Design of Experiments in Industry: Classical to Modern Methodologies, Bias Control, and Emerging Trends

Kumarjit Pathak
AIMS
[1]Kumarjit.pathak@theaims.ac.in

Abstract*— Design of Experiments (DOE) constitutes a foundational methodology for systematic investigation of complex systems across diverse industrial applications. This comprehensive review synthesizes classical and modern experimental design approaches, providing practitioners with a structured framework for methodology selection and implementation. We present a systematic comparison of nineteen distinct design methodologies spanning factorial designs, response surface methods, space-filling approaches, optimal designs, and emerging adaptive frameworks. Two comprehensive comparison tables characterize these methodologies across multiple dimensions including run efficiency, estimable effects, model compatibility, bias mitigation strategies, and practical constraints. The review establishes mathematical foundations underlying major design classes, documents industrial applications across manufacturing, pharmaceutical development, quality engineering, and computer experiments, and analyzes trade-offs between run economy and information content. Key findings demonstrate that no single design dominates across all contexts—optimal selection depends on experimental objectives, factor dimensionality, model assumptions, and operational constraints. Classical factorial and response surface designs remain essential for physical experiments with moderate factor counts, while space-filling designs prove indispensable for computer experiments and flexible nonparametric modeling. Modern innovations including Definitive Screening Designs achieve remarkable efficiency by unifying screening and optimization, while adaptive Bayesian approaches enable intelligent sequential experimentation. Integration with machine learning and autonomous systems represents a significant frontier, promising accelerated innovation through closed-loop experimental learning. This review provides both a comprehensive reference for current practice and identification of emerging research directions in high-dimensional experimentation, multi-objective optimization, robustness quantification, and reproducible experimental science. The structured comparison framework and methodological synthesis equip researchers and industrial practitioners to navigate the rich landscape of experimental design options, making informed decisions that maximize information extraction while respecting resource constraints and operational realities.

Index Terms— *Design of Experiments, Factorial Designs, Response Surface Methodology, Optimal Design, Space-Filling Designs, Experimental Optimization, Industrial Statistics, Bayesian Optimization, Machine Learning Integration, Adaptive Experimentation.*

I. INTRODUCTION

Design of Experiments (DOE) has evolved into a cornerstone of industrial statistics, offering a systematic framework for understanding, optimizing, and controlling complex processes in manufacturing, engineering, and technological innovation. By enabling experimenters to assess the simultaneous effects of multiple input variables, DOE drastically improves the efficiency, reliability, and interpretability of industrial tests, surpassing traditional one-factor-at-a-time or purely observational approaches. The modern industrial landscape—characterized by increasing product complexity, stringent quality standards, and a growing need for rapid innovation—demands robust methodologies that can extract maximal information while minimizing resource expenditure and experimental bias.

Since its origins in agricultural research in the early 20th century, DOE has become foundational to disciplines ranging from chemical engineering and biotechnology to electronics, materials science, and large-scale manufacturing. The powerful principles developed by Fisher—factorial structure, blocking, randomization, and replication—not only laid the groundwork for the classical designs still widely used today but also catalyzed the development of advanced methodologies tailored for modern industrial systems with high-dimensional, non-linear, or highly constrained design spaces.

The fundamental objective of experimental design is to establish causal relationships between input factors (independent variables) and responses (dependent variables) while controlling for sources of variability that might obscure these relationships. In industrial contexts, experiments are often expensive, time-consuming, or operationally disruptive, making efficient experimental planning not merely desirable but economically essential. A well-designed experiment can provide definitive answers about factor effects, interactions, and optimal settings with minimal runs, whereas poorly planned experimentation may consume substantial resources while yielding ambiguous or misleading conclusions

Despite this methodological richness, practitioners often face challenges in selecting appropriate designs for specific applications. The proliferation of design options, each with distinct assumptions, strengths, and limitations, necessitates structured guidance for matching methodology to experimental context. Furthermore, the integration of experimental design with modern computational statistics, machine learning, and autonomous systems is creating new opportunities and challenges that extend beyond traditional DOE frameworks

This review synthesizes a broad spectrum of DOE methodologies as applied to industrial experimentation. Beginning with classical full and fractional factorial designs and Plackett–Burman screening, it advances to discuss response surface designs and recent innovations such as space-filling and clustering-based designs for computer simulations and adaptive experimentation. Special focus is given to methods for identifying and mitigating biases (including confounding, selection bias, measurement error, and temporal drift), as these are of critical importance for ensuring the generalizability and robustness of industrial findings.

The aims of this review are threefold:

Run	Factor A	Factor B	Factor C
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

Fig. 1. 2^3 factorial design example.

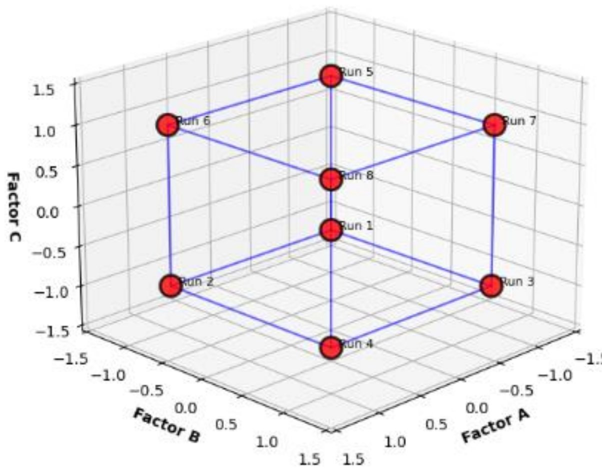


Fig. 2. Geometric representation (Cube of Factor Space).

To provide a consolidated, comparative guide to established and emerging experimental design methodologies

To highlight key mathematical foundations, statistical properties, and comparative strengths and weaknesses of each method

To offer practical recommendations and decision criteria for selecting designs suited to specific industrial goals, constraints, and future integration with predictive modeling and AI systems

By integrating extensive literature, real-world case studies, and structured comparison tables, this article intends to serve both as a reference for practitioners and a springboard for further methodological research. The subsequent sections present technical foundations, comparative evaluation, industrial application cases, and emerging trends that define the frontier of DOE for the decades ahead. This structured presentation aims to serve multiple audiences: industrial practitioners seeking guidance on design selection and implementation, academic researchers pursuing methodological advances, and students learning experimental design principles. The comprehensive comparison tables, mathematical frameworks, and application examples collectively provide both a reference for current practice and a foundation for advancing the frontiers of experimental methodology

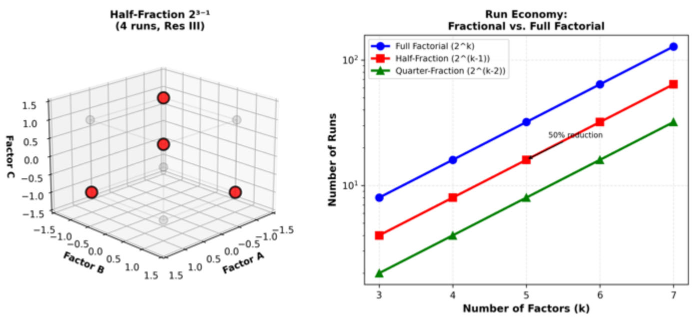
.

II. Literature Review

A. Classical DOE Methodologies

The foundation of Design of Experiments (DOE) was established by R.A. Fisher in the early 20th century, who introduced essential principles such as randomization, replication, and blocking to ensure unbiased and efficient estimation of experimental effects [1]. Classical factorial designs, which investigate all possible combinations of multiple factors, have played a pivotal role in enabling researchers and industrial practitioners to estimate main effects and interactions systematically and comprehensively [1,4]. However, the exponential increase in the number of runs required with increasing factors renders full factorial designs impractical for large, complex systems commonly encountered in modern industrial applications [18].

To mitigate this burden, fractional factorial designs were developed, allowing experimenters to estimate key effects using a fraction of the full factorial runs by confounding higher-order interactions, which are typically assumed negligible [3],[18]. This tradeoff between run size and information completeness is a hallmark of classical DOE strategy. Plackett–Burman designs further enhance screening efficiency by enabling the estimation of main effects with very few runs, albeit with the strong assumption that interactions do not substantially affect responses [3].



