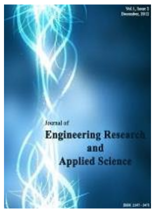
Journal of Engineering Research and Applied Science
Available at www.journaleras.com
Volume 15 (1), June 2026, pp 52-67
ISSN 2147-3471 © 2026



Statistical Methods in Optimizing Industrial Systems

Adinife Patrick Azodo[1,*], Oluwadamilola Oludare Ojebode[2]

[1]Mechanical Engineering, Federal University Wukari, Nigeria
[2]Mathematics and Statistics, Federal University of Agriculture Abeokuta, Nigeria

Accepted 25 May 2026

Abstract
Optimizing industrial energy systems is vital for achieving sustainability goals, reducing costs, and improving operational performance. This paper examines the integration of advanced statistical methods, such as regression analysis, time-series forecasting, Monte Carlo simulations, and machine learning, into industrial energy management. These approaches enable predictive maintenance, accurate energy demand forecasting, and efficient resource allocation, while machine learning models in particular enhance reliability by predicting failures, minimizing downtime, and extending asset life. Statistical techniques also strengthen risk assessment and process control under uncertainty, supporting proactive energy management. Despite these advantages, challenges such as data quality, model scalability, and system complexity remain. Addressing these requires advances in data integration, algorithm interpretability, and interdisciplinary collaboration. By synthesizing statistical approaches with emerging technologies like artificial intelligence and the Internet of Things, this paper underscores the transformative role of statistically informed optimization in building resilient, cost-effective, and sustainable industrial energy systems.

Keywords: *Industrial Optimization; Statistical Process Control (SPC); Design of Experiments (DOE), Regression Analysis, Lean and Six Sigma; Predictive Maintenance*

1. Introduction
Industrial system optimization plays a pivotal role in enhancing performance, reducing operational costs, and advancing sustainability across manufacturing and production sectors. As industries face growing pressures to improve productivity, minimize waste, and adapt to volatile market demands, the adoption of enabling technologies such as Artificial Intelligence (AI), the Internet of Things (IoT), and cloud computing has become indispensable. These technologies support real-time data acquisition, predictive maintenance, and intelligent decision-making, contributing to more agile and responsive operations [1 – 2]. Complementary methodologies such as Six Sigma and Lean manufacturing continue to provide structured approaches to quality improvement and process streamlining [3].

Over the past few decades, industrial optimization has shifted from trial-and-error adjustments to data-driven, system-wide strategies. Early efforts primarily targeted discrete components, such as individual machines or production stages. In contrast, modern optimization emphasizes holistic systems thinking, where improvements derive from the coordinated integration of physical infrastructure, human resources, decision logic, and enabling technologies [4 – 5]. Current optimization goals are multidimensional, encompassing efficiency, cost-effectiveness, quality assurance, environmental stewardship, and resilience.

For instance, AI-based control systems and Lean principles enhance adaptive efficiency [3, 6], while energy-conscious design and circular economy strategies promote sustainability [7, 8]. Structured approaches such as goal programming and Six Sigma strengthen product quality and process reliability [9], and decarbonization targets increasingly shape operational decisions [10]. However, the accelerating complexity of modern industrial systems, marked by high interconnectivity, nonlinear feedback loops, and real-time variability, poses challenges that traditional deterministic methods are ill-equipped to address [11, 12]. Addressing these challenges requires analytical techniques capable of handling uncertainty, uncovering hidden patterns, and adapting continuously to dynamic environments.

Within this evolving landscape, statistical methods have advanced from auxiliary tools to essential

*Corresponding author: azodopat@gmail.com

enablers of robust industrial optimization. Techniques such as regression modeling, cluster analysis, Design of Experiments (DOE), multivariate analysis, and simulation-based approaches provide the quantitative foundation for understanding system behavior and guiding decisions. This review will not only describe these methods but also critically evaluate their mathematical principles, comparative performance, and limitations, including issues such as overfitting risks and data quality constraints in practical applications. At the same time, the emergence of machine learning, big data analytics, quantum computing, and swarm intelligence offers opportunities to augment statistical approaches, enabling more accurate, scalable, and interpretable models for industrial performance improvement [13 – 14]. Despite this potential, integration remains fragmented, particularly in developing contexts, due to barriers such as poor data quality, infrastructural limitations, and skills shortages.

This review was conducted through a systematic search of major academic databases, including Scopus, Web of Science, and Google Scholar. Keywords such as "statistical optimization," "Design of Experiments," and "industrial systems" were used to identify relevant peer-reviewed studies, ensuring the inclusion of both foundational and contemporary research.

This paper addresses a critical gap in the literature: the absence of a comprehensive synthesis linking traditional statistical optimization methods with emerging digital technologies to enable system-wide, adaptive industrial optimization. Specifically, it aims to:

- Examine the individual and combined effectiveness of statistical and digital methodologies in industrial optimization.
- Analyze the barriers hindering integration, including infrastructural limitations, data quality issues, and scalability concerns in resource-constrained industrial settings.
- Propose a strategic roadmap for flexible, efficient, and sustainable optimization practices that meet the needs of both modern and emerging industrial sectors.

2. Core Statistical Methodologies for Industrial System Optimization

The strategic application of statistical techniques is essential for evidence-based decision-making, performance enhancement, and continuous improvement in industrial systems. This section reviews three core methodologies widely adopted across industries: Statistical Process Control (SPC), Regression Analysis, and Design of Experiments (DOE). We examine their underlying principles and key mathematical models, which together provide a rigorous foundation for monitoring processes, identifying patterns, and optimizing operations [15, 16].

SPC is a cornerstone of quality control, using control charts and process capability analysis to detect and prevent deviations in manufacturing and service processes [17]. Regression analysis, both linear and nonlinear, enables the modeling of relationships between variables, facilitating predictive analytics and root-cause analysis in industrial settings [18]. DOE provides a structured framework for experimentation, allowing practitioners to systematically explore factor effects and interactions to optimize process parameters [19].

In addition to these core methods, complementary approaches such as Monte Carlo simulations and time-series forecasting are particularly relevant for optimizing dynamic and energy-intensive environments. Monte Carlo methods simulate probabilistic outcomes by sampling from input distributions, offering insights into risk and uncertainty in system performance [20]. Time-series forecasting, including ARIMA and machine learning-based models, supports predictive maintenance, demand forecasting, and energy load optimization [21].

These techniques form a comprehensive framework for improving industrial performance through data-driven insights and structured experimentation. Despite their effectiveness, applying these methodologies in practice presents challenges. Ensuring data integrity, addressing resource limitations, and navigating the complexities of statistical design and interpretation remain critical barriers that must be overcome for successful deployment [22].

2.1 Statistical Process Control (SPC)

Statistical Process Control (SPC) is a cornerstone of industrial quality management, designed to monitor, regulate, and improve process stability by distinguishing between common-cause and special-cause variation. By applying statistical methods to production data, SPC ensures that processes remain consistent and capable of meeting specification requirements [23].

Core SPC tools include X-bar and R charts, which monitor process averages and ranges, and P and C charts, which are applied to attribute data. The mathematical foundation of SPC lies in the calculation of control limits, typically set at three standard deviations from the process mean. For example, the upper and lower control limits (UCL and LCL) for an X-bar chart are given by:

$$UCL = \overline{X} + A_2 \cdot \overline{R}, LCL = \overline{X} - A_2 \cdot \overline{R}$$

where $\overline{X}$ is the process mean, $\overline{R}$ is the average range, and A_2 is the average range, and A_2 is a constant based on sample size [24]. These formulas enable SPC to detect statistically significant deviations while minimizing false alarms.

SPC has been widely adopted in industries such as automotive manufacturing, where it ensures adherence to tight dimensional tolerances (e.g., piston diameters, bolt torque), and in food packaging, where it maintains fill-level consistency and sealing integrity [25]. Its integration with automated data collection and real-time analytics further enhances responsiveness and defect prevention [25].

