (3^4) design	[32]
8	Production of biogas from raw vegetable wastes	Biogas generation (m^3 /ton of RVW)	a) Plastic content (%) b) h/D ratio c) Water content (ml) d) Digestion period (Week	Taguchi L_{16} (4^4)orthogonal array (OA) design	[26]
9	Production of proteolytic enzymes using agro-industrial by-products	Proteolytic enzyme activity in a submerged medium (U)	a. Inorganic nitrogen source (0.5%) b. Metal ions c. Agitation speed (rpm) d. Initial pH	Taguchi L_{16}orthogonal array (OA) design	[33]
10	α-Amylaseproduction by *Bacillus subtilis*	Amylase-smf (U/mg)	a. Time (hr) b. Carbon source (1%) c. Nitrogen source (1%) d. Amino acid (0.01%)	orthogonal array $L_{16}(4^5)$	[34]

4.1.3 Central composite design

In order to address the limitations of PBD that only considers main effects and ignores interactions among factors, the Central Composite Design (CCD) was introduced. CCD is widely used in Response Surface Methodology (RSM) to build quadratic models for the response variable without requiring a complete three-level factorial experiment [35]. CCD involves a combination of factorial design (two levels), center points, and star points. There are three types of CCD: Circumcentered CCD (CCC), Inscribed CCD (CCI), and Face-centered CCD (CCF) [4]. The design requires a full factorial or fractional factorial design, additional points at a distance α from the center (often in a star configuration), and a central point. The total number of experiments follows a specific formula, and the α-values are calculated based on the number of variables [36]. All factors are studied at five levels: -α, -1, 0, +1, and +α. CCD allows for studying both main effects and interactions while optimizing the response variable in a structured and efficient manner and so widely used for designing many biochemical experiments (Table.3)

4.1.4 Box-Behnken design

Box-Behnken design is an alternative to CCD and does not require a quadratic design or an embedded factorial or fractional factorial design. In this design, treatment combinations are located at the midpoints of the edges of the process space and the center. The rotatable design uses three levels for each factor [37]. However, it has limited capability for orthogonal blocking compared to central composite designs. Box-Behnken designs have experimental points on a hypersphere equidistant from the center point [35]. Key characteristics include a specific formula for the number of experiments and adjustment of factor levels at three levels (-1, 0, +1) with equal spacing [4] and also widely used for many Biochemical experiments (Table.3).

4.1.5 Three-level factorial

A full three-level factorial design is not commonly used in response surface methodology (RSM) when the number of factors exceeds 2. This is because the number of experiments required for this design (calculated as N = 3k, where N is the number of experiments and k is the number of factors) becomes very large, resulting in decreased efficiency for modeling quadratic functions [6].

4.1.6 Doehlert designs

This design utilizes circular, spherical, or hyper-spherical domains based on the number of variables, ensuring uniformity within the experimental domain. The design requires specific experiments based on factors and central point replication. Each variable can be studied at different levels, accommodating restrictions or preferences [6]. The intervals between levels exhibit a uniform distribution. The design allows the matrix to be shifted to another experimental region using adjacent points.

Table 3: Design of Experiments (DoE) in Biochemical Research: A Table of Key Examples.