Resolution IV Aliasing
(Main effects clear of 2FI)

Effect	Aliased With
A	2FI (e.g., BCD)
B	2FI (e.g., ACD)
AB	CD
AC	BD

Resolution V Aliasing
(Main & 2FI clear)

Effect	Aliased With
A	3FI+ (e.g., BCDE)
AB	3FI+ (e.g., CDE)
ABC	DE

Fig. 2. Fractional Factorial

The design of these experiments relies fundamentally on orthogonality and balance, mathematical properties that ensure uncorrelated estimates of factor effects and minimal variance of estimators [1],[18]. These classical designs remain central to DOE practice largely because they provide clear statistical interpretation and are supported by a mature analytical framework centered around analysis of variance (ANOVA) and regression methods Citation 18.

Building upon these foundations, response surface methodology (RSM) was developed to explore nonlinearities and optimize process parameters by fitting polynomial models to experimental data. Designs such as central composite and Box–Behnken are commonly used for such purposes, allowing efficient estimation of quadratic effects and interaction terms [4], [5],[12], [13]. These methodologies extend the classical DOE framework to accommodate curvature in responses and enable systematic optimization.

Despite the rise of more modern methodologies, classical DOE techniques are still widely taught, used in industry, and implemented in software due to their solid theoretical basis, ease of interpretation, and broad applicability across industries [1], [18].

Screening designs are specifically developed for situations where experimenters face a large number of potential factors but need to identify efficiently which factors significantly influence the response [2],[3]. This is a common scenario in early-stage industrial research, product development, and process optimization, where resource constraints demand rapid identification of critical variables before investing in more detailed studies [18].

Plackett–Burman designs represent a classical screening approach, allowing the estimation of main effects for N-1 factors using N runs where N is a multiple of 4 [3]. These designs sacrifice information on interactions to achieve maximum economy in the number of experimental runs. The underlying assumption—that main effects dominate, and interactions are negligible—is reasonable in many industrial

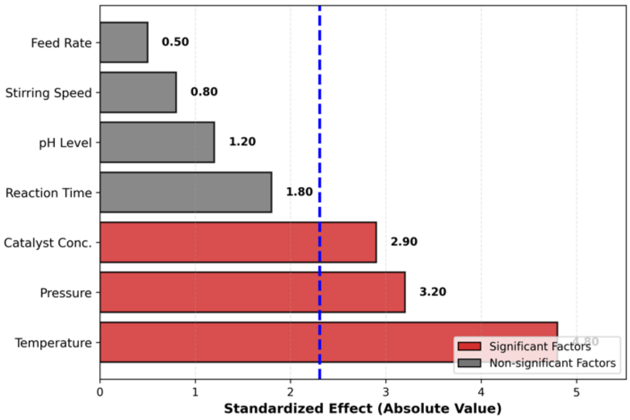


Fig. 4. Pareto chart for factor screening (illustrative example)

contexts, especially during initial factor screening [3],[18].

Fractional factorial designs, particularly those with resolution III, IV, and V, offer another powerful screening tool [18],[19]. Resolution III designs allow main effects to be estimated free from one another but confound them with two-factor interactions. Resolution IV designs provide main effects clear of two-factor interactions but confound some two-factor interactions with each other [18], [19]. The choice of design resolution depends on the experimenter's assumptions about the importance of interactions and the acceptable level of confounding [19].

Modern screening methodologies have evolved to include definitive screening designs, which efficiently estimate main effects, two-factor interactions, and quadratic effects with fewer runs than traditional approaches [20]. These designs are particularly valuable when both screening and optimization objectives must be met within a single experiment.

The statistical analysis of screening experiments often employs effect sparsity principles, normal probability plots, and Pareto charts to identify the most influential factors visually and statistically [2],[18]. Advanced methods also incorporate Bayesian variable selection and regularization techniques (such as LASSO) to enhance factor identification in high-dimensional settings [17].

Screening designs remain indispensable in industrial practice for their ability to reduce experimental costs dramatically while maintaining statistical rigor in identifying influential factors for subsequent detailed investigation Citation 2, Citation 3, Citation 18.

B. *Response Surface Methodology and Optimization*

Response Surface Methodology (RSM) emerged from the

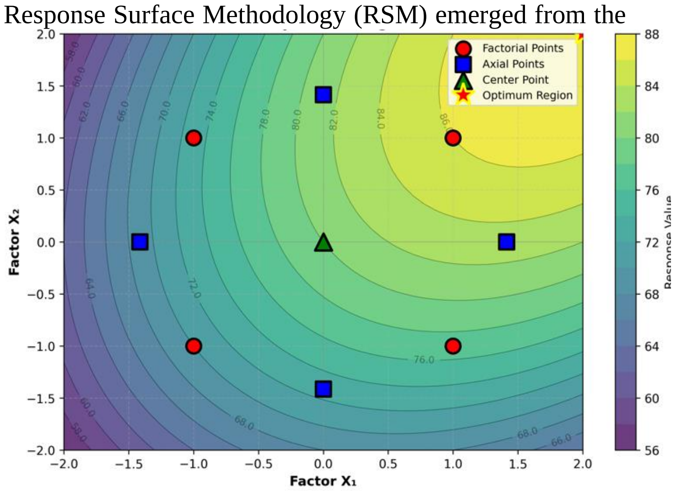


Fig. 4. Contour Plot with CCD design points and optimum regions).

foundational work of Box and Wilson in 1951 as a systematic approach to explore relationships between multiple input factors and one or more response variables, particularly when curvature in the response function is suspected [4]. Unlike factorial designs that primarily estimate linear main effects and interactions, RSM designs enable the fitting of second-order polynomial models that capture quadratic effects and interaction curvature, which are essential for identifying optimal operating conditions [4],[5],[13].

Central Composite Designs (CCD) represent the most widely implemented RSM design in industrial practice [4],[13]. A CCD combines factorial or fractional factorial points with axial (star) points positioned at distance α from the design center, plus center point replicates to estimate pure error [4],[13]. The choice of α determines design properties such as rotatability (uniform prediction variance at all points equidistant from the center) and orthogonality (uncorrelated parameter estimates). Face-centered, inscribed, and circumscribed variants of CCD provide flexibility when factor levels must respect physical or safety constraints.

Box-Behnken Designs (BBD) offer an alternative RSM approach that requires fewer runs than CCD for three or more factors by placing design points at the midpoints of edges of the factor space rather than at corners [5]. This feature makes BBD particularly valuable when extreme factor level combinations are infeasible, expensive, or potentially hazardous. BBD maintains rotatable or near-rotatable properties and efficiently estimates quadratic models, though the absence of corner points may limit extrapolation beyond the experimental region [5].

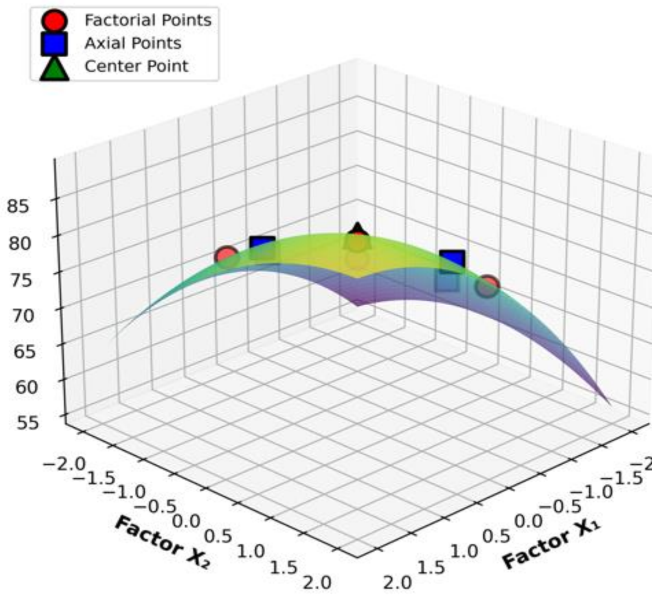


Fig. 3. 3D response surface with CCD points

The analysis of RSM experiments typically proceeds through sequential stages: fitting an initial first-order model to screen factors, conducting steepest ascent (or descent) experiments to move toward the optimal region, and finally fitting a full second-order model to characterize the response surface near the optimum [2],[4],[18]. Canonical analysis and ridge analysis are advanced techniques used to characterize the nature of stationary points (maximum, minimum, or saddle point) and to identify optimal factor settings under constraints.

Modern extensions of classical RSM include designs for multiple responses with potentially conflicting objectives, robust parameter design to minimize sensitivity to noise factors, and integration with computational optimization algorithms for large-scale industrial systems [11],[12],[16]. The development of D-optimal and other computer-generated optimal designs has further enhanced RSM capability by accommodating irregular experimental regions, mixture constraints, and non-standard model structures that classical geometric designs cannot address efficiently.