However, SPC is not without limitations. Traditional Shewhart charts assume data normality and independence, which may not hold in modern, highly dynamic production environments. When applied to skewed or autocorrelated data, SPC can produce misleading signals [26]. Short production runs, common in customized manufacturing, also limit the amount of data available for reliable charting. Additionally, overreliance on SPC may obscure underlying systemic issues if practitioners focus solely on control limits without addressing root causes.

Despite these challenges, SPC remains an essential methodology, particularly when complemented with advanced techniques such as multivariate control charts and machine learning–driven anomaly detection, which extend its applicability to complex, high-variability industrial systems [27].

2.2 Regression Analysis and Predictive Modeling

Regression analysis is a fundamental statistical tool for identifying and quantifying relationships between process variables, enabling industries to make predictions, optimize operations, and prevent failures. At its core, the general form of a multiple linear regression model is expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

where Y is the dependent variable (e.g., energy consumption or production output), X_i are the independent variables (e.g., temperature, pressure, production rates), β_i are the regression coefficients representing the effect of each predictor, and ϵ\epsilonϵ is the error term accounting for unexplained variation.

In industrial applications, linear regression is commonly used for simple relationships, while multiple and logistic regression are applied to more complex problems, such as predicting equipment failure probability, modeling production quality deviations, or forecasting supply chain delays [28]. For example, manufacturers use regression to estimate energy consumption based on production and environmental conditions, while logistics firms apply it to predict cost fluctuations and delivery times.

Despite its wide applicability, regression modeling faces critical challenges. Multicollinearity arises when independent variables are highly correlated, making coefficient estimates unstable and difficult to interpret. Overfitting occurs when a model captures noise in the training data, reducing its ability to generalize to unseen data. Outliers can further distort regression coefficients, leading to misleading predictions. These limitations highlight the trade-off between model complexity and reliability [30].

To mitigate such issues, industries adopt advanced techniques. k-fold cross-validation is widely used to assess generalizability by partitioning data into training and testing subsets. Regularization methods such as Ridge regression (L_2 penalty) and Lasso regression (L_1 penalty) constrain coefficient estimates to reduce overfitting, while robust regression frameworks enhance resistance to outliers. Together, these methods strengthen the predictive power of regression models, ensuring a balance between accuracy and interpretability.

2.3 Design of Experiments (DOE)

Design of Experiments (DOE) is a structured statistical methodology used to investigate how multiple input variables influence system outcomes, making it especially valuable for process optimization and quality improvement. Mathematically, DOE is based on the principle that the response variable (Y) can be expressed as a linear combination of factors (X_i) and their interactions:

$$Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ij} X_i X_j + \epsilon$$

where β_0 is the intercept, β_i are the main effects of individual factors, β_{ij} represent interaction effects, and ϵ is the random error term [30].

Full factorial designs allow for the estimation of all main effects and interactions but require a large number of runs as factors increase. In contrast, fractional factorial designs strategically reduce the number of experimental runs by sacrificing the estimation of higher-order interactions, which are often negligible in industrial practice [31]. For fine-tuning and optimization, Response Surface Methodology (RSM) is employed to model nonlinear relationships. However, RSM assumes a smooth and continuous response surface, an assumption that may not always hold in real-world processes, particularly when abrupt changes or nonlinearity dominate system behavior [32].

Applications of DOE span from optimizing weld parameters in automotive fabrication to maximizing yield in chemical processing. Its benefits include reducing the number of experiments required, minimizing resource consumption, and extracting maximum information about factor effects and interactions. However, successful application requires statistical expertise and careful planning. Missing a key factor during design can bias results and limit insights, while resource constraints (time, cost, and equipment availability) may restrict the use of more complex designs. DOE facilitates evidence-based decision-making by enabling efficient experimentation that balances resource use with the need for accurate and actionable insights.

2.4 Time-Series Forecasting and Monte Carlo Simulations

Time-Series Forecasting

Time-series forecasting is a statistical approach that uses historical data to predict future values of a variable over time. One widely used method is the Autoregressive Integrated Moving Average (ARIMA) model, expressed as ARIMA(p, d, q), where:

p: the order of the autoregressive (AR) part, which captures the influence of past values on the current value;

d: the degree of differencing, which ensures stationarity of the series;

q: the order of the moving average (MA) part, which accounts for dependencies on past forecast errors.

Mathematically, a simplified ARIMA model can be written as:

$$Y_i = \alpha + \sum_{i=1}^{p} \emptyset_i Y_{t-i} + \sum_{j=1}^{q} \theta_i \epsilon_{t-j} + \epsilon_t$$

where Yt is the predicted value at time t, ϕi are AR coefficients, θj are MA coefficients, and ϵt is the error term.

Industrial applications include forecasting energy load demand, maintenance scheduling, and renewable energy output (e.g., predicting solar or wind generation for grid balancing) [33, 34]. However, ARIMA and similar models assume linearity and may perform poorly with non-stationary data or sudden, unforeseen disruptions. Model selection (choosing p, d, q) can be complex, and overfitting is a persistent challenge, especially in volatile industrial environments [35].

Monte Carlo Simulations

Monte Carlo simulation is a probabilistic method that evaluates uncertainty in complex systems by performing repeated random sampling. The general algorithm involves:

- Defining the system variables of interest,
- Assigning probability distributions to uncertain inputs,
- Running thousands (or more) of random iterations,
- Aggregating the outcomes to obtain a probability distribution of possible results.

Mathematically, if f(x) represents the outcome of a system with random input x, then the Monte Carlo estimate of the expected value is:

$$\hat{E}[f(x)] = \frac{1}{N}\sum_{i=1}^{N} F(x_i)$$

where N is the number of simulations and xi are random samples from the defined input distributions.

In industrial settings, Monte Carlo is widely applied to failure prediction under variable operating conditions, maintenance optimization, and investment risk assessment [36, 37]. For example, in supply chains, it helps evaluate the probability of cost overruns or delays under uncertain demand scenarios. However, the accuracy of results strongly depends on the quality of input data and assumptions about probability distributions. Additionally, Monte Carlo can be computationally intensive, making real-time implementation challenging [38].

Time-series forecasting and Monte Carlo simulations provide powerful complementary tools: the former captures temporal patterns for prediction, while the latter quantifies uncertainty in decision-making. Their growing adoption in energy systems planning, maintenance strategies, and supply chain management highlights their practical importance but also underscores the need for careful model selection, validation, and computational considerations.

3. Theoretical Foundations of Statistical Optimization in Industrial Systems

The effective application of statistical methods in industrial optimization is underpinned by rigorous mathematical and statistical theory. These foundational concepts provide the structure necessary for modeling, analyzing, and improving complex industrial processes under uncertainty. This section highlights the key theoretical constructs convex analysis, duality, probabilistic modeling, and inferential statistics and explains their relevance to the statistical techniques employed in industrial systems.

3.1 Principles of Convex Analysis and Duality Theory

Convex analysis forms the theoretical backbone of many optimization methods because it ensures that local optima are also global optima in convex problem spaces. Formally, a set C ⊆ Rn is convex if, for any two points x_1, x_1, ∈ C, the line segment connecting them is also in C:

$$\lambda \mathrm{x}_1 + (1-\lambda)\mathrm{x}_2 \in \mathrm{C} \qquad \forall x_{1,} x_2 \in \mathrm{C}, 0 \leq \boldsymbol{\lambda} \leq 1$$

Similarly, a function f: Rn → R is convex if for all x_1, x_1, ∈ dom(f) and $0 \leq \boldsymbol{\lambda} \leq 1$

$$f(\lambda \mathrm{x}_1 + (1-\lambda)\mathrm{x}_2) \leq \lambda f(x_1) + (1-\lambda) f(\mathrm{x}_2)$$

This property guarantees that convex optimization problems are computationally tractable and have well-defined global solutions, making them central to industrial techniques such as regression modeling and experimental design.