SI. No	Name of Process	Dependent variable	Independent variable	Design of Experiment	References
1	Bioethanol production	Bioethanol yield (g/L)	a. HCl concentration (%v/ v) b. Sonication time (min) c. Yeast inoculum size (g/ L)	Box-Behnken rotatable design	[38]
2	Phytoremediation of arsenic	Total arsenic removal from soil (%)	a. Arsenic Conc. in soil (mg /kg) b. Sampling days(days) c. Aeration rate(L/min)	Box-Behnken rotatable design	[39]
3	Tannase production	Tannase activity (U/ml)	a. Tannic acid conc. (%) b. Fermentation period (h) c. Temperature (^oC) d. pH	Central composite design (CCD)	[40]
4	Ethanol production	Ethanol conc. (g/L)	a. Temperature (◦ C) b. Inoculum level (%) c. TRS (%) d. pH	Box–Behnken design	[24]
5	Lipase production	Lipase production	a. Olive oil b. Tween 80 c. KH_2PO_4	Central composite rotatable design	[41]
6	Production of vanillic acid	Vanillic acid yield (mg /g)	a. L-asparagine concentration (mmol /L) b. pH of solid medium c. Moisture content (%)	Box–Behnken experimental design	[42]
7	Lipase production	Lipase Activity (IU/ml)	a. Aeration (vvm) b. Agitator speed (rpm) c. Medium volume (l)	Face-centered central composite design (FCCCD)	[43]
8	Enzymatic Hydrolysis of Recycled Paper	CES (%) (conversion efficiency of the substrate)	a. Enzyme concentration (FPU) b. Temperature (°C) c. Stirring rate (rpm)	Central composite design (CCD	[44]
9	Production of *Metarhizium anisopliae* conidiospores	*M. anisopliae* conidial yield	a. Moisture content (%) b. Yeast extract (g) c. pH	2^3 full factorial central composite design	[13]
10	Lipase production	Lipase activity (U /mL)	a. Glucose (g/L) b. Sesame oil (Ml/L) c. Peptone (g/L) d. NaCl (g/L) e. $MnSO_4.H_2O$ (g/L)	2^5 Central Composite Design(CCD)	[23]

4.2 Optimization Techniques

4.2.1 Response surface methodologies (RSM)

RSM (Response Surface Methodology) is a collection of statistical techniques for optimizing processes influenced by multiple variables. It involves performing designed experiments, estimating model coefficients, and predicting and validating the response [45]. The RSM is commonly employed in biotechnology to optimize culture conditions and processing parameters. It generates mathematical models to describe the underlying processes and is often visualized through response surface graphs [46]

The relationship between the response and the input is given in Eq. (2)

$$\eta = f(x_1, x_2, x_3 \ldots x_n) + \varepsilon \quad (2)$$

Whereas, η is the response, f is the unknown response function, $x_1, x_2, .x_n$ denote the independent variables, n is the number of the independent variables, and ε is the statistical error. The model used in RSM is generally a full quadratic equation or the diminished form of this equation. The second-order model can be written as follows

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j X_j + \sum_{j=1}^{k} \beta_{ij} X_j^2 + \sum \sum_{i<1} \beta_{ij} X_i X_j \quad (3)$$

Whereas β_0, β_i, β_{ii}, and β_{ij} are regression coefficients for intercept, linear, quadratic, and interaction coefficients, respectively, and X_i and X_j are coded independent variables. The matrix notation of the model is given in Eq. (4):

$$y = X\beta + \varepsilon \quad (4)$$

The above equations can be solved by the least squares method (MLS). In MLS, it is assumed that random errors are identically distributed with a zero mean and a common unknown variance and are independent of each other.

$$\varepsilon_i = y_i - \hat{y}_i \quad (5)$$

Our criterion for choosing the estimates is that they should minimize the sum of the squares of the residuals, which is often called the sum of squares of the errors and is denoted by SSE.

$$SSE = \sum_{i=1}^{n} \varepsilon_i^2 \quad (6)$$

The overall predictive capability of the model is commonly explained by the coefficient of determination (R^2) and Absolute average deviationAAD.

$$AAD = \{[\sum_{i=1}^{p}(|y_{i,exp} - y_{i,exp}|)]/p\} \times 100 \quad (7)$$

The response surface plot and contour plot can visualize the predicted model equation. The response surface plot is the theoretical three-dimensional plot showing the relationship between the response and the independent variables. The two-dimensional display of the surface plot is called a contour plot, and in the contour plot, lines of constant response are drawn in the plane of the independent variables [47]. RSM has been successfully applied in various Biochemical studies, including medium formulation and improving microbial processes (Table.4). However, RSM has limitations in predicting optimal formulations for complex systems and studying interactions involving multiple variables. Researchers have increasingly turned to Artificial Neural Networks (ANN) to address these limitations as an alternative technique.

4.2.2 Artificial Neural Network (ANN)

An artificial neural network (ANN) is a computational model inspired by biological neural networks. It can be used to estimate, predict, control, and adjust process parameters [48]. ANN learns from data without the need for prior knowledge or equations and can handle large amounts of data, excel at pattern recognition, and work with complex systems [49]. The architecture of ANN includes input, hidden, and output layers of neurons. ANN can be trained in supervised, unsupervised, or reinforcement learning conditions [50]. It has been successfully applied in system design, modeling, optimization, and control. ANN's ability to filter noisy signals and generalize information through training makes it valuable. However, proper training is crucial for efficient operation, and input data quality determines output data quality.