C. *Space-Filling Designs and Computer Experiments*

The rise of computer simulation and deterministic computational models has fundamentally transformed experimental design practice, giving rise to space-filling designs that prioritize uniform coverage of the factor space over the orthogonality properties central to classical DOE [6],[10]. Unlike physical experiments where random error necessitates replication and blocking, computer experiments are typically deterministic, producing identical outputs for identical inputs, which eliminates the need for classical error estimation but introduces different design considerations.

Latin Hypercube Sampling (LHS), introduced by McKay, Beckman, and Conover in 1979, represents a foundational space-filling method that stratifies the sample space to ensure each factor level is represented equally across runs [6]. LHS guarantees that the projection of the design onto any single factor axis is uniformly distributed, providing variance reduction compared to simple random sampling and enabling efficient exploration of high-dimensional factor spaces [10]. Enhancements such as maximin LHS (maximizing the minimum distance between points) and orthogonal LHS (ensuring low correlation between factors) further improve design properties for different modeling objectives.

Quasi-random sequences, particularly Halton and Sobol sequences, offer deterministic alternatives to random sampling with superior uniformity properties measured by discrepancy theory [7],[14]. These low-discrepancy sequences fill space more uniformly than pseudorandom methods, especially in moderate dimensions, and possess the valuable property of sequential extendability, allowing experimenters to add points without regenerating the entire design. Halton sequences, constructed using prime number bases, provide reproducible designs suitable for sensitivity analysis and uncertainty quantification in simulation studies.

Distance-based optimal designs, including maximin and minimax distance criteria, explicitly optimize the geometric distribution of design points to avoid clustering while ensuring adequate coverage [10],[20]. Maximin designs maximize the minimum pairwise distance between points, preventing under-sampling of any region, whereas minimax designs minimize the maximum distance from any point in the factor space to the nearest design point. These designs are particularly valuable for building surrogate models (metamodels) such as Gaussian process emulators and radial basis function approximations that depend critically on point spacing [10],[20].

Cluster-based designs generated through k-means or similar algorithms on large candidate sets approximate centroidal Voronoi tessellations, which represent theoretically optimal space-filling configurations [10]. By partitioning a dense candidate point cloud into clusters and selecting cluster centroids as design points, these methods achieve near-optimal space-filling properties while accommodating complex constraints and irregular experimental regions [20].

The selection among space-filling designs depends on modeling objectives, dimensionality, computational budget, and whether sequential experimentation is anticipated [6], [10], [14], [20]. Space-filling designs have become indispensable in engineering fields requiring expensive simulations, including computational fluid dynamics, finite element analysis, and complex systems modeling where each simulation may require hours or days of computation time.

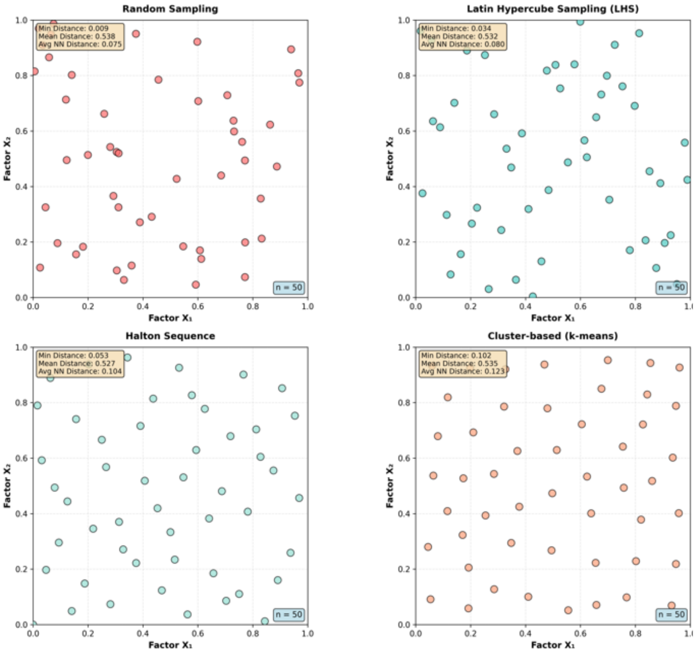


Fig. 4. Comparison of Space filling Designs Methods (2D Illustrative Example)

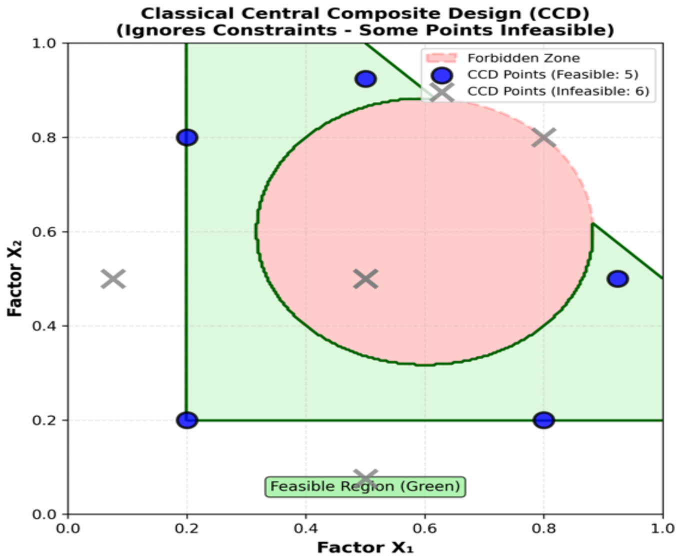
D. Optimal Designs and Model Based Approach

Optimal design theory provides a rigorous mathematical framework for constructing experimental designs that maximize specific statistical criteria related to parameter estimation precision, prediction accuracy, or model discrimination [12],[17]. Unlike classical factorial and response surface designs that follow fixed geometric patterns, optimal designs are algorithmically generated to optimize explicit objectives defined through the information matrix of the assumed model, enabling unprecedented flexibility in addressing complex experimental constraints.

D-optimal designs maximize the determinant of the Fisher information matrix, which is equivalent to minimizing the volume of the confidence ellipsoid for parameter estimates [12],[17]. This criterion produces designs that provide the most precise parameter estimates on average and has become the most widely implemented optimality criterion in industrial practice due to its general applicability and computational tractability [12]. D-optimal designs excel when the experimental region is irregular, when mixture constraints apply, when factors have differing numbers of levels, or when certain factor combinations are infeasible or prohibitively expensive [17].

A-optimal designs minimize the average variance of parameter estimates by minimizing the trace of the inverse information matrix [12],[17]. While A-optimality focuses on parameter estimation precision similarly to D-optimality, the two criteria can produce different designs depending on model structure and the relative importance of different parameters. A-optimal designs may be preferred when all model parameters are of equal scientific interest and when prediction variance at specific points is a secondary concern [17].

I-optimal designs and V-optimal designs (also called G-optimal) prioritize prediction precision over parameter estimation precision [12]. I-optimal designs minimize the average prediction variance across the design space,



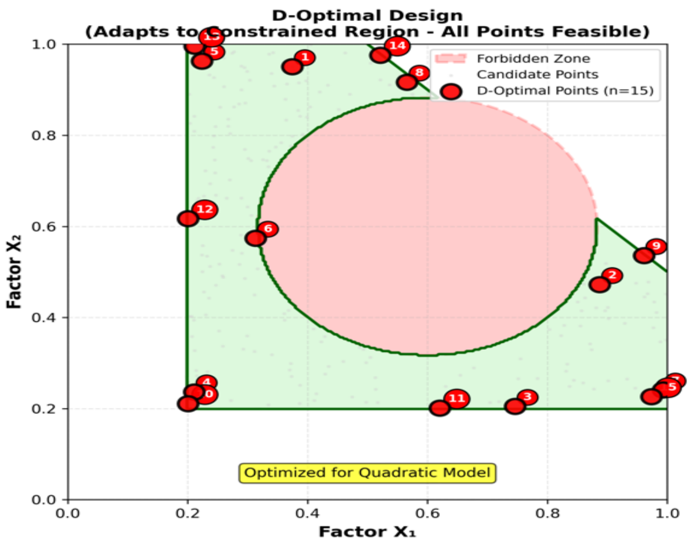


Fig. 4. CCD vs D-Optimal

making them particularly suitable when the primary experimental objective is building accurate predictive models rather than isolating individual factor effects. These designs have gained prominence in applications involving surrogate modeling, emulation of computer experiments, and machine learning contexts where prediction accuracy dominates parameter interpretability [12],[20].