Duality theory complements convex analysis by associating each optimization problem (the primal) with a corresponding dual problem. The dual solution provides a lower bound to the primal solution, and under strong duality conditions, the two solutions coincide. Duality not only facilitates the derivation of efficient algorithms but also provides insights into sensitivity analysis, particularly how changes in constraints affect the optimal solution [39].

However, these principles are not universally applicable. Many real-world industrial optimization problems are non-convex, involving discrete variables, nonlinearities, or multiple local minima. In such cases, convexity-based guarantees break down, and solutions may require heuristic, approximate, or global optimization methods. Recognizing these limitations is essential for applying convex analysis and duality appropriately in industrial systems.

3.2. Probabilistic Modeling and Stochastic Processes

Industrial systems are inherently uncertain, with variability in processes, demand, and environmental conditions. Probabilistic modeling provides a framework to quantify this uncertainty by representing inputs and outputs through random variables. Formally, a random variable is a function that maps the outcomes of a random experiment to numerical values, capturing the inherent randomness of industrial processes.

A stochastic process extends this concept by considering a collection of random variables indexed by time (or another parameter), i.e., $\{X_t: t \in T\}$. This models the random evolution of a system over time, making it a fundamental tool for analyzing production variability, equipment reliability, and demand fluctuations. For example, Monte Carlo simulations rely on repeated sampling from probability distributions defined by random variables to evaluate uncertainty in system performance [40]. Similarly, time-series forecasting methods are rooted in stochastic process theory, as they model temporal interdependence and randomness in industrial data such as energy demand or equipment degradation [41].

Despite their versatility, these theoretical tools have limitations. Model assumptions play a critical role: probabilistic models are only as reliable as the assumed probability distributions of inputs. Incorrect assumptions can lead to misleading forecasts or risk assessments. Stochastic process modeling also presents challenges, as selecting an appropriate framework (e.g., Markov chains for discrete states, Wiener processes for continuous dynamics) can be difficult, and more sophisticated models often demand substantial computational resources. Furthermore, the accuracy of probabilistic models heavily depends on the quality and quantity of available data, poor or insufficient data can compromise the reliability of estimated distributions and derived predictions.

Overall, probabilistic modeling and stochastic processes provide powerful theoretical foundations for industrial decision-making under uncertainty. However, their effective application requires careful validation of assumptions, judicious model selection, and robust data support to balance analytical rigor with practical feasibility.

3.3. Inferential Statistics and Hypothesis Testing

Inferential statistics form the backbone of statistical optimization by allowing conclusions to be drawn about entire processes based on sample data. At the core of this framework is hypothesis testing, a formal procedure used to evaluate assumptions about system parameters. In this approach, the null hypothesis (H_0) represents the assumption of no effect or no difference, while the alternative hypothesis (H_0) reflects the presence of a statistically significant effect. The strength of evidence is commonly assessed through the p-value, which quantifies the probability of observing results as extreme as the data obtained, assuming H_0 is true. A low p-value (typically < 0.05) provides grounds for rejecting H_0, thereby lending confidence to experimental or model-based conclusions.

In the context of Design of Experiments (DOE), hypothesis testing is used to determine whether changes in factor levels significantly influence system performance. This ensures that identified factor effects are not due to random variation but represent meaningful drivers of system outcomes. Similarly, in regression analysis, inferential statistics support the evaluation of parameter estimates, assessment of model fit, and quantification of uncertainty, enabling practitioners to separate genuine patterns from spurious correlations [30].

However, these methods are not without limitations. Hypothesis tests rely on key assumptions, such as independence of observations, normally distributed residuals, and homogeneity of variance. Violations of these assumptions can undermine the validity of statistical inferences, leading to biased or misleading conclusions. Additionally, the reliance on p-values has been widely criticized due to their potential for misinterpretation, most notably the tendency to conflate statistical significance with practical significance. This highlights the need for complementary measures such as effect sizes, confidence intervals, and practical relevance assessments to ensure robust decision-making in industrial optimization.

3.4. Optimization Under Constraints: Lagrange Multipliers and Linear Programming

Optimization under constraints plays a central role in industrial systems where decisions must satisfy operational, regulatory, or resource limitations. Two foundational approaches are Lagrange multipliers and linear programming (LP), both of which integrate statistical models with optimization frameworks.

The method of Lagrange multipliers is designed to find the maximum or minimum of an objective function subject to equality constraints. The central principle is that, at the optimum, the gradient of the objective function is proportional to the gradient of the constraint. Mathematically, for an objective function f(x) subject to a constraint g(x) = 0, the condition is expressed as:

$$\nabla f(x) = \lambda \nabla g(x)$$

where λ is the Lagrange multiplier. In statistical optimization, this method is particularly useful for parameter estimation problems with bounded feasibility, such as maximizing likelihood functions under physical or regulatory constraints. However, its applicability is limited to differentiable functions and equality constraints, making it less practical for complex, non-smooth, or inequality-constrained industrial problems.

In contrast, linear programming focuses on optimizing a linear objective function subject to linear equality and inequality constraints. A standard LP formulation is:

$$\text{Maximize (or Minimize) } cT\,x$$
$$\text{subject to } Ax \leq b,\ x \geq 0,$$

where c is the vector of objective coefficients, A represents constraint coefficients, and b defines the constraint bounds. LP is widely used in industrial planning, scheduling, and resource allocation, particularly when decision variables such as production volumes, material flows, or costs are derived from statistical forecasts or predictive models. Its primary limitation lies in the assumption of linearity, which often oversimplifies real-world industrial systems characterized by nonlinear dynamics, stochastic variability, and uncertainty.

Critically, both methods demonstrate the tension between mathematical tractability and real-world complexity. While they provide elegant frameworks for constrained optimization, their assumptions can limit applicability. Recent research has focused on extending these classical tools into nonlinear programming, stochastic programming, and robust optimization, where uncertainty and complexity are explicitly modeled [42]. Thus, the integration of statistical analysis with optimization theory ensures that constrained decision-making in industrial systems remains both rigorous and practically relevant.

3.5. Advanced Statistical Approaches for Robustness and Risk

The pursuit of robustness in industrial optimization emphasizes not only achieving optimal performance but also ensuring stability under variability and uncertainty. Robust statistical approaches are designed to identify parameter settings that deliver consistent results despite uncontrollable "noise" factors such as environmental variation, material inconsistencies, or equipment degradation. These methods extend beyond traditional regression, DOE, and SPC by explicitly accounting for unpredictability in complex, data-rich industrial systems [43 – 45].

A prominent technique for robustness is the Taguchi Method, which uses orthogonal arrays to systematically investigate multiple factors while minimizing the number of experiments. A central metric in Taguchi's framework is the Signal-to-Noise (S/N) ratio, which quantifies robustness by maximizing desirable signal effects while minimizing the influence of noise. This approach is particularly effective for identifying factor levels that maintain stable performance despite external disturbances [46]. Similarly, Response Surface Methodology (RSM) builds predictive models of system behavior by fitting a response surface to experimental data. RSM employs a sequence of designed experiments to locate optimal operating conditions, making it especially useful for fine-tuning multivariate industrial processes [19]. However, both methods face limitations: they are most efficient with a relatively small number of factors and rely on assumptions about the shape of the response surface, which may not hold in highly nonlinear or complex systems [44 – 45].

Another widely adopted tool is the Monte Carlo simulation, which evaluates probabilistic outcomes by repeatedly sampling from input distributions defined by random variables. By simulating thousands of possible scenarios, Monte Carlo provides insight into the range and likelihood of system performance under uncertainty. In industrial contexts, it supports risk-informed decision-making in areas such as system design, production scheduling, and logistics planning [20]. While powerful, Monte Carlo methods are computationally intensive and highly dependent on the quality of assumed probability distributions; poorly specified inputs can lead to misleading results [47, 48].