4.2.2.1 Genetic algorithm (GA)

In the optimization process using Genetic Algorithm (GA), a trained mathematical model is used as a fitness function to determine the optimal concentration of medium components [4]. GA simulates the mutation process and is based on the principle of "survival of the fittest." It modifies a population of individual solutions by selecting parents, combining them through crossover, and applying random changes through mutation [51]. This iterative process allows the population to evolve towards the most favorable solution and is widely used for optimization in Biochemical experiments (Table.4). GA is particularly useful for solving optimization problems that involve non-differentiable, discontinuous, stochastic or highly nonlinear objective functions [52]

Table 4: A summary of designs and optimization techniques used for Biochemical studies

S. N.	Name of Process	Design	Optimization technique	References
1	Biocoagulation-flocculation of municipal solid waste	CCD	ANN, GA	[52]
2	Bioethanol production from food industry waste	BBD	ANN, GA	[38]
3	ε-polylysine production	CCD	ANN, GA	[48]
4	Biogas production	CCD	ANN, GA	[53]
5	Bioethanol production from sunflower stalk	CCD	RSM	[54]
6	Hydrogen production	BBD	RSM	[55]
7	Hydrogen production	BBD	RSM	[56]
8	Lipase production	FCCCD	RSM	[43]
9	lipase production	CCD	RSM	[23]
10	Ethanol production from wheat straw	BBD	RSM	[24]
11	Enzyme production	CCD	ANN, GA	[57]
12	Ethanol production from breadfruit	BBD	ANN, GA	[58]

5. Comparison of different optimization techniques

The Taguchi Method assumes that the input factors significantly impact the output response and that the factors can be studied independently [59]. RSM assumes that the relationship between input factors and output response is continuous and can be modeled using polynomial equations [60]. Conversely, ANN does not make strong assumptions about the relationship between inputs and outputs, allowing for more flexibility in modeling complex and nonlinear relationships [48]. The Taguchi Method employs orthogonal arrays to design a limited number of experiments that efficiently cover the entire parameter space [61]. RSM typically uses factorial designs or central composite designs to explore the response surface and estimate model coefficients. ANN does not require a specific experimental design as it can learn from existing data, but it may benefit from a well-designed dataset for training. The Taguchi Method focuses on determining the optimal parameter settings that minimize the effects of uncontrollable factors and maximize the signal-to-noise ratio [7]. RSM aims to model and predict the response surface using regression analysis, enabling optimization based on the estimated model. ANN is a versatile modeling technique that can capture complex relationships between inputs and outputs without requiring explicit functional forms, making it suitable for nonlinear and high-dimensional problems. The Taguchi Method employs an objective function based on the signal-to-noise ratio to evaluate and optimize the process. It uses methods like parameter design and tolerance design to find the optimal parameter settings [62]. RSM utilizes statistical techniques such as gradient-based optimization or desirability functions to find the optimal response values. ANN employs optimization algorithms such as back propagation or genetic algorithms to adjust network weights and biases iteratively, aiming to minimize specific loss or error functions.

6. Conclusion

Due to its many benefits over the more conventional one-variable-at-a-time optimization approach, statistical approaches have been widely accepted for optimizing analytical processes. These advantages include deriving useful knowledge from constrained experiments and figuring out how various variables interact to influence results. In this method, the experimental optimization can be selected by selecting an adequate experimental design, fitting an appropriate mathematical function, and assessing how well the model predicts experimental data. By analyzing the effects of the experimental conditions and their statistical

significance, experimental design is essential in discovering important factors. Based on the fitted model, it also aids in determining the location and characteristics of prospective local optima. The review illustrated examples of factorial and optimization experimental design applications.In addition to helping to assess model performance and spot potential outlier experiments, using statistically significant model coefficients helps to prevent over fitting. Both newcomers and experienced practitioners must understand the basic ideas of linear algebra and statistics that underlie experimental design because specialized experimental design software is anticipated to become increasingly common in the future.