The construction of optimal designs typically proceeds through computer algorithms that search over a candidate set of potential experimental points, iteratively selecting or exchanging points to improve the chosen optimality criterion

[12],[17]. Exchange algorithms, such as the Fedorov exchange algorithm and coordinate exchange, balance computational efficiency with design quality [12]. Modern software implementations have made optimal design generation accessible to practitioners, though the model-dependent nature of optimal designs requires careful specification of the assumed response model structure.

Bayesian optimal design extends classical optimal design theory by incorporating prior information about model parameters and explicitly accounting for uncertainty in both the model and parameters [17]. Bayesian approaches use expected utility criteria that average over the prior distribution, producing designs that balance information gain against prior knowledge. These methods are particularly valuable in sequential experimentation where each stage of data collection informs the design of subsequent stages, enabling adaptive designs that learn efficiently from accumulating evidence [16],[17].

The integration of optimal design principles with modern computational capabilities has enabled designs for increasingly complex scenarios, including designs for multiple competing models, designs for parameter subset selection, and designs that explicitly account for model uncertainty. This flexibility makes optimal design methodology essential for contemporary industrial experimentation where standard catalog designs often prove inadequate [12],[17].

E. Modern Advances and Emerging Methodologies

The past two decades have witnessed significant innovations in experimental design methodology driven by computational advances, machine learning integration, and the increasing complexity of industrial systems [16].[20]. These modern approaches extend classical DOE principles while addressing limitations inherent in traditional methods, particularly for high-dimensional problems, expensive experiments, and dynamic systems.

Definitive Screening Designs (DSD), introduced by Jones and Nachtsheim in 2011, represent a breakthrough in efficient factor screening by providing a three-level design that can estimate main effects, two-factor interactions, and quadratic effects with remarkable economy [20]. For k factors, a DSD requires only 2k+1 runs while ensuring that main effects are not aliased with each other or with any two-factor interaction, a property not achievable with traditional screening designs. The key innovation lies in the strategic construction that exploits effect sparsity assumptions—the principle that only a few factors and interactions typically dominate response variation. DSDs have rapidly gained industrial adoption as they unify screening and optimization objectives within a single efficient design structure [20].

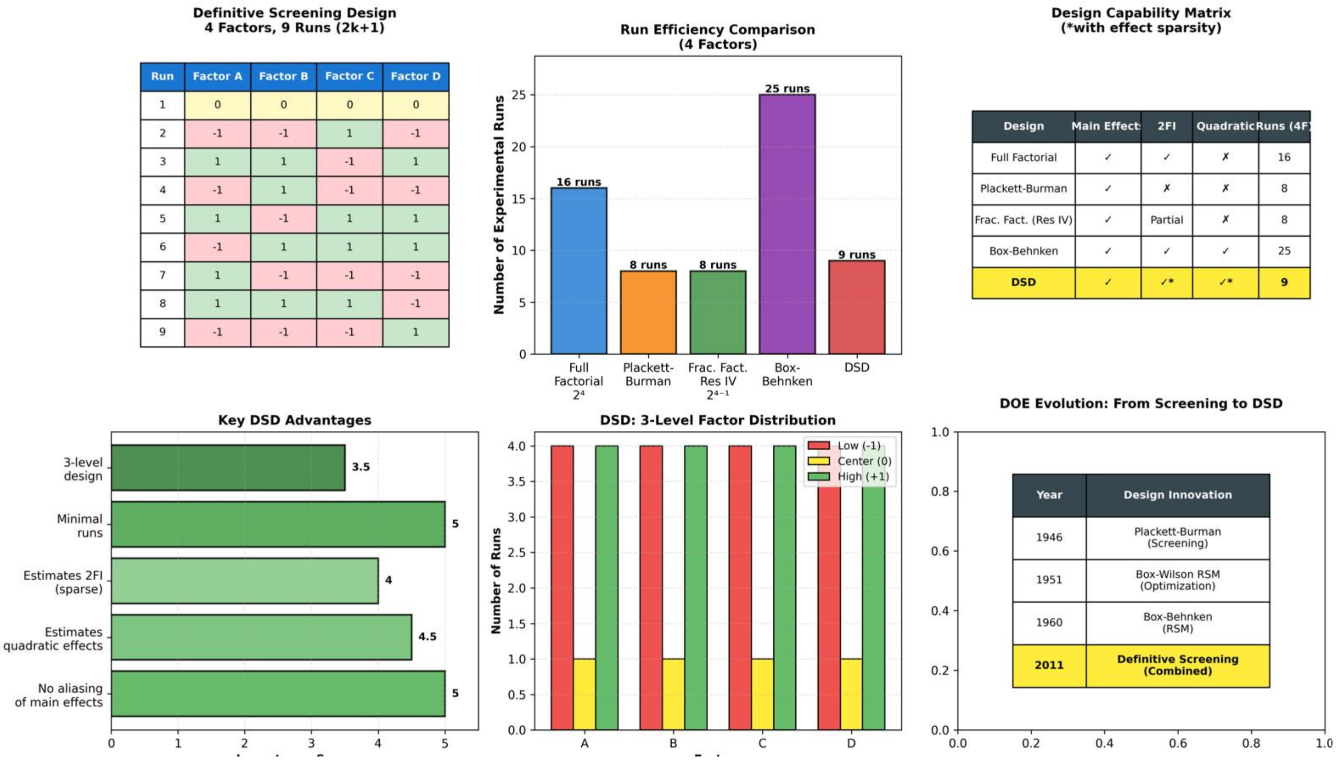


Fig. 4. Definitive Screening Designs (DSD – an Illustrative example

Adaptive and sequential experimental designs leverage Bayesian updating and active learning principles to intelligently select subsequent experiments based on information gained from previous runs [16],[17],[20]. These methods are particularly valuable when experiments are expensive, time-consuming, or when the response surface structure is uncertain [17]. Bayesian optimization algorithms employing Gaussian process surrogates with acquisition functions such as expected improvement or upper confidence bound have become standard tools for optimizing expensive black-box functions in engineering, drug discovery, and materials science [16],[17],[20].

Machine learning integration with DOE represents an emerging frontier where classical experimental design principles merge with modern statistical learning methods [16],[20]. Neural networks, random forests, and gradient boosting machines offer flexible nonparametric alternatives to polynomial response surface models, capable of capturing complex nonlinear relationships without pre-specifying functional forms. However, these models require careful consideration of design point placement since their performance depends critically on training data coverage and density. Space-filling designs naturally complement machine learning approaches by ensuring adequate sampling across the input space [10], [16],[20].

Split-plot and multi-stratum designs have evolved to accommodate the hierarchical and nested structures common in modern industrial settings where some factors are difficult or expensive to vary [2], [18]. Extensions to more complex restriction patterns, including strip-plot and split-split-plot designs, provide frameworks for experiments where multiple levels of randomization constraints exist simultaneously. Proper analysis of these designs using mixed-effects models and restricted maximum likelihood estimation ensures valid inference despite departure from classical complete randomization assumptions [2], [18].

Robust parameter design and Taguchi methods emphasize minimizing response variability and sensitivity to uncontrollable noise factors, representing a shift from mean response optimization to variance reduction and robustness [11]. While traditional Taguchi orthogonal arrays and signal-to-noise ratio analysis have received methodological criticism, the underlying principles of robust design—using control factors to reduce the impact of noise factors—remain valuable. Modern approaches combine robust design concepts with response surface methodology and dual-response optimization to achieve both target mean performance and minimal variability [11],[18].

The convergence of experimental design with computational optimization, uncertainty quantification, and artificial intelligence is creating new paradigms for industrial experimentation [16],[20]. Autonomous experimentation systems that integrate DOE principles with real-time sensor data, closed-loop control, and reinforcement learning represent the frontier of smart manufacturing and Industry 4.0 initiatives. These systems promise to accelerate innovation cycles while reducing experimental costs through intelligent, data-driven exploration of complex process and product design spaces [16],[20].