The integration of advanced statistical techniques with algorithmic optimization methods further enhances robustness and responsiveness. For example, statistical forecasting models can be embedded within routing algorithms or scheduling heuristics to improve adaptability in real-time industrial operations [21]. Yet, the effectiveness of these hybrid approaches is constrained by practical challenges. High-quality, representative data are essential for reliable model building, and inadequate infrastructure or insufficient statistical expertise can hinder implementation. Additionally, as methods become more computationally demanding, organizations must invest in digital infrastructure and workforce training to fully leverage their potential [49 – 51].

Advanced statistical approaches such as the Taguchi Method, RSM, and Monte Carlo simulations offer powerful frameworks for enhancing robustness and managing risk in industrial optimization. However, their benefits are balanced by assumptions, complexity, and data requirements. Addressing these challenges is critical for achieving sustainable, high-performance industrial systems that balance precision with resilience.

3.5. Advanced Statistical Approaches for Robustness and Risk

Monte Carlo simulation supports risk-informed decisions by repeatedly sampling from specified input distributions to produce probabilistic outcome estimates for design, scheduling, and logistics under uncertainty; its insights depend critically on well-specified distributions and can be computationally demanding [47, 48].

Blending these statistical tools with algorithmic optimization (e.g., embedding forecasts within routing/scheduling heuristics) can improve real-time responsiveness, but success hinges on high-quality, representative data, skilled personnel, and adequate digital infrastructure to handle modeling and computation. These implementation constraints remain a practical barrier in many organizations [49 – 50].

4. Integrating Statistical Methods with Industrial Engineering Practices

The optimization of industrial systems is increasingly driven by the strategic integration of statistical methodologies with core industrial engineering (IE) practices. Statistical tools are the analytical cornerstone of operational excellence, quality assurance, and process improvement frameworks rather than functioning alone. This integration has improved the ability of engineering experts to make data-driven decisions, identify the root causes of inefficiencies, and secure the long-term viability of industrial systems.

4.1 Statistical Foundations of Lean and Six Sigma

Lean and Six Sigma, two widely recognized methodologies in industrial engineering and operations management, are grounded in statistical thinking. While Lean emphasizes the elimination of waste and the pursuit of value from the customer's perspective, Six Sigma formalizes problem-solving through the DMAIC cycle (Define, Measure, Analyze, Improve, Control). The integration of statistical tools within this cycle provides a rigorous foundation for process improvement.

Statistical Process Control (SPC): In the Measure and Control phases, SPC enables practitioners to establish baseline performance, monitor variability, and detect deviations in real time, ensuring that improvements are sustained [52].

Design of Experiments (DOE): Applied in the Analyze and Improve phases, DOE systematically tests factor interactions, identifies root causes of variation, and determines optimal operating conditions, thereby moving beyond trial-and-error approaches.

Regression Analysis: Regression models quantify relationships between process inputs and outputs, supporting predictive analysis and guiding decision-making in both Analyze and Improve phases [53].

Although Lean is often criticized for its reliance on qualitative observations (e.g., waste identification or value stream mapping), the incorporation of statistical methods strengthens its rigor. By quantifying waste levels, cycle times, and process variation, statistical tools transform subjective assessments into objective, data-driven insights. This integration bridges the gap between Lean's qualitative principles and Six Sigma's quantitative rigor, enabling a balanced and evidence-based approach to continuous improvement.

4.2 Strengthening Supply Chain and Logistics through Statistical Insight

Statistical models underpin core decision-making processes in supply chain and logistics, particularly in demand forecasting, inventory control, and routing optimization. Time-series methods such as ARIMA and exponential smoothing capture historical patterns and seasonality, enabling firms to anticipate future demand more accurately and thereby reduce both stockouts and excess holding costs [54]. Regression models and machine learning–based predictive analytics are increasingly integrated into these forecasting systems to enhance accuracy in volatile markets.

In inventory management, the determination of safety stock levels illustrates the statistical basis of operational decision-making. Here, demand variability is often modeled using Normal or Poisson distributions, and managers balance the probability of stockouts against the cost of carrying extra inventory. For instance, the adoption of RFID technologies, when assessed using statistical performance metrics, has been shown to reduce inventory inconsistencies by as much as 30% [55]. Similarly, data-driven project scheduling improvements, validated through key performance indicators (KPIs), contributed to a 25% reduction in time-to-market in integrated management systems [56].

Routing optimization also benefits from statistical insight, where predictive models are combined with optimization algorithms to forecast delivery times and account for traffic variability. This integration strengthens logistical efficiency but introduces computational complexity, particularly in large-scale, global networks. Moreover, while these models improve operational robustness, their reliability is contingent on high-quality, timely data. Forecasting approaches, for example, are vulnerable to disruptions such as pandemics or sudden geopolitical shocks, highlighting the limitations of relying solely on historical trends.

Thus, statistics in supply chain management not only support routine operational choices but also provide evaluative tools for measuring the impact of technological and managerial interventions. However, scalability, model accuracy, and data reliability remain persistent challenges that require careful consideration in both research and practice.

4.3. Statistical Rigor in Emerging Industrial Engineering Paradigms

The growing integration of Industry 4.0 technologies, including Artificial Intelligence (AI), Machine Learning (ML), and the Internet of Things (IoT), has amplified the demand for statistical rigor in industrial engineering. These technologies generate vast and complex datasets, but without robust statistical validation, their outputs risk being unreliable or misleading [57, 58]. Ensuring statistical rigor is therefore central to building trustworthy, explainable, and scalable solutions.

Statistical validation plays a critical role in evaluating and deploying AI and ML models. For classification tasks, metrics such as Receiver Operating Characteristic (ROC) curves, Precision, Recall, and F1-score provide quantitative measures of model performance, while regression-based models are typically assessed using R-squared or Mean Squared Error (MSE). Confidence intervals and error distributions further establish the reliability of these models when applied to predictive maintenance systems, where accurate forecasts of equipment failure can prevent costly downtime. In real-time process control, statistical principles guide adaptive decision-making: for example, Kalman filters, rooted in statistical estimation theory, enable continuous updates of process parameters, while control charts provide automated detection of deviations and trigger corrective actions.

Table 1: Summary of Case Studies in Statistical Optimization Across Various Industries

Case Study	Industry / Application	Methodology	Key Results	Authors
Optimization of Rubber-Metal Bonded Products	Manufacturing	Full Factorial DOE	Reduced defects and improved product quality	[60]
Linagliptin Manufacturing Process	Pharmaceutical manufacturing	DOE Techniques	Achieved high yield and purity, removing impurities	[61]
Inventory Demand Forecasting	E-commerce / logistics	Uni-Regression Deep Forecasting (BiLSTM, NARX)	MAE = 1.73, $R^2 = 0.94$ (high forecasting accuracy)	[62]
Inventory Optimization in E-Commerce	Retail / supply chain	BO-CNN-LSTM (Bayesian-Optimized CNN-LSTM)	Improved forecasting accuracy and operational efficiency	[63]
Energy Efficiency in Plastic Manufacturing	Industrial manufacturing	Energy Audits, Data-Driven Decision Making	Annual savings of $45,824.56 (23.68% reduction)	[64]
Energy Management in Cement Production	Manufacturing	SPC, Energy Audits	Improved energy efficiency and reduced costs	[65]
Predictive Maintenance in Manufacturing	Industrial maintenance	Predictive Analytics, Regression	Reduced downtime and maintenance costs	[66]
Healthcare Disease Detection	Healthcare	Statistical Analysis, Machine Learning	Early detection and personalized care via EHR	[67]
Aerospace Manufacturing Optimization	Aerospace	Simulation Modeling, Linear Programming	Optimized processes, reduced costs, improved quality	[68]

Despite their promise, the integration of statistical methods with AI and ML presents significant challenges. One persistent trade-off is between accuracy and interpretability: complex models such as deep neural networks often deliver superior predictive power but operate as “black boxes,” limiting trust and explainability in industrial contexts. Moreover, the performance of these approaches is heavily dependent on the availability of clean, representative, and unbiased data. Flawed or incomplete datasets can lead to erroneous predictions, undermining decision-making at scale. Finally, industrial organizations often face barriers related to computational infrastructure and workforce capabilities. The deployment of advanced statistical and AI-driven tools requires not only substantial digital infrastructure but also skilled personnel capable of interpreting and maintaining these systems [59].