7. Acknowledgment

The authors gratefully acknowledge the funding and facilities provided by the Department of Renewable Energy Engineering, Agricultural Engineering College & Research Institute, Tamil Nadu Agricultural University, Coimbatore to perform this work

8. References

1. Garcia-Lopez E, Alcazar P, Cid C. Identification of biomolecules involved in the adaptation to the environment of cold-loving microorganisms and metabolic pathways for their production. Biomolecules, 2021;11(8):1155.
2. Ahn Y, Logan BE. Effectiveness of domestic wastewater treatment using microbial fuel cells at ambient and mesophilic temperatures. Bioresource technology. 2010;101(2):469-475.
3. Alavi-Borazjani SA, Da Cruz Tarelho LA, Capela MI. Parametric optimization of the dark fermentation process for enhanced biohydrogen production from the organic fraction of municipal solid waste using Taguchi method. International Journal of Hydrogen Energy. 2021;46(41):21372-21382.
4. Singh V. Strategies for fermentation medium optimization: an in-depth review. Frontiers in microbiology. 2017;7:2087.
5. Sivakumar T. Statistical optimization of keratinase production by Bacillus cereus. Global Journal of Biotechnology and Biochemistry. 2011;6(4):197-202.
6. Bezerra MA. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta. 2008;76(5):965-977.
7. Karna SK, Sahai R. An overview on Taguchi method. International journal of engineering and mathematical sciences. 2012;1(1):1-7.
8. Mäkelä M. Experimental design and response surface methodology in energy applications: A tutorial review. Energy Conversion and Management. 2017;151:630-640.
9. Tsui K.-L. An overview of Taguchi method and newly developed statistical methods for robust design. Iie Transactions. 1992;24(5):44-57.
10. Desai KM. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. Biochemical Engineering Journal. 2008;41(3):266-273.
11. Roy R, Hinduja S, Teti R. Recent advances in engineering design optimization: Challenges and future trends. CIRP annals. 2008;57(2):697-715.
12. Fard Masoumi HR. Statistical optimization of process parameters for lipase-catalyzed synthesis of triethanolamine-based esterquats using response surface methodology in 2-liter bioreactor. The Scientific World Journal; c2013.
13. Prakash GB, Padmaja V, Kiran RS. Statistical optimization of process variables for the large-scale production of Metarhizium anisopliae conidiospores in solid-state fermentation. Bioresource Technology. 2008;99(6):1530-1537.
14. Sim JH, Kamaruddin AH. Optimization of acetic acid production from synthesis gas by chemolithotrophic bacterium–Clostridium aceticum using statistical approach. Bioresource technology. 2008;99(8):2724-2735.
15. Reddy L. optimization of alkaline protease production by batch culture of Bacillus sp. RKY3 through Plackett–Burman and response surface methodological approaches. Bioresource technology. 2008;99(7):2242-2249.
16. Taiwo AE, Madzimbamuto TN, Ojumu TV. Optimization of process variables for acetoin production in a bioreactor using Taguchi orthogonal array design. Heliyon. 2020, 6(10).
17. Haloui T. Application of Plackett-Burman design in the essential oil extraction by hydro-distillation process of *Pistacia lentiscus* L. leaves. Journal of Materials and Environmental Science. 2015;6(4):942-948.
18. Das SP, Das D, Goyal A. Statistical optimization of fermentation process parameters by Taguchi orthogonal array design for improved bioethanol production. Journal of Fuels; c2014.
19. Israni N, Shivakumar S. Polyhydroxyalkanoate (PHA) biosynthesis from directly valorized ragi husk and sesame oil cake by Bacillus megaterium strain Ti3: Statistical optimization and characterization. International journal of biological macromolecules. 2020;148:20-30.
20. Niladevi K. Optimization of laccase production from a novel strain-Streptomyces psammoticus using response surface methodology. Microbiological research. 2009;164(1):105-113.
21. Vanaja K, Shobha Rani R. Design of experiments: concept and applications of Plackett Burman design. Clinical research and regulatory affairs. 2007;24(1):1-23.
22. Jacyna J, Kordalewska M, Markuszewski MJ. Design of Experiments in metabolomics-related studies: An overview. Journal of pharmaceutical and biomedical analysis. 2019;164:598-606.
23. Rajendran A, Thangavelu V. Sequential optimization of culture medium composition for extracellular lipase production by *Bacillus sphaericus* using statistical methods. Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology. 2007;82(5):460-470.
24. Singh A, Bishnoi NR. Ethanol production from pretreated wheat straw hydrolyzate by Saccharomyces cerevisiae via sequential statistical optimization. Industrial Crops and Products. 2013;41:221-226.
25. Vijayalakshmi K, Rajakumar S. Antimicrobial protein production by *Bacillus amyloliquefaciens* MBL27: An application of statistical optimization technique. African Journal of Microbiology Research. 2010;4(22):2388-2396.
26. Mukhopadhyay D, Sarkar JP, Dutta S. Optimization of process factors for the efficient generation of biogas from raw vegetable wastes under the direct influence of plastic

