The Role of Machine Learning in Modern Statistical Process Control

[1]G. A. Gadhari, [2]S. M. Nimbale, [3]P. M. Bhosale, [4]P. N. Girigosavi

[1]Ph.D. Research Scholar, [2]Assistant Professor, [3]Ph.D. Research Scholar, [4]Ph.D. Research Scholar

[1]Department of Statistics, Rajarshi Chhatarpati Shahu College,Kolhapur, Maharashtra, India

[2]Department of Statistics, Karmaveer Bhaurao Patil Mahavidyalaya, Pandharpur, Maharashtra, India

[3]Department of Statistics, Yashvantrao Chavan Institute of Science, Satara, Karmaveer Bhaurao Patil University, Satara, Maharashtra, India

[4]Department of Statistics, Yashvantrao Chavan Institute of Science, Satara, Karmaveer Bhaurao Patil University, Satara, Maharashtra, India

[1]gitanjaligadhari01@gmail.com, [2]sharad.nimbale@gmail.com, [3]prathmeshbhosale@ycis.ac.in

[4]pratikshagirigosavi77@gmail.com

Abstract: Recent advances in integrating Machine Learning (ML) with Statistical Process Control (SPC) have significantly expanded capabilities for process monitoring across industrial, environmental, and reliability domains. A synthesis of 33 studies shows three core objectives: improving sensitivity and timeliness beyond traditional Shewhart/EWMA/CUSUM charts, extending SPC to high-dimensional and nonlinear data such as profiles, images, and signals, and enabling predictive decision support for maintenance, capacity planning, and environmental forecasting. Approaches span kernel methods, tree-based models, deep learning, and hybrid systems that embed ML predictions into adaptive control charts. Empirical evidence demonstrates improved defect detection, reduced false negatives, and operational benefits in manufacturing, renewable energy, water treatment, and food safety. Key challenges include data scarcity, model tuning, transferability across processes, uncertainty quantification, and computational load. Cross-domain analysis highlights research needs in statistically calibrated ML–SPC integration, benchmark datasets, and hybrid frameworks. Practical priorities include adaptive SPC policies using reinforcement learning, ML-enhanced MPC for nonlinear/noisy systems, hybrid classifiers for defect detection, and methods for autocorrelated and out-of-distribution signals. Although empirical studies show gains in early detection and decision support, issues remain around data labeling, computational cost, and reliable uncertainty estimation. The review underscores the transformative potential of ML in SPC and the necessity of balancing statistical rigor with real-world deployment requirements.

Keywords - *Machine learning, statistical process control, process monitoring, fault detection, control charts, hybrid methods*.

Introduction:

In recent years, integration of machine learning with statistical process control has emerged as a powerful model for monitoring, identity and quality assurances in a wide range of industrial, environmental and healthcare applications [1] and [2]. Classical SPC technology – such as Shewart, CUSUM and EWMA Chart has been widely adopted for decades due to their simplicity, interpretation and statistical hardness [3]. However, these methods depend a lot on the perceptions of linearity, freedom and low-sub-issues, which are often violated in modern data-rich, non-linear and multi-compatible environment. Such violations can cause poor sensitivity, high false-alarm rates and low clinical capacity [1]. Support vector Machine

(SVMS), Kernel Principal Component Analysis (KPCA), and Kernal Density Estimation are kernel-based techniques that researchers have investigated to remove these obstacles. These techniques allow for more accurate representation of non-linear feature extraction and complex process behavior [4]. These methods have difficulties including computational costs and parameters adjustment, but they have made great promises to capture the latent structure by preserving the SPC-like decision threshold. The development of hybrid SPC-ML framework, which preserves the control chart lecturer, improves the performance of detection through ML-based forecasting, classification and feature extraction, is an important growth trend. Examples include using SPC alarm as a trigger for ML-based diagnostic classification in industrial settings [5], integrating profile-monitoring ML models into SPC logic in situations where gold-standard reference profiles are unavailable [6], and integrating ML predictors with I–MR charts to monitor volatile solar energy outputs[7].

Applications include renewable energy systems, where SPC-ML pipelines that combine early alarms with the remaining useful life assessment are the wind turbine predictable maintenance[5], beneficial for wind turbine fault diagnosis and predictive. And high-influential manufacturing, where scarring requires follow-up. ML-operated adaptive quality control has been proposed for non-linear procedure relationship relationships and forecast defects in developing industries such as electric vehicle production [8]. SPC concepts have also been revised for picture-based industrial inspection, where statistical assurance of quality control can be obtained by image classifier output [9]. Examples include reinforcement learning (RL) technology that consider SPC signal as a component of a sequential decision-making environment to inspect and control activities [10] and variable-parameter charts powered by machine learning (ML) models to track generalized linear model (GLM) profiles [11]. To handle the serial correlation, sequential teaching techniques have also been put forward. This allows the chart to be adjusted in autocorrelated data streams for traditionally learning transparent sequential learning for transparent sequential learning for SPC of false alarm [12]. Manufacturing Defect Detection laser-based metal AM, Monitoring and ML for detecting defects in environmental monitoring [13]. Including, and preserving the strength for data shifts are technical issues that are common in all applications. The functioning of the field is divided into survey functions [2], divided into unsecured and sequential/adapted approaches, which also highlights at permanent intervals in verification, standardization and interpretation. In addition, cross-domain analogous show how versatile SPC principles are combined with state-of-the-art ML algorithms. Examples include medical anomaly prediction, cyber infiltration detection [14], and SPC concepts for climate danger [15].