In sum, statistical rigor provides the essential foundation for embedding AI, ML, and IoT into industrial engineering. While these paradigms expand the frontiers of predictive maintenance, real-time control, and system optimization, their effectiveness ultimately hinges on balancing predictive accuracy with interpretability, ensuring data quality, and addressing the infrastructural and skill-related barriers to implementation.

5. Case Studies in Data-Driven Industrial Optimization

This section bridges the gap between statistical theory and industrial application by showcasing how core methodologies are successfully deployed in real-world settings. Through diverse case studies, it becomes clear how statistical techniques such as DOE, regression, SPC, forecasting, and simulation directly enhance manufacturing quality, operational efficiency, sustainability, and cost reduction. Table 1 provides a structured summary of recent research, outlining the applied methodologies, industrial contexts, and key findings.

The reviewed case studies demonstrate the versatility and value of statistical optimization across industrial contexts. In logistics and e-commerce, advanced forecasting models have markedly improved inventory planning and operational responsiveness [62 – 63]. Manufacturing industries continue to benefit from DOE and SPC, which drive significant improvements in product quality and process robustness [60, 61, 65]. Similarly, data-driven audits in energy-intensive sectors have delivered substantial cost savings, emphasizing the role of statistical approaches in advancing industrial sustainability goals [64].

The integration of machine learning with traditional statistical methods has also proven effective in predictive maintenance and real-time optimization, particularly within complex environments such as large-scale industrial systems and aerospace manufacturing [66, 68]. The healthcare case study [67], while slightly outside traditional industrial domains, demonstrates broad applicability. Its relevance becomes clearer when considered in pharmaceutical operations or hospital logistics, where optimization plays a direct role in efficiency and patient outcomes.

Table 2: Challenges and Limitations of Statistical Methods in Industrial Optimization

Category	Challenge / Limitation	Explanation	Authors
Technical Challenges	Data quality issues	Reliable analysis depends on high-quality data.	[69]
	Incompleteness	Missing values skew interpretations.	[70]
	Noise and bias	Data errors distort results.	[71]
	Timeliness	Outdated data undermines real-time decisions.	[70]
	Model complexity	Advanced models are difficult to implement, interpret, and validate.	[72]
	Computational demands	High dimensionality requires significant computing power.	[69, 73]
	Model selection complexity	Choosing suitable models for nonlinear systems remains a challenge.	[74]
Managerial Issues	Lack of expertise	Gaps in statistical literacy hinder effective use.	[72, 75]
	Resistance to change	Organizations may resist new methods due to tradition or uncertainty.	[73]
Resource Constraints	High implementation costs	Infrastructure for advanced analytics is costly.	[76]
	Lack of standardization	Absence of unified frameworks complicates validation and reproducibility.	[77]
	Specialized knowledge gaps	High-dimensional data demands advanced expertise.	[78, 79]
Contextual Challenges	Generalized model limitations	Models may overlook industry-specific dynamics.	[80, 81]
	Dynamic system complexity	Dynamic behaviors in real systems are difficult to model accurately.	[82, 83]

Overall, these applications demonstrate that statistical techniques not only improve technical performance but also shape strategic decision-making in modern industries. At the same time, they show that persistent difficulties remain, including data preparedness, model interpretability, and the need for domain-specific expertise. Addressing these challenges will be essential for scaling statistical optimization methods to increasingly complex industrial systems.

6. Challenges and Limitations in Implementing Statistical Optimization

Although statistical techniques are powerful tools for industrial optimization, their practical deployment is often hindered by a range of barriers. These include technical, managerial, resource-related, and contextual difficulties that can limit the accuracy, applicability, and scalability of optimization initiatives. A comprehensive overview of these challenges is provided in Table 2, which categorizes key limitations identified in recent research.

While these barriers vary in scope, several emerge as particularly critical across the literature. Data quality issues, spanning incompleteness, noise, bias, and timeliness, represent the most fundamental challenge, as even the most advanced models cannot compensate for poor input data. Model complexity and computational demands further exacerbate these difficulties, particularly when addressing high-dimensional or nonlinear systems. At the organizational level, the lack of expertise and resistance to change remain persistent barriers that hinder adoption, while high costs and the absence of standardized frameworks constrain scalability and reproducibility across industries.

These challenges are often interdependent: for example, poor data quality amplifies computational load, while limited expertise reduces the ability to manage complex models effectively. Taken together, they demonstrate that addressing isolated barriers is insufficient. A holistic approach, integrating technical improvements, organizational change management, and investment in skills and infrastructure, is essential for realizing the full potential of statistical optimization in industrial practice.

7. Future Trajectories and Emerging Trends in Statistical Optimization

The future of industrial optimization is being reshaped by advances in artificial intelligence (AI), machine learning (ML), the Internet of Things (IoT), and big data technologies. These developments are not peripheral to statistical optimization, they are fundamentally transforming how statistical methods are used, scaled, and interpreted in complex industrial systems (Figure 1). The framework shows the convergence of several key technologies: IoT and Big Data serve as the foundational data source, which is processed and analyzed by Statistical Methods and AI/ML. The insights from these analyses inform the development and refinement of Digital Twins, which act as a central hub for simulation and diagnostics. This integrated approach ultimately leads to a robust and comprehensive integrated optimization framework.

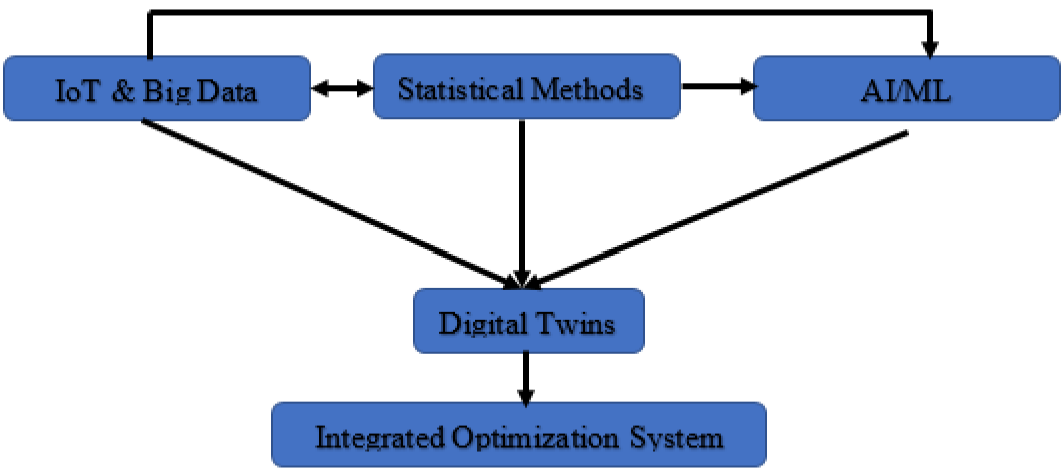


Figure 1. A proposed conceptual framework for data-driven industrial optimization.

One major trajectory is the convergence of traditional statistical techniques with AI and ML to form hybrid systems. This integration combines the theoretical rigor and transparency of statistical inference with the adaptability and performance of intelligent algorithms in high-dimensional, nonlinear contexts [84]. For example, regression-based time-series models can be enhanced with neural networks to improve demand forecasting while still preserving statistical metrics for confidence and error estimation.

The rise of big data analytics further expands the scope of statistical optimization. The sheer volume and velocity of IoT-enabled industrial data demand scalable frameworks capable of near real-time analysis. This shift accelerates the adoption of robust outlier detection, online learning, and stream-based inference, challenging the conventional reliance on sample-based methods. Real-time optimization is increasingly essential in smart manufacturing, where AI-enhanced Statistical Process Control (SPC) systems detect irregularities, adjust parameters instantly, and prevent quality deviations [85, 86].