III. METHODOLOGICAL FOUNDATION

Let's look at mathematical underpinnings of major experimental design methodologies, establishing the theoretical framework that guides design construction, analysis, and interpretation in industrial applications

A. *Factorial Designs*

The fundamental model for a two-level factorial experiment with "k" factors can be expressed as:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \sum_{i<j<l} \beta_{ijl} x_i x_j x_l + \cdots + \mathcal{E}$$

Where 'y' is the response $x_i \in \{-1, +1\}$ represents the coded level of factor 'i', β_i are the main effects coefficient, β_{ij} are two-factor interaction coefficients, and $\mathcal{E}$ is random error [1], [18]. The orthogonality property of full factorial designs ensures that the design matrix X satisfies $\mathbf{X^{TX}} = n\mathbf{I}$ where n is the number of runs, guaranteeing uncorrelated parameter estimates.

For fractional factorial designs with resolution R, the defining relation determines the aliasing structure Citation 3, Citation 18, Citation 19. A design with generator I=ABCD creates alias chains where each effect is confounded with effects obtained by multiplying by ABCD [18],[19]. The resolution characterizes the minimum length of words in the defining relation: Resolution III designs confound main effects with two-factor interactions, Resolution IV designs confound main effects with three-factor interactions but two-factor interactions with other two-factor interactions, and Resolution V designs provide clear estimation of all main effects and two-factor interactions

B. *Response Surface Design Methodology*

The second-order polynomial model underlying response surface designs takes the form:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j} \beta_{ij} x_i x_j + .. + \mathcal{E}$$

This model contains $p = 1 + k + k + (2k) = \frac{(k+1)(k+2)}{2}$ parameters Central Composite Designs provide efficient estimation by combining 2^k or $2^{k\text{-}p}$ factorial points, 2k axial points at distance α from the center, and N_c center points.

The rotatable property is achieved when $\alpha = (2^k)^{1/4}$ for a full factorial core, ensuring that prediction variance $Var(\hat{y}(x))$ depends only on the distance from the design center. $\mathrm{Var}(\hat{y}(x)) = \sigma^2 f(d)$ where $d = \sqrt{\sum_{i=1}^{k} x_i^2}$

This property provides uniform prediction quality in all directions from the center. Box-Behnken designs achieve similar objectives with fewer runs by placing points at midpoints of edges of the *k*-dimensional factor space, requiring approximately $k^2 + k + n_c$ runs. The absence of corner points makes BBDs safer when extreme factor combinations are hazardous[5].

C. *Optimal Design Method*

The Fisher information matrix for a design ε with model matrix X is given by:

$$\mathbf{M}(\xi) = \mathbf{X^{TX}} = \sum_{i=1}^{n} w_i \mathbf{f}(x_i)\mathbf{f}(x_i)^T$$

where $f(x_i)$ is the vector of model terms evaluated at design point x_i , and w_i is the weight (proportion of observations) at that point [12], [17].

D-optimality maximizes $\det(\mathrm{M}(\varepsilon))$, minimizing the volume of the confidence ellipsoid for parameter estimates. The D-efficiency of a design relative to the optimal design is:

$$\mathrm{Eff}_D = \left(\frac{\det\big(\mathbf{M}(\xi)\big)}{\det\big(\mathbf{M}(\xi^*)\big)}\right)^{1/p} \times 100\%$$

Where ε^* is the D-optimal design and p is the number of parameters. A-optimality minimizes trace $\min \operatorname{trace}\,(\mathbf{M}(\xi)^{-1})$, which is equivalent to minimizing the average variance of parameter estimates. I-optimality minimizes the average prediction variance over the design region:

$$\min \int_{\mathcal{X}} \mathrm{Var}\big(\hat{y}(x)\big)\ dx = \sigma^2 \int_{\mathcal{X}} \mathbf{f}(x)^T \mathbf{M}(\xi)^{-1} \mathbf{f}(x)\ dx$$

making it particularly suitable for prediction-focused applications.

D. *Space-Filling Design Criteria*

For computer experiments, space-filling quality is assessed through discrepancy measures [6],[10],[14]. The L_2-discrepancy quantifies how uniformly points fill the design space compared to a uniform distribution Citation 14. For Latin Hypercube Sampling, the design ensures marginal uniformity through stratification

$$x_{ij} = \frac{\pi_j(i) - U_{ij}}{n}$$

Where π_j is a random permutation of {1,2,3, …n} for factor j and $U_{ij} \sim$ Uniform (0,1) . Maximin distance designs maximize the minimum pairwise Euclidean distance: $\phi_{\text{maximin}} = \max_{\xi} \min_{i \neq j} |\mathbf{x_i} - \mathbf{x_j}|$ while minimax designs minimize the maximum distance from any point in the design region to the nearest design point [10], [20]. These criteria ensure no region is under-sampled.

E. *Bias Mitigation Through Randomization and Blocking*

Randomization serves as the foundation for valid statistical inference by distributing the effects of uncontrolled variables randomly across experimental units [1],[2]. When complete randomization is infeasible, blocking partitions experimental units into homogeneous groups. The blocked design model includes block effects: $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$ where τ_i is the treatment effect, β_j is the block effect and ε_{ij} is random error. Proper blocking reduces unexplained variability and increases precision of treatment effect estimates.

For split-plot designs with hard-to-change factors, the model includes multiple error terms reflecting different levels of randomization:

$$y_{ijk} = \mu + \alpha_i + \delta_{ik} + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Where δ_{ik} is whole-plot error and ε_{ijk} is subplot error [2]. The presence of two error variances requires restricted maximum likelihood (REML) or mixed-effects model estimation for valid inference

IV. COMPARATIVE EVALUATION

Let's review appropriate experimental design methodology depends critically on experimental objectives, resource constraints, factor dimensionality, model assumptions, and practical considerations specific to each industrial application Citation attached_file:1. This section synthesizes the comparative analysis presented in Tables 1 and 2, providing a structured framework for design selection and highlighting the fundamental trade-offs that practitioners must navigate.

Table 1 presents a comprehensive comparison of some major design methodologies spanning classical factorial designs, response surface methods, and modern space-filling approaches. The table systematically compares designs across multiple dimensions including typical run count formulas, estimable effects, resolution or confounding properties, space-filling characteristics, compatible model types, primary applications, key advantages, limitations, and bias mitigation strategies. This structured comparison enables practitioners to quickly identify candidate designs matching their experimental context and constraints

Table 2 extends this analysis to some advanced and specialized methodologies including optimal designs, definitive screening designs, mixture designs, split-plot designs, and adaptive approaches. These methods address specific challenges that classical catalog designs cannot efficiently handle, such as irregular experimental regions, compositional constraints, hierarchical randomization restrictions, and sequential experimentation under uncertainty.

A primary discriminator among design methodologies is the relationship between run count and information gained. Full factorial designs provide complete information about all main effects and interactions up to order k but require exponentially growing runs (2^k), making them practical only for k ≤ 4 factors in most industrial settings. In contrast, fractional factorial designs achieve dramatic run reduction through strategic aliasing, with a 2^(k-p) design reducing runs by a factor of 2^p while maintaining estimability of critical lower-order effects under sparsity assumptions.

Plackett-Burman designs represent an extreme point on the efficiency spectrum, screening up to N-1 factors in N runs (where N is a multiple of 4), but at the cost of complete confounding between main effects and two-factor interactions. This makes PB designs suitable only when interaction effects can be confidently assumed negligible or when they serve as preliminary screening before follow-up experimentation.

Definitive Screening Designs achieve a remarkable balance by requiring only 2k+1 runs for k factors while estimating main effects, selected two-factor interactions, and quadratic effects with no aliasing of main effects among themselves or with two-factor interactions [20]. This efficiency has driven rapid adoption in pharmaceutical development and other time-constrained applications where the traditional sequential approach of screening followed by optimization proves too resource-intensive