materials using Taguchi methodology. Desalination and Water Treatment. 2013;51(13-15):2781-2790.
27. Paul T. Downstream Process Development for Extraction of Prodigiosin: Statistical Optimization, Kinetics, and Biochemical Characterization. Applied Biochemistry and Biotechnology. 2022;194(11):5403-5418.
28. Popoola LT, Yusuff AS. Biodegradation effects of crude oil-polluted water using bacteria isolates from sunflower husk on fish growth: parametric optimization using taguchi approach. Carpathian journal of earth and environmental sciences. 2021;16(1):211-222.
29. Sharma R. Statistical optimization of process parameters for improvement of phycobiliproteins (PBPs) yield using ultrasound-assisted extraction and its kinetic study. Ultrasonics sonochemistry. 2020;60:104762.
30. Korumilli T, Mishra S, Korukonda JR. Production of astaxanthin by Xanthophyllomyces dendrorhous on fruit waste extract and optimization of key parameters using Taguchi method. Journal of Biochemical Technology. 2020;11(1):25.
31. Zavareh MSH. Bioremediation of crude oil using bacterium from the coastal sediments of Kish Island, Iran. Iranian journal of public health. 2016;45(5):670.
32. Abd El Aty AA, Wehaidy HR, Mostafa FA. Optimization of inulinase production from low-cost substrates using Plackett–Burman and Taguchi methods. Carbohydrate polymers. 2014;102:261-268.
33. Kotlar C, Agüero M, Roura S. Statistical optimization of a novel low-cost medium based on regional agro-industrial by-products for the production of proteolytic enzymes by Bacillus cereus. Preparative Biochemistry and Biotechnology. 2012;42(5):406-425.
34. Uysal E. optimization of α-amylase production by Bacillus subtilis RSKK96: using the Taguchi experimental design approach. Preparative Biochemistry & Biotechnology. 2010;41(1):84-93.
35. Breig SJM, Luti KJK. Response surface methodology: A review on its applications and challenges in microbial cultures. Materials Today: Proceedings. 2021;42:2277-2284.
36. Elfghi F. A hybrid statistical approach for modeling and optimization of RON: A comparative study and combined application of response surface methodology (RSM) and artificial neural network (ANN) based on design of experiment (DOE). Chemical Engineering Research and Design. 2016;113:264-272.
37. Rajasimman M, Sangeetha R, Karthik P. Statistical optimization of process parameters for the extraction of chromium (VI) from pharmaceutical wastewater by emulsion liquid membrane. Chemical Engineering Journal. 2009;150(2-3):275-279.
38. Suresh T. Process intensification and comparison of bioethanol production from food industry waste (potatoes) by ultrasonic-assisted acid hydrolysis and enzymatic hydrolysis: Statistical modeling and optimization. Biomass and Bioenergy. 2020;142:105752.
39. Titah HS. Statistical optimization of the phytoremediation of arsenic by Ludwigia octovalvis-in a pilot reed bed using response surface methodology (RSM) versus an artificial neural network (ANN). International journal of phytoremediation. 2018;20(7):721-729.
40. Mohan S, Viruthagiri T, Arunkumar C. Statistical optimization of process parameters for the production of tannase by Aspergillus flavus under submerged fermentation. 3 Biotech. 2014;4:159-166.
41. Kumar R. Identification of variables and value optimization for optimum lipase production by Bacillus pumilus RK31 using statistical methodology. New Biotechnology. 2011;28(1):65-71.
42. Gadhe A. Statistical optimization of process parameters for the production of vanillic acid by solid-state fermentation of groundnut shell waste using response surface methodology. Journal of Chemical Technology & Biotechnology. 2011;86(12):1535-1541.
43. Dutta S, Mandal NK, Ray L. Statistical Optimisation of Fermentation Process Parameters for the Production of an Alkaline, Thermostable Lipase from an Isolated Strain of Bacillus cereus C7. Indian Chemical Engineer. 2010;52(2):106-115.
44. Liu Q. Statistical optimization of recycled-paper enzymatic hydrolysis for simultaneous saccharification and fermentation via central composite design. Applied biochemistry and biotechnology. 2010;160:604-612.
45. Baş D, Boyacı İH. Modeling and optimization I: Usability of response surface methodology. Journal of food engineering. 2007;78(3):836-845.
46. Sattayasamitsathit S, Prasertsan P, Methacanon P. Statistical optimization for simultaneous production of 1, 3-propanediol and 2, 3-butanediol using crude glycerol by newly bacterial isolate. Process Biochemistry. 2011;46(2):608-614.
47. Kiran B. Statistical optimization using central composite design for biomass and lipid productivity of microalga: a step towards enhanced biodiesel production. Ecological Engineering. 2016;92:73-81.
48. Bhattacharya S. Improvement of ε-polylysine production by marine bacterium Bacillus licheniformis using artificial neural network modeling and particle swarm optimization technique. Biochemical engineering journal. 2017;126:8-15.
49. Liyanaarachchi VC. Artificial neural network (ANN) approach to optimize cultivation conditions of microalga Chlorella vulgaris in view of biodiesel production. Biochemical Engineering Journal. 2021;173:108072.
50. Betiku E. Mathematical modeling and process parameters optimization studies by artificial neural network and response surface methodology: A case of non-edible neem (*Azadirachta indica*) seed oil biodiesel synthesis. Energy. 2014;72:266-273.
51. Desai K. Optimization of fermentation media for exopolysaccharide production from Lactobacillus plantarum using artificial intelligence-based techniques. Process Biochemistry. 2006;41(8):1842-1848.
52. Igwegbe CA. Bio-coagulation-flocculation (BCF) of municipal solid waste leachate using picralima nitida extract: RSM and ANN modeling. Current Research in Green and Sustainable Chemistry. 2021;4:100078.
53. Sathish S, Vivekanandan S. Parametric optimization for floating drum anaerobic bio-digester using Response Surface Methodology and Artificial Neural Network. Alexandria Engineering Journal. 2016;55(4):3297-3307.
54. Manmai N, Unpaprom Y, Ramaraj R. Bioethanol production from sunflower stalk: application of chemical and biological pretreatments by response surface methodology (RSM). Biomass Conversion and Biorefinery. 2021;11:1759-1773.
55. Shi X.-Y. Optimization of conditions for hydrogen production from brewery wastewater by anaerobic sludge