A related advancement is the integration of digital twin technology, which depends on precise statistical models to replicate physical systems in virtual environments. These twins enable scenario testing, diagnostics, and simulation before deployment, but they also require dynamic models capable of self-correcting in real time as feedback is received [87].

Looking ahead, several emerging challenges will shape future research and practice. These include uncertainty quantification in autonomous systems, algorithmic fairness, and the interpretability of increasingly complex models [88 – 89]. Addressing these issues will require stronger collaboration between industrial engineers, statisticians, data scientists, and computer scientists.

Finally, statistical optimization is moving beyond static, model-centric approaches toward flexible, data-intensive systems that emphasize sustainability, robustness, and real-time responsiveness [90]. The long-term effectiveness of statistical methods in driving innovation and competitiveness will depend on their ability to adapt to the growing autonomy and interconnectedness of industrial systems.

8. Conclusions

This review has examined the fundamental and emerging importance of statistical methods in industrial system optimization, highlighting both their broad applicability and the need for more flexible, integrated strategies. Core techniques such as Statistical Process Control (SPC), Design of Experiments (DOE), and Regression Analysis remain central to data-driven decision-making in areas like process stability, quality improvement, and predictive maintenance. At the same time, the incorporation of advanced approaches such as machine learning, time-series forecasting, and Monte Carlo simulations reflects a shift toward dynamic, real-time optimization frameworks. The central finding of this review is the critical need to integrate established statistical techniques with cutting-edge technologies such as artificial intelligence, the Internet of Things, and digital twin systems. This convergence is essential for addressing the increasing complexity, scale, and speed of modern industrial environments. Yet, persistent challenges, including data quality, model scalability, and interpretability, demonstrate that optimization is not only a technical issue but also an organizational and methodological one. Future progress will rely on strengthening interdisciplinary collaboration, enhancing model transparency, and building robust systems capable of real-time learning and adaptation. By advancing a comprehensive, statistically grounded optimization paradigm, industrial systems can move closer to achieving operational resilience, sustainable growth, and long-term competitive advantage in the digital era.

Acknowledgement
Not applicable

Authors' Contributions
APA handled conceptualization, methodology, data analysis, writing—original draft, supervision, funding acquisition. OOO contributed to data curation, investigation, formal analysis, writing—review & editing, visualization. Both authors have read and approved the final manuscript and agree to be accountable for their contributions and the integrity of the work.

Conflict of interest
The authors disclose no financial, professional, or personal conflicts of interest that could have affected this manuscript's findings or interpretations. No organization or individual has influenced the study's objectivity.

Funding
The authors confirm that this research was conducted without any financial support from external funding bodies. No funding was received from any organization or institution that could have influenced the content, findings, or conclusions of this study.

Data availability
As this is a review article, no new datasets were created. All data discussed are from published works cited in the references.

Ethical statement
This study followed scientific research and publication ethics. The study had no human participants, animals, or institutional review board permission. The research process followed integrity, transparency, and neutrality to meet the highest academic standards.

References

[1]. Nwabueze MO, Aliyu A, Adegbo KJ, Ikemefuna CD. Enhancing machine optimization through AI-driven data analysis and gathering: Leveraging integrated systems and hybrid technology for industrial efficiency. World Journal of Advanced Research and Reviews. 2024;23(3):1919–1943. doi:10.30574/wjarr.2024.23.3.2882.

[2]. Nachappa MN, Faujdar PK, Agarwal P. Industrial systems optimization from combining Internet of Things and cloud computing. In: Proceedings of the 2024 International Conference on Optimization Computing and Wireless Communications (ICOCWC). IEEE; 2024. p. 1–7. doi:10.1109/ICOCWC60930.2024.10470880.

[3]. Das T. Productivity optimization techniques using industrial engineering tools: A review. International Journal of Scientific Research and Archives. 2024;12(1):375–385. doi:10.30574/ijsra.2024.12.1.0820.

[4]. Batool A, Dai Y, Ma H, Yin S. Industrial machinery components classification: A case of D-S pooling. Symmetry. 2023;15(4):935. doi:10.3390/sym15040935.

[5]. Ito T, Fujita J, Tada A, Matsuzaki H. Industrial machinery and control method thereof. United States Patent No. US 11,256,229. Shibaura Machine Co., Ltd.; 2022.

[6]. Arinze CA, Jacks BS. A comprehensive review on AI-driven optimization techniques enhancing sustainability in oil and gas production processes. Engineering Science & Technology Journal. 2024;5(3):962–973. doi:10.51594/estj.v5i3.950.

[7]. Qiao H, Lin J, Guo Z, Gao L, Ye J, Xu J. Equipment system optimization method based on cost-efficiency analysis. In: Proceedings of the 2023 6th International Conference on Computer Network, Electronic and Automation (ICCNEA). IEEE; 2023. p. 325–329. doi:10.1109/ICCNEA60107.2023.00076.

[8]. Segovia-Hernández JG, Hernández S, Cossío-Vargas E, Sánchez-Ramírez E. Challenges and opportunities in process intensification to achieve the UN's 2030 agenda: Goals 6, 7, 9, 12, and 13. Chemical Engineering and Processing - Process Intensification. 2023;192:109507. doi:10.1016/j.cep.2023.109507.

[9]. Sutanto W, Prajitna SH, Harito C. Enhancing efficiency in industrial metal roof replacement: Utilizing goal programming for cost optimization. JTI: Jurnal Teknik Industri. 2024;10(1). doi:10.24014/jti.v10i1.29476.

[10]. Sadollah A, Nasir M, Geem ZW. Sustainability and optimization: From conceptual fundamentals to applications. Sustainability. 2020;12(5):2027. doi:10.3390/su12052027.

[11]. Hintermüller M, Kunisch K, Leugering G, Rocca E. Challenges in optimization with complex PDE-systems. Oberwolfach Reports. 2021;18(1):419–506. doi:10.4171/OWR/2021/9.

[12]. Dass N, Kumar A, Sungroya M. Optimizing real-time systems with parallel computing: Techniques and challenges. In: Futuristic Trends in Information Technology. Vol. 3. IIP Series; 2024. p. 246–258. doi:10.58532/V3BBIT1P2CH7.

[13]. Bestle D. Optimization processes for automated design of industrial systems. In: Nachbagauer K, Held A, editors. Optimal Design and Control of Multibody Systems. IUTAM 2022. Vol. 42. Cham: Springer; 2024. doi:10.1007/978-3-031-50000-8_1.

[14]. Deeb A, Khokhlovskiy V, Prokofiev V, Shkodyrev V. Optimizing multi-level industrial systems through archive-based data-driven configuration. In: Proceedings of the 2024 International Russian Automation Conference (RusAutoCon). IEEE; 2024. p. 252–259. doi:10.1109/RusAutoCon61949.2024.10694339.

[15]. Montgomery DC. Introduction to Statistical Quality Control. 8th ed. Hoboken (NJ): John Wiley & Sons; 2019.

[16]. Antony J. Design of Experiments for Engineers and Scientists. Oxford: Elsevier; 2006.

[17]. Woodall WH, Montgomery DC. Some current directions in the theory and application of statistical process monitoring. Journal of Quality Technology. 2014;46(1):78–94. doi:10.1080/00224065.2014.11917961.

[18]. Kutner MH, Nachtsheim CJ, Neter J, Li W. Applied Linear Statistical Models. 5th ed. Boston (MA): McGraw-Hill Irwin; 2005.

[19]. Myers RH, Montgomery DC, Anderson-Cook CM. Response Surface Methodology: Process and Product Optimization Using Designed Experiments. 4th ed. Hoboken (NJ): John Wiley & Sons; 2016.

[20]. Kroese DP, Brereton T, Taimre T, Botev ZI. Why the Monte Carlo method is so important today. Wiley Interdisciplinary Reviews: Computational Statistics. 2014;6(6):386–392. doi:10.1002/wics.1314.

[21]. Zhang Y, Li K, Wang S. A hybrid forecasting and optimization approach for production scheduling under uncertainty. Computers & Industrial Engineering. 2020;139:105570. doi:10.1016/j.cie.2019.105570.