Design Method	Typical Run Count	Factors Estimable	Resolution/ Confounding	Space-Filling	Model Type	Primary Application	Key Advantages	Key Limitations	Bias Mitigation
Full Factorial	2^k for k factors (e.g., 8, 16, 32)	All main effects, all interactions up to order k	No confounding; Resolution ∞	Poor (corner points only)	Linear, first-order	Small-scale experiments (≤4 factors); interaction exploration	Complete information; unbiased estimates; no aliasing	Exponential run growth; impractical for k>5	Randomization, blocking, replication
Fractional Factorial	$2^{(k-p)}$ (e.g., 8, 16, 32)	Main effects; selected interactions	Resolution III–V; some aliasing	Poor (subset of corners)	Linear, first-order	Screening (6–15 factors); early-stage optimization	Drastically reduced runs; efficient screening	Confounding/aliasing of effects; assumption of negligible interactions	High-resolution designs; fold-over; sequential augmentation
Plackett–Burman	Multiples of 4 (12, 20, 24, 28, etc.)	Main effects only	Resolution III; heavy confounding	Poor (structured, non-factorial)	Main effects only	High-dimensional screening (>10 factors)	Extreme run efficiency; up to N-1 factors in N runs	Cannot estimate interactions; severe aliasing	Confirmatory runs; dummy factors; follow-up experiments
Central Composite Design (CCD)	$2^k + 2k + n_c$ (e.g., 15–30)	Main, quadratic, two-factor interactions	Minimal for 2nd-order model	Moderate (axial + factorial)	Second-order polynomial (quadratic)	Process optimization (3–5 factors); RSM	Efficient quadratic modeling; rotatability options; well-studied	Star points may exceed factor bounds; not space-filling	Blocking for star/factorial; randomization; center point replication
Box–Behnken Design (BBD)	$2k^2 + 2k + 1$ (e.g., 13, 25, 41)	Main, quadratic, two-factor interactions	Minimal for 2nd-order model	Moderate (edge midpoints + center)	Second-order polynomial (quadratic)	Optimization (3–7 factors); avoiding extreme corners	Fewer runs than CCD; avoids corner extremes; rotatable	No corner points; limited for 3+ factor levels	Randomization; center point replication
Latin Hypercube Sampling (LHS)	Flexible (user-defined N)	Model-free; all factors uniformly sampled	No confounding (model-agnostic)	Good (stratified in each dimension)	Non-parametric; surrogate models (GP, neural nets)	Computer experiments; global sensitivity analysis	Flexible sample size; good 1D projections; variance reduction	Random gaps possible; may cluster in high dimensions	Optimized LHS (maximin, orthogonal)
Halton Sequence	Flexible (deterministic sequence)	Model-free; quasi-random coverage	No confounding	Excellent (low-discrepancy)	Non-parametric; surrogate models	Simulation-based DOE; sequential experimentation	Deterministic; reproducible; sequentially extendable; superior uniformity	May show patterns in high dimensions; not strictly LHS	Deterministic uniformity ensures unbiased coverage
Random k-Means Cluster	Flexible (N clusters from M>>N samples)	Model-free	No confounding	Excellent (CVT-like)	Non-parametric; surrogate models	Space-filling for complex domains; adaptive sampling	Approximates optimal space-filling; handles constraints; flexible	Computationally intensive for large M; requires re-clustering for expansion	Uniform random initial sampling; multivariate uniformity

Table1. Classical and Space-Filling DOE Methods comparison

A. *Model Compatibility and Assumption Requirement*

Design selection must align with the assumed or desired response model structure [4],[5],[12],[13],[17]. Two-level factorial designs (full and fractional) are inherently limited to first-order linear models and cannot detect curvature unless center points are added for lack-of-fit testing [1],[8]. When quadratic effects are anticipated, response surface designs (CCD and Box-Behnken) become necessary, requiring three or more factor levels to enable second-order polynomial fitting.

Space-filling designs (LHS, Halton, Sobol, cluster-based) are explicitly model-agnostic, making no assumptions about functional form and instead prioritizing uniform coverage for subsequent flexible modeling approaches including Gaussian processes, neural networks, or other machine learning methods [6],[7],[10],[14],[20]. This model-free philosophy proves essential for computer experiments where response surfaces may exhibit complex nonlinearity not captured by low-order polynomials.

Optimal designs (D-, A-, I-optimal) maximize specific statistical criteria for a pre-specified model, providing maximum efficiency when the assumed model is correct but potentially poor performance under model misspecification [12],[17]. The model-dependent nature of optimal designs necessitates careful specification and validation, though their flexibility in handling constraints and irregular regions often outweighs this limitation in practice.

B. *How to Handle Practical Constraints*

Industrial experiments frequently encounter constraints that standard catalog designs cannot easily accommodate [2],[12]. D-optimal designs excel in these scenarios by algorithmically selecting runs that maximize information content while respecting feasibility constraints, mixture constraints, or other restrictions on the factor space. The flexibility of D-optimal methodology enables experimentation in irregular regions where classical geometric designs would place points outside the feasible domain.

Split-plot designs address the pervasive industrial challenge of hard-to-change factors by incorporating restricted randomization structure into the experimental framework [2],. Proper analysis using mixed-effects models accounts for the presence of multiple error terms at different randomization levels, ensuring valid inference despite departure from complete randomization.

Mixture designs (simplex-lattice and simplex-centroid) provide specialized structures for experiments where factors represent component proportions constrained to sum to unity, common in formulation industries including pharmaceuticals, chemicals, food products, and materials. These designs employ Scheffé polynomial models appropriate for compositional data rather than standard factorial or response surface polynomials.

Design Method	Typical Run Count	Factors Estimable	Special Features	Model Type	Primary Application	Key Advantages	Key Limitations	Bias Mitigation
D-Optimal Design	Variable (optimized for specific N)	User-defined (main, interactions, quadratics)	Maximizes determinant of information matrix; handles constraints	Flexible (polynomial, custom)	Constrained design spaces; irregular regions; mixture experiments	Handles complex constraints; mixed factor types; irregular boundaries; optimal for specific models	Model-dependent; requires candidate set; computationally intensive; no standard run formula	Model adequacy checking; augmentation strategies; sequential refinement
A-Optimal Design	Variable (optimized)	User-defined	Minimizes average variance of parameter estimates	Flexible (polynomial, custom)	Prediction-focused applications; parameter estimation precision	Minimizes average prediction variance; handles constraints	Model-dependent; different objectives than D-optimal; requires optimization algorithms	Model validation; cross-validation; sequential augmentation
Definitive Screening Design (DSD)	2k + 1 for k factors (e.g., 7, 13, 17)	Main effects, two-factor interactions, quadratics (with sparsity)	3-level design; no confounding of main effects; economical	Second-order with interaction selection	Combined screening and optimization; early product development	Estimates main, interactions, quadratics in one design; minimal runs; no aliasing of main effects	Requires effect sparsity assumption; 3 levels required; relatively new (limited long-term validation)	Randomization; augmentation if needed; effect heredity principles
Taguchi Orthogonal Arrays	Standard arrays (L8, L9, L12, L16, L18, L27, etc.)	Main effects primarily; some interactions	Inner (control) and outer (noise) array structure	Signal-to-noise ratio optimization	Robust parameter design; quality engineering; Six Sigma	Emphasizes robustness to noise; widely adopted in industry; standard arrays available	Heavy confounding; limited interaction estimation; complex analysis for interactions	Noise factor explicit consideration; confirmation runs; robust design philosophy
Mixture Designs (Simplex-Lattice)	C(q+m-1, m) for q levels, m components	Mixture effects; blending interactions	Factors are proportions summing to 1 (or 100%)	Scheffe polynomials (mixture models)	Formulation (food, pharma, chemicals, cosmetics); blending	Handles compositional constraints naturally; specialized models for mixtures	Constrained to simplex space; cannot test zero proportions; specialized analysis	Pseudo-components for constrained regions; replication of interior points
Mixture Designs (Simplex-Centroid)	2^m - 1 for m components	Mixture effects including centroids	Includes all vertices, edge centers, and overall centroid	Scheffe polynomials	Formulation with emphasis on blend centers	More points at interior; better for nonlinear blending; includes overall center	More runs than simplex-lattice; constrained space	Interior point replication; pseudo-components
Split-Plot Designs	Depends on whole-plot and sub-plot structure	Main effects and interactions (with separate error terms)	Two levels of randomization; hard-to-change and easy-to-change factors	Mixed-effects models	Experiments with hard-to-change factors (temperature, batch, equipment)	Reflects practical constraints; reduces cost of changing hard factors; realistic industrial model	Complex analysis (two error terms); restricted randomization; requires mixed-model expertise	Proper randomization at both levels; blocking; REML analysis

Table1. Advanced and Specialized DOE Methods – a comparison

C. How to Handle Practical Constraints

All design methodologies in Tables 1 and 2 incorporate strategies for mitigating various sources of experimental bias, though approaches differ based on design structure. Randomization serves as the universal foundation for distributing effects of uncontrolled variables, with complete randomization in factorial designs, restricted randomization in split-plot designs, and deterministic (but uniform) placement in space-filling quasi-random sequences.

Blocking addresses known sources of heterogeneity by partitioning experimental units into homogeneous groups, though the specific blocking strategy depends on design type. Factorial designs allow orthogonal or partially confounded blocking with high-order interactions, while response surface designs require more complex blocking patterns to maintain model estimability.