using desirability function approach. Renewable Energy. 2010;35(7):1493-1498.
56. Gadhe A, Sonawane SS, Varma MN. Optimization of conditions for hydrogen production from complex dairy wastewater by anaerobic sludge using desirability function approach. International Journal of Hydrogen Energy. 2013;38(16):6607-6617.
57. Ousaadi MI. Valorization and optimization of agro-industrial orange waste for the production of enzyme by halophilic Streptomyces sp. Environmental Research. 2021;201:111494.
58. Betiku E, Taiwo AE. Modeling and optimization of bioethanol production from breadfruit starch hydrolyzate vis-à-vis response surface methodology and artificial neural network. Renewable energy. 2015;74:87-94.
59. Fei NC, Mehat NM, Kamaruddin S. Practical applications of Taguchi method for optimization of processing parameters for plastic injection molding: a retrospective review. International Scholarly Research Notices; c2013.
60. Zaid H. Optimization of Different Chemical Processes Using Response Surface Methodology-A Review: Response Surface Methodology. Journal of Engineering and Sustainable Development. 2022;26(6):1-12.
61. Patel N, Parihar PL, Makwana JS. Parametric optimization to improve the machining process by using Taguchi method: A review. Materials Today: Proceedings. 2021;47:2709-2714.
62. Nalbant M, Gökkaya H, Sur G. Application of Taguchi method in the optimization of cutting parameters for surface roughness in turning. Materials & design. 2007;28(4):1379-1385.