[22]. Antony J. Six Sigma vs Lean: Some perspectives from leading academics and practitioners. International Journal of Productivity and Performance Management. 2011;60(2):185–190. doi:10.1108/17410401111101494.

[23]. Kumar R, Singh A, Sharma P. Monte Carlo simulation for risk assessment in industrial investment planning. Computers & Industrial Engineering. 2021;159:107594. doi:10.1016/j.cie.2021.107594.
[24]. Singh R, Jain V. Statistical control limits and their role in process monitoring: A comparative study. Measurement. 2022;190:110636. doi:10.1016/j.measurement.2021.110636.
[25]. Almeida FA, Silva RM, Costa JM. Application of SPC tools in food packaging: A case study in fill-level control. Journal of Quality in Maintenance Engineering. 2023;29(1):45–60. doi:10.1108/JQME-03-2022-0015.
[26]. Chen Y, Wang H, Li X. Real-time SPC integration with IoT-based manufacturing systems. Procedia CIRP. 2020;93:112–117. doi:10.1016/j.procir.2020.03.019.
[27]. Zhou Y, Zhang T, Huang L. Machine learning-enhanced SPC for anomaly detection in smart factories. Computers & Industrial Engineering. 2024;185:108493. doi:10.1016/j.cie.2023.108493.
[28]. Prasad S. Regression. In: Advanced Statistical Methods. Singapore: Springer; 2024. p. 1–29. doi:10.1007/978-981-99-7257-9_1.
[29]. Kalita JK, Bhattacharyya DK, Roy S. Regression. In: Fundamentals of Data Science: Theory and Practice. Elsevier; 2024. p. 69–89. doi:10.1016/B978-0-32-391778-0.00012-0.
[30]. Manea D-I, Țițan E, Șerban RR, Mihai M. Statistical applications of optimization methods and mathematical programming. Proceedings of the International Conference on Applied Statistics. 2020;1(1).
[31]. Desouky O, Al Hamidi YM, Khraisheh MK. Student-led multi-disciplinary approach for the design of experiments in engineering: A methodology. In: Proceedings of the 2024 ASEE Annual Conference and Exposition; Portland, Oregon; 2024 Jun. doi:10.18260/1-2--48023.
[32]. Sequeira SE, Graells M, Puigjaner L. Integration of available CAPE tools for real-time optimization systems. Computer Aided Chemical Engineering. 2001;9:1077–1082. doi:10.1016/S1570-7946(01)80173-5.
[33]. Alam MS, Rahman MM, Chowdhury S. ARIMA-based forecasting of solar energy output for smart grid applications. Energy Reports. 2023;9:1123–1135. doi:10.1016/j.egyr.2023.01.045.
[34]. Zhang Y, Li K. Time-series forecasting for energy demand using ARIMA and Prophet models. Applied Energy. 2021;285:116382. doi:10.1016/j.apenergy.2020.116382.
[35]. Chen Y, Zhang H, Liu Q. Hybrid ARIMA and LSTM model for predictive maintenance in manufacturing. Journal of Intelligent Manufacturing. 2022;33(4):987–1002. doi:10.1007/s10845-021-01800-9.
[36]. Rahman T, Adeyemi A, Yusuf M. Probabilistic modeling of maintenance strategies using Monte Carlo simulation. Reliability Engineering & System Safety. 2024;241:108034. doi:10.1016/j.ress.2024.108034.
[37]. Singh R, Adepoju O. Monte Carlo-based supply chain risk modeling under demand uncertainty. International Journal of Production Research. 2020;58(14):4321–4335. doi:10.1080/00207543.2020.1712493.
[38]. Kumar R, Singh A, Sharma P. Enhancing process capability through SPC: A review of recent industrial applications. International Journal of Productivity and Performance Management. 2021;70(4):923–940. doi:10.1108/IJPPM-06-2020-0294.
[39]. Bazaraa MS, Shetty CM. Foundations of Optimization. Vol. 122. Springer Science & Business Media; 2012.
[40]. Bengtsson L. Statistics. In: Electrical Measurement Techniques. Singapore: Springer; 2024. doi:10.1007/978-981-99-8187-8_13.
[41]. DelSole T, Tippett M. Basic concepts in probability and statistics. In: Statistical Methods for Climate Scientists. Cambridge: Cambridge University Press; 2022. p. 1–29. doi:10.1017/9781108659055.002.
[42]. Deprez M, Robinson EC. Regression. In: Machine Learning for Biomedical Applications: With Scikit-Learn and PyTorch. Elsevier; 2024. p. 67–86. doi:10.1016/B978-0-12-822904-0.00008-X.
[43]. Montgomery DC. Design and Analysis of Experiments. 9th ed. Hoboken (NJ): John Wiley & Sons; 2017.
[44]. Vakrilov NV, Kafadarova NM. A review of industry quality improvement strategies through DoE statistical techniques. In: Proceedings of the 2023 XXXII International Scientific Conference Electronics (ET). IEEE; 2023. p. 1–5. doi:10.1109/ET59121.2023.10278638.
[45]. Archana K, Kumar S. Enhancing manufacturing quality through statistical process control. Journal of Advanced Science and Technology. 2023;20(2):176–181. doi:10.29070/3xd2x680.
[46]. Roy RK. A Primer on the Taguchi Method. 2nd ed. Dearborn (MI): Society of Manufacturing Engineers; 2010.
[47]. Cerda-Flores SC, Rojas-Punzo AA, Nápoles-Rivera F. Applications of multi-objective optimization to industrial processes: A literature review. Processes. 2022;10(1):133. doi:10.3390/pr10010133.