Confounding bias represents an inherent characteristic of fractional factorial designs that must be explicitly managed through resolution selection and follow-up experimentation when needed. Tables 1 and 2 clearly identify which designs suffer from aliasing and under what conditions, enabling practitioners to make informed trade-offs between run economy and effect resolution

D. Design Selection Framework

The comprehensive comparison in Tables 1 and 2 supports a structured decision framework for design selection:

1. For screening 5+ factors with minimal runs: Use Plackett-Burman or low-resolution fractional factorial designs if interactions are negligible; use Definitive Screening Designs if interactions and curvature matter.

2. For optimization with 3-5 factors: Use Central Composite or Box-Behnken designs for quadratic modeling; consider D-optimal if constraints exist

3. For computer experiments or model-free exploration: Use Latin Hypercube Sampling for moderate dimensions; use Halton or Sobol sequences for deterministic uniformity; use cluster-based designs for complex constraints.

4. For constrained or irregular regions: Use D-optimal or other algorithmic optimal designs that respect feasibility constraints.

5. For hard-to-change factors: Use split-plot or strip-plot designs with appropriate mixed-effects analysis.

6. For mixture or compositional data: Use simplex-lattice, simplex-centroid, or optimal mixture designs .

7. For sequential or adaptive learning: Use Bayesian optimal designs or active learning approaches integrated with surrogate models.

This synthesis of design capabilities, limitations, and appropriate use cases provides practitioners with a comprehensive reference for navigating the rich landscape of experimental design methodologies documented in Tables 1 and 2.

V. INDUSTRIAL APPLICATION AND CASE STUDIES

Experimental design methodologies have been successfully deployed across diverse industrial sectors, demonstrating their practical value in solving real-world optimization, quality improvement, and process understanding challenges Citation attached_file:1. This section presents representative applications that illustrate how different DOE approaches address specific industrial objectives under varying resource constraints and operational conditions.

A. Manufacturing Process Optimization

In manufacturing environments, Central Composite Designs and Box-Behnken Designs are extensively used for process parameter optimization where curvature effects are anticipated [4], [5], [13]. A typical application involves optimizing injection molding parameters (temperature, pressure, cooling time) to minimize defect rates while maximizing throughput [4]. The quadratic modeling capability of RSM designs enables identification of optimal operating windows that balance multiple quality characteristics simultaneously. For example, a semiconductor fabrication facility might use a CCD to optimize etching parameters across three factors in 20 runs rather than exploring the entire parameter space exhaustively, achieving significant time and cost savings while establishing robust process control limits.

B. Chemical and Pharmaceutical Development

Definitive Screening Designs have gained rapid adoption in pharmaceutical formulation development where screening and optimization objectives must be achieved efficiently under tight development timelines [20]. A case study in tablet formulation involved screening seven excipients and processing parameters using a 15-run DSD, successfully identifying critical factors affecting dissolution rate and mechanical strength while simultaneously detecting quadratic relationships. This unified approach eliminated the traditional sequential screening-then-optimization paradigm, reducing total experimental effort by approximately 60% compared to conventional Plackett-Burman followed by CCD methodology.

D-Optimal designs prove indispensable in mixture experiments common to chemical industries, where component proportions must sum to 100% and standard geometric designs cannot accommodate the simplex-constrained factor space [12],[17]. Polymer blend optimization, paint formulation, and alloy composition studies routinely employ D-optimal approaches to handle irregular feasible regions while maintaining high statistical efficiency for parameter estimation .

C. Quality Engineering and Robust Design

Taguchi methods and split-plot designs address robustness objectives in product and process design where minimizing sensitivity to uncontrollable noise factors is paramount [11]. Automotive component testing frequently uses robust parameter designs with control factors (design parameters) and noise factors (environmental conditions, material variability) to identify configurations that deliver consistent performance across operating conditions Citation 11. Modern implementations combine Taguchi's philosophical emphasis on robustness with response surface methodology's statistical rigor, employing dual-response optimization to simultaneously target mean performance and minimize variance [18].

Split-plot designs naturally accommodate the hierarchical structure of industrial experiments where certain factors are difficult or expensive to change [2], Citation attached_file:1. A batch chemical process might treat reactor temperature as a hard-to-change whole-plot factor (requiring hours to adjust) while catalyst concentration and reaction time serve as easy-to-change subplot factors, resulting in a design that respects operational constraints while maintaining statistical validity through proper mixed-effects model analysis.

D. Computer Experiments and Simulations Based Design

Space-filling designs have become standard practice in engineering fields relying on expensive computational simulations [6], [10], [20]. Aerospace applications employ Latin Hypercube Sampling and Sobol sequences to explore design spaces for computational fluid dynamics simulations of airfoil performance, where each simulation requires hours of high-performance computing [7],[14]. The uniform coverage properties of these designs ensure efficient sampling for subsequent surrogate model construction using Gaussian processes or neural networks [10],[16],[20].

Automotive crash safety analysis illustrates adaptive sequential design integration with simulation, where Bayesian optimization guides finite element model runs to rapidly converge on designs meeting regulatory requirements while minimizing structural weight. This active learning approach reduces the number of expensive crash simulations by intelligently selecting subsequent design points based on accumulated information [16],[17].

Emerging Applications in Advanced Manufacturing

Industry 4.0 and smart manufacturing initiatives are driving integration of DOE principles with real-time data analytics and closed-loop control [16],[20].Additive manufacturing parameter optimization employs hybrid designs combining space-filling exploration with response surface refinement to map relationships between process parameters (laser power, scan speed, layer thickness) and part properties (density, surface finish, mechanical strength). The multi-stage nature of these applications benefits from sequential experimentation where initial screening identifies critical parameters before detailed optimization in reduced factor spaces.

These industrial case studies demonstrate that successful DOE implementation requires matching design methodology to experimental objectives, resource constraints, and operational realities while maintaining statistical rigor in analysis and interpretation

VI. Discussions and Further Directions

The evolution of experimental design methodologies from classical factorial principles to modern adaptive and machine learning-integrated approaches reflects both theoretical advances in statistical science and practical demands of increasingly complex industrial. This review has documented a rich methodological landscape spanning deterministic geometric designs, algorithmic optimal designs, model-free space-filling designs, and emerging intelligent experimentation frameworks. Looking forward, several key trends and research frontiers are reshaping the practice and theory of design of experiments.

Integration with Machine Learning and Artificial Intelligence

The convergence of experimental design with machine learning represents perhaps the most significant contemporary development in the field. Traditional DOE assumes parametric response models (polynomials), whereas modern industrial systems increasingly require flexible nonparametric models capable of capturing complex, high-dimensional relationships. Gaussian process emulators, random forests, gradient boosting machines, and neural networks offer unprecedented modeling flexibility, but their performance depends critically on training data coverage and distribution.

Space-filling designs naturally complement these machine learning approaches by ensuring adequate sampling across input spaces. However, optimal experimental strategies for training deep neural networks or ensemble methods remain an active research area with limited theoretical guidance compared to the well-established theory for linear and polynomial models. Developing design criteria and methodologies specifically optimized for machine learning objectives represents a major frontier for both the statistics and engineering communities.

Active learning and Bayesian optimization have emerged as powerful frameworks for sequential experimentation, particularly when individual experiments are expensive or time-consuming. These approaches use acquisition functions (expected improvement, upper confidence bound, probability of improvement) to intelligently select subsequent experiments based on accumulated information from previous runs While computationally intensive, these methods can dramatically reduce total experimental effort by focusing resources on the most informative regions of the factor space.

Autonomous Experimentation and Closed-Loop Systems

Industry 4.0 and smart manufacturing initiatives are driving development of autonomous experimentation systems that integrate experimental design principles with real-time data acquisition, closed-loop control, and reinforcement learning. These systems can conduct experiments without human intervention, continuously learning and adapting based on streaming sensor data and automatically adjusting process parameters to optimize multiple objectives simultaneously.

Materials discovery and drug development are witnessing deployment of robotic laboratories where automated systems perform thousands of experiments guided by adaptive design algorithms. These self-driving laboratories promise to accelerate innovation cycles by orders of magnitude while simultaneously exploring vastly larger design spaces than traditional human-in-the-loop experimentation can achieve. However, ensuring robustness, interpretability, and scientific rigor in these autonomous systems remains a significant challenge requiring continued methodological development.