[48]. Perez HD, Grossmann IE. Recent advances in computational models for the discrete and continuous optimization of industrial process systems. In: Quintela Estévez P, Coll B, Crujeiras RM, Durany J, Escudero L, editors. Advances on Links Between Mathematics and Industry. Vol. 15. Cham: Springer; 2021. p. 1–31. doi:10.1007/978-3-030-59223-3_1.
[49]. Heim HP. The statistical regression calculation in plastics processing: Process analysis, optimization, and monitoring. Macromolecular Materials and Engineering. 2002;287(11):773–783. doi:10.1002/mame.200290006.
[50]. Martínez EC. The statistical simplex method for experimental optimization with process data. Computer Aided Chemical Engineering. 2005;20:31–36. doi:10.1016/S1570-7946(05)80127-0.
[51]. Aly NA, Murti G. Statistical process control (SPC) implementation: A step-by-step program. International Journal of Materials and Product Technology. 1991;6(1):1–8. doi:10.1504/IJMPT.1991.036644.
[52]. Pauji NI, Dewi, Nugraha. Peningkatan kualitas kemeja menggunakan metode statistical process control (SPC). Bandung Conference Series: Industrial Engineering Science. 2024;4(2):776–786. doi:10.29313/bcsies.v4i2.14689.
[53]. Smirnova OO, Karlina EP. Theoretical foundations of building management systems for modern enterprises. Vestnik of Astrakhan State Technical University. Series: Economics. 2024;(3).
[54]. Ishola CY, Olabode AS. Fuzzy approach for determining statistical process control (SPC) tools location on the production floor. Data Science Journal of Computer Applications and Informatics. 2024;8(1):25–36. doi:10.32734/jocai.v8.i1-17137.
[55]. Yee JT, Oh SC. Technology Integration to Business: Focusing on RFID, Interoperability, and Sustainability for Manufacturing, Logistics, and Supply Chain Management. Springer Science & Business Media; 2012. doi:10.1007/978-1-4471-4391-8.
[56]. Costa I, Fernandes G, Tereso A. Integration of project management with NPD process: A metalworking company case study. In: Proceedings of the 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC). 2017. p. 1191–1200. doi:10.1109/ICE.2017.8280016.
[57]. Echempati R, Mahajerin E, Sala A. Effective integration of mathematical and CAE tools in engineering. In: Proceedings of the 2008 ASEE Annual Conference and Exposition; 2008. p. 13–467.
[58]. Vadlaman S, Kankanampati PK, Goel O, Jain A, Khan S. Integrating AI and machine learning for optimized supply chain and procurement systems. University Research Report. 2022;9(4):540–560. Available from: University Research Report.
[59]. Rane NL, Mallick SK, Kaya Ö, Rane J. Emerging trends and future directions in machine learning and deep learning architectures. In: Applied Machine Learning and Deep Learning: Architectures and Techniques. Deep Science Publishing; 2024. p. 192–211. doi:10.70593/978-81-981271-4-3_10.
[60]. Chaurasia AK, Aggarwal A, Khanna P. Mathematical modelling to predict bead geometry in MIG welded aluminium 6101 plates. Materials Today: Proceedings. 2024;113:152–158.
[61]. Wang H, Chen K, Lin B, Kou J, Li L, Wu S. Process development and optimization of linagliptin aided by the design of experiments (DoE). Organic Process Research & Development. 2022;26(12):3254–3264. doi:10.1021/acs.oprd.2c00230.
[62]. Aldahmani E, Alzubi A, Iyiola K. Demand forecasting in supply chain using uni-regression deep approximate forecasting model. Applied Sciences. 2024;14(18):8110. doi:10.3390/app14188110.
[63]. Liu R, Vakharia V. Optimizing supply chain management through BO-CNN-LSTM for demand forecasting and inventory management. Journal of Organizational and End User Computing. 2024;36(1):1–25. doi:10.4018/JOEUC.335591.
[64]. Anoune K, Ghazi M, Ghazouani M, Nasiri B, Zebraoui O. Empowering industry through energy auditing: A case study of savings and sustainability. International Journal of Applied Power Engineering. 2024;13(4):952–962. doi:10.11591/ijape.v13.i4.pp952-962.
[65]. Akhiyar D, Nofriadiman. Study of energy management and efficiency in industrial processes: Analysis of implementation and its impact on productivity. Jurnal Teknik dan Teknologi Tepat Guna. 2024;3(2). doi:10.62357/j-t3g.v3i2.347.
[66]. Ramadhan GJM, Niam S. Big data analytics: Techniques, tools, and applications in various industries. Jurnal Ar Ro'is Mandalika (ARMADA). 2023;3(2):56–65. doi:10.59613/armada.v3i2.2835.
[67]. Ponnusamy V, Nallarasan V, Rajasegar RS, Arivazhagan N, Gouthaman P. Modern real-world applications using data analytics and machine learning. In: Singh P, Mishra AR, Garg P, editors. Data Analytics and Machine Learning. Vol. 145. Studies in Big Data. Singapore: Springer; 2024. doi:10.1007/978-981-97-0448-4_11.
[68]. Uzair M, Khurshid SH, Iftikhar MU, Ahmed H, Hameed H, Ali H. Optimizing industrial processes in beverages and aerospace sector: A case study analysis. MATEC Web of Conferences. 2024;398:01015. doi:10.1051/matecconf/202439801015.

[69]. Grossmann IE. Challenges in the application of mathematical programming in the enterprise-wide optimization of process industries. Theoretical Foundations of Chemical Engineering. 2014;48(5):555–573. doi:10.1134/S0040579514050182.
[70]. Gehrmann J, Beyan O. Data quality in medical real-world data: An oncological use case. Studies in Health Technology and Informatics. 2024;316:9–13. doi:10.3233/SHTI240332.
[71]. Eckel F, Stampa J, Timkó M, Vecsey L. Data quality and availability tests of public seismometer data in Europe. In: Proceedings of the EGU General Assembly 2023; Vienna, Austria; 2023 Apr 24–28. doi:10.5194/egusphere-egu23-7852.
[72]. Montgomery DC. Opportunities and challenges for industrial statisticians. Journal of Applied Statistics. 2001;28(3–4):427–439. doi:10.1080/02664760120034153.
[73]. Filipić B, Tušar T. Challenges of applying optimization methodology in industry. In: Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation. ACM; 2013. p. 1103–1104. doi:10.1145/2464576.2482688.
[74]. Murari A, Lungaroni M, Rossi R, Spolladore L, Gelfusa M. Development of information theoretic model selection criteria for the analysis of experimental data. Complexity. 2022:9518303. doi:10.1155/2022/9518303.
[75]. Fang D, Varshney KR, Wang J, Ramamurthy KN, Mojsilovic A, Bauer JH. Quantifying and recommending expertise when new skills emerge. In: Proceedings of the 2013 IEEE 13th International Conference on Data Mining Workshops. IEEE; 2013. p. 672–679. doi:10.1109/ICDMW.2013.33.
[76]. Salmeron JL, Correia MB, Palos-Sanchez PR. Complexity in forecasting and predictive models. Complexity. 2019:8160659. doi:10.1155/2019/8160659.
[77]. DiRenzo GV, Hanks E, Miller DAW. A practical guide to understanding and validating complex models using data simulations. Methods in Ecology and Evolution. 2023;14(1):203–217. doi:10.1111/2041-210X.14030.
[78]. Uribe C, Isaza C. Expert knowledge-guided feature selection for data-based industrial process monitoring. Revista Facultad de Ingeniería Universidad de Antioquia. 2012;65:112–125.
[79]. Gurcan F, Sevik S. Expertise roles and skills required by the software development industry. In: Proceedings of the 2019 1st International Informatics and Software Engineering Conference (UBMYK). IEEE; 2019. p. 1–4. doi:10.1109/UBMYK48245.2019.8965571.
[80]. Deman V, Ciantar M, Naudin L, Castera P, Beignon A-S. Combining Zhegalkin polynomials and SAT solving for context-specific Boolean modeling of biological systems. IEEE/ACM Transactions on Computational Biology and Bioinformatics. 2024. doi:10.1109/TCBB.2024.3456302.
[81]. Stremersch S, Gonzalez J, Valenti A, et al. The value of context-specific studies for marketing. Journal of the Academy of Marketing Science. 2023;51(1):50–65. doi:10.1007/s11747-022-00872-9.
[82]. Briand L, Bianculli D, Nejati S, Pastore F, Sabetzadeh M. The case for context-driven software engineering research: Generalizability is overrated. IEEE Software. 2017;34(5):72–75. doi:10.1109/MS.2017.3571562.
[83]. Edmonds B. Complexity and context-dependency. Foundations of Science. 2013;18(4):745–755. doi:10.1007/s10699-012-9303-x.
[84]. Zhao J, Li L, Luo LX, Liu F, Wang YG. Analyzing and applying big data technologies based on industrial internet data. Journal of Physics: Conference Series. 2023;2665(1):012003. doi:10.1088/1742-6596/2665/1/012003.
[85]. Jiang Y. Research on improving production efficiency in smart manufacturing using IoT. In: Proceedings of the 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE). 2024. p. 450–454. doi:10.1109/MLISE62164.2024.10674615.
[86]. Parashare GR. AI-enabled statistical process control for semiconductor manufacturing quality improvement. International Journal of Scientific Research and Management. 2025;13(6):2279–2300. doi:10.18535/ijsrm/v13i06.ec07.
[87]. Mosleuzzaman M, Arif I, Siddiki A. Design and development of a smart factory using Industry 4.0 technologies. Academic Journal of Business Administration, Innovation and Sustainability. 2024;4(4):89–108. doi:10.69593/ajbais.v4i04.131.
[88]. Torabi M. Interdisciplinary approaches in business studies: Applications and challenges. International Journal of Business and Management. 2023;18(5):1–33. doi:10.5539/ijbm.v18n5p33.
[89]. Hensel M, Bier H. Interdisciplinary data-integrated approaches. SPOOL. 2022;9(1):3–4. doi:10.47982/spool.2022.1.00.
[90]. Krzywański J, Sosnowski M, Grabowska K, Zylka A, Lasek L, Kijo-Kleczkowska A. Advanced computational methods for modeling, prediction and optimization: A review. Materials. 2024;17(14):3521.