Multi-Objective and Constrained Optimization

Industrial applications increasingly require simultaneous optimization of multiple, often competing objectives subject to complex constraints. Traditional response surface methodology focuses on optimizing a single response, with multi-response problems typically handled through weighted utility functions or desirability indices that reduce the problem to single-objective optimization. Modern approaches employing Pareto optimality and evolutionary algorithms offer more sophisticated handling of trade-offs, but integration with experimental design theory remains incomplete.

Constrained optimization where feasible regions are defined by complex nonlinear constraints, safety requirements, or regulatory limits requires design methodologies that respect these restrictions while maintaining statistical efficiency. D-optimal and other algorithmic designs provide flexibility, but theoretical guarantees and practical performance under constraint violations need further investigation.

High-Dimensional and Massive-Scale Experimentation

As industrial systems grow in complexity, experiments involving tens or hundreds of factors are becoming more common, particularly in genomics, materials science, and complex manufacturing processes. Traditional factorial and response surface designs scale poorly to high dimensions due to the curse of dimensionality—the exponential growth in required runs.

Supersaturated designs, which estimate fewer effects than the number of factors by exploiting sparsity assumptions even more aggressively than screening designs, represent one response to this challenge. Space-filling designs combined with variable selection algorithms and regularized regression methods (LASSO, elastic net) offer alternative pathways for high-dimensional exploration. However, theoretical understanding of when and why these approaches succeed or fail remains limited, particularly regarding the required sparsity levels and sample sizes for reliable inference.

Robustness and Uncertainty Quantification

Modern industrial practice increasingly emphasizes robustness—designing products and processes that perform consistently despite variation in uncontrolled factors. While Taguchi methods popularized robust design concepts, integration of robustness objectives with optimal design theory and response surface methodology remains an active area. Dual-response optimization (targeting both mean and variance) and robust optimal designs that account for uncertainty in model parameters represent promising directions.

Uncertainty quantification in computer experiments and simulation-based design is gaining prominence as computational models become more complex and their parameter uncertainty more significant. Designing experiments to efficiently characterize uncertainty propagation through complex systems, particularly under expensive simulation constraints, requires integration of space-filling designs with Bayesian inference and sensitivity analysis techniques.

Reproducibility and Open Science

The broader scientific reproducibility crisis has implications for experimental design practice. Proper randomization, blinding where feasible, pre-registration of analysis plans, and transparent reporting of design choices and analysis decisions are increasingly recognized as essential for credible inference. Development of standardized reporting guidelines for designed experiments, similar to CONSORT for clinical trials, could enhance reproducibility and meta-analysis across industrial applications.

Open-source software ecosystems for design generation, data analysis, and visualization have democratized access to sophisticated DOE methodologies, but also risk misapplication by practitioners without adequate statistical training Citation attached_file:1. Educational initiatives and expert systems that guide design selection based on experimental context represent important complementary developments to methodological research.

CONCLUSION

This comprehensive review has synthesized the diverse landscape of experimental design methodologies employed in industrial applications, from classical factorial approaches to modern adaptive and machine learning-integrated frameworks. The systematic comparison presented in Tables 1 and 2, encompassing nineteen distinct design methodologies, provides practitioners with a structured reference for navigating design selection decisions based on experimental objectives, resource constraints, and operational realities.

Several key insights emerge from this analysis. First, no single design methodology dominates across all experimental contexts. The optimal choice depends critically on the interplay between run efficiency requirements, model assumptions, factor dimensionality, space-filling properties, and practical constraints. Full factorial designs offer complete information but scale poorly beyond four factors. Fractional factorials achieve dramatic efficiency gains through strategic confounding. Response surface designs enable quadratic modeling for optimization. Space-filling approaches prioritize uniform coverage for flexible modeling. Optimal designs maximize statistical criteria while accommodating complex constraints.

Second, bias mitigation remains central to valid experimental inference across all design types. Randomization distributes effects of uncontrolled variables, blocking addresses known heterogeneity, replication enables error estimation, and careful attention to confounding patterns prevents misattribution of effects. The mathematical foundations establish that orthogonality, information matrix properties, and discrepancy measures provide rigorous frameworks for evaluating design quality.

Third, modern methodological advances are expanding experimental capabilities in response to emerging industrial challenges. Definitive Screening Designs unify screening and optimization in remarkably efficient structures. Adaptive and sequential approaches leverage Bayesian updating to intelligently allocate experimental effort. Integration with machine learning enables flexible nonparametric modeling, while autonomous experimentation systems promise to revolutionize innovation cycles through closed-loop learning.

Looking forward, several research frontiers merit continued attention. High-dimensional experimentation challenges traditional designs as factor counts grow. Multi-objective optimization under complex constraints requires integrated frameworks. Robustness and uncertainty quantification demand designs that characterize not just mean behavior but variability. Reproducibility necessitates standardized reporting practices and transparent documentation.

The fundamental principles articulated by Fisher, Box, and other pioneers—randomization, replication, factorial thinking, and systematic variation—remain as relevant today as when first formulated. What has evolved is the methodological toolkit available for implementing these principles across increasingly diverse experimental contexts. This review provides both a comprehensive reference for current practice and a foundation for advancing experimental methodology, equipping researchers and practitioners to make informed design choices while recognizing opportunities where new approaches are needed.

As industrial systems grow in complexity and experimental costs increase, the strategic application of sound experimental design principles becomes essential for efficient innovation and competitive advantage**.** Continued integration across statistics, machine learning, optimization, and domain sciences promises further advances in our ability to extract maximum information from systematic experimentation, accelerating scientific discovery and industrial innovation in the decades ahead

REFERENCES

[1] R.A. Fisher, "The Design of Experiments," 1st ed., Oliver & Boyd, Edinburgh, 1935.
[2] NIST/SEMATECH e-Handbook of Statistical Methods, "Design of Experiments," National Institute of Standards and Technology, Gaithersburg, MD, 2023.
[3] R.L. Plackett and J.P. Burman, "The Design of Optimum Multifactorial Experiments," Biometrika, vol. 33, no. 4, pp. 305–325, 1946.
[4] G.E.P. Box and K.B. Wilson, "On the Experimental Attainment of Optimum Conditions," Journal of the Royal Statistical Society Series B, vol. 13, no. 1, pp. 1–45, 1951.
[5] G.E.P. Box and D.W. Behnken, "Some New Three Level Designs for the Study of Quantitative Variables," Technometrics, vol. 2, no. 4, pp. 455–475, 1960.
[6] M.E. McKay, R.J. Beckman, and W.J. Conover, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, vol. 21, no. 2, pp. 239–245, 1979.
[7] S. Owen, "Variance Components and Quasi-Monte Carlo," in Monte Carlo and Quasi-Monte Carlo Methods, Springer, Berlin, 2023.
[8] Tirthajyoti Sarkar, "Design of Experiments with Python," GitHub, 2022.
[9] "Design of Experiments (DOE)," pyDOE documentation, 2022. Available:
[10] J. Liu, "Space-Filling Designs in Complex Industrial Systems," Journal of Quality Technology, vol. 51, no. 1, pp. 52–68, 2019.
[11] S.M. Taguchi, "Introduction to Quality Engineering," Asian Productivity Organization, Tokyo, 1986.
[12] L. Vandenberghe et al., "Optimal Design for Industrial Experiments under Model Uncertainty," Journal of Statistical Planning and Inference, vol. 215, pp. 192–214, 2021.
[13] Altair Engineering, "Central Composite and Response Surface Designs," Altair Help, 2022.
[14] "Computer-Aided Experimental Design: Halton and Sukharev Grids," pyDOE documentation, 2022.
[15] "On the Use of Blocking in Factorial and Response Surface Designs," National Institute of Statistical Sciences, Technical Report, 2020.
[16] S. Biswas and D. Rakshit, "Machine Learning Integrated DOE for Robust Industrial Optimization," International Journal of Production Research, vol. 59, no. 18, pp. 5421–5438, 2021.
[17] K. Chaloner and I. Verdinelli, "Bayesian Experimental Design: A Review," Statistical Science, vol. 10, no. 3, pp. 273–304, 1995.
[18] D.C. Montgomery, "Design and Analysis of Experiments," 9th ed., Wiley, Hoboken, NJ, 2020.
[19] H. Xu and C.F.J. Wu, "Generalized minimum aberration for factorial designs with quantitative and qualitative factors," Annals of Statistics, vol. 29, no. 3, pp. 1072–1106, 2001.
[20] B. J. Williams et al., "Latest Advances in Space-Filling, Clustered, and Adaptive Experimental Design for Industry," Applied Statistics, vol. 72, no. 1, pp. 49-73, 2022.