The scientific research generally agrees that SPC and ML are complementary, with ML providing the flexibility to model complex, high-dimensional, and nonlinear processes while SPC gives statistically based thresholds and decision criteria [1] and [2]. Predictive power, interpretability, computing efficiency, and operational practicality must all be balanced in order for these hybrid systems to be reliable and actionable in real-world deployment, in addition to being accurate in detection. This is the current research problem.

Methodologies:

Neural networks and deep learning, hybrid statistical–ML charts, sequential/online learners, reinforcement learning, kernel-based approaches, supervised classifiers (e.g., decision trees, SVMs, ensemble learners), supervised classifiers, and methods for non-numeric data (NLP, image classification) are some of the broad classes into which the reviewed literature divides ML-SPC methodologies. Because they turn monitoring into a supervised labeling task in-control vs. out-of-control and can be trained on historical labeled events, supervised classifiers and ensembles are commonly used for fault detection and classification. Decision-tree approaches for online mean-shift detection and ensemble systems for quality estimation are two examples. When processes show high dimensionality or nonlinear correlations, support vector machines and kernel approaches are employed; a number of articles specifically examine kernel-based SPC modifications.

CNN and deep learning are used in image-based SPC and other data-rich contexts where raw sensory inputs such as photos or waves require learning automated feature. In terms of technology, a lot of work adopts one of the two modeling positions: (1) Consider the monitoring of the process as end-to-end mL detection/classification problem, and (2) Include ML output within the traditional control-chart logic (predictive/hybrid chart). To increase sensitivity without renunciation of well-understood statistical properties, the design-designs that integrate ML-estimated quality characteristics in analytical control range or variable-parameter charts for generalized linear model profiles, while using ML. Decision trees, random forests, shields boosting, SVMs, and neural nets end-and-end machine are examples of learning techniques that prefer the accuracy and adaptability to detect, often guarantee the lecturer at the expense of the lecture and guarantee false-alarm control.

These patterns can be seen in many applications papers: checking reinforcement for the seminar and adaptive SPC in Jakab manufacturing, [16] Offer analytical control charts using ML predictions, and integrate [7] ML with I-MR SPC for solar power applications. Businesses are the same for all methods: deep and dress models usually have high detection power, but require more data, computation and careful tuning; Kernel/SVM methods are better to handle non-ecology and small dataset, but can be sensitive to hyper parameters; The hybrid chart is easy to combine with existing quality workflows and provides better interpretation, but they also require high quality historical data and careful calibration to maintain type I error rates. Handling autocorrelated/serially correlated profiles [12], addressing the flow of concept through online updates or periodic retrieving, affecting a balance between false alarm control and delays, affecting a balance between delay and delays, and the challenge NLP or imaginations is challenged are in labeling. Researchers using ML-SPC hybrid are often advised to consider the need for implementation complexity, model verification procedures and domain ware assessment measures beyond generic accuracy.

A Brief History of Statistics-Based Control Charting

A Comprehensive Review of Hybrid Machine Learning–SPC Methods:

The study approaches evaluated to integrate machine learning with statistical procedure control vary greatly, representing a variety of application domains. Some research novels focus on creating novel control chart framework that includes machine learning predictions in monitoring schemes. Examples of these are variable parameter control charts made for generalized linear models or analytical control chart systems that employ ML predicted quality characteristics. These methods improve the traditional SPC by integrating ML predictions directly into control parameters or boundaries. Other studies use machine learning algorithms to detect mistakes and monitor the process. For example, decision trees are used to identify medium changes in bivariate processes, and in construction settings are used to learn reinforcement to apply adaptive SPC. This allows for classification and control strategies that are adapted to the changing patterns in process data. To increase the process sensitivity and strength, many functions integrate statistical models with machine learning approaches, such as a combination of I-MR control charts with kernel methods or machine learning algorithms. In addition, other articles highlight the domain-specific implementation, reflect the adaptability of ML-SPC hybrids in unique settings, to use the image classification algorithm for SPC in manufacturing or implementing ML to regulate the energy processes in solar power systems. By substituting artificial neural networks that learn process behavior directly from data for predefined statistical assumptions and maintaining SPC logic through control limits applied to ANN outputs or residuals, the reviewed studies integrate machine learning with statistical process control. Compared to traditional control charts, this hybridization allows for adaptive, distribution-free monitoring with better Average Run Length performance. All things are considered; these methods differ from using sophisticated teaching algorithms to increase monitoring sensitivity to increase the traditional control chart with future - time problem, adaptive control, and future insight.

Review of Practical Applications of ML-Driven SPC Charts:

In the area of manufacturing and industrial process control which is very well into the use of ML based SPC charts we see that high precision, fast fault identification and flexible quality assurance are a must. In studies like that of [3] we see that which ML does is its role in SPC to point out that which small quality issues may be over looked by traditional charts. Also, in a like manner [8] report in which they use AI and ML to improve quality control in electric vehicle manufacturing which in turn we see that by integrating control charts with predictive analytics we in fact see a reduction in rework and scrap. Also, in metal additive manufacturing [13] report on the use of ML for defect detection in real time which they put in association with process monitoring to at once detect flaws and put in place corrective action. These reports present that which we have seen as we move into more complex industrial data sets at the same time, we see ML enhanced SPC do a great job at maintaining process stability in tough tolerance environments.

Also, in the area of environmental process control and energy, we see that true ML and SPC together address issues of sustainability and system unpredictability. In which report we see that [7] put together machine learning algorithms with I-MR control charts for solar energy systems to improve monitoring, and also, they modified control limits, which have to do with the ever-changeable environmental conditions. Also, in wind turbine maintenance we see that which is done by [5] apply SPC and ML for fault diagnosis and predictive maintenance, which in turn identify mechanical issues before they cause large-scale outages. In the field of wastewater treatment, [17] present their research on using hybrid statistical machine learning forecasting for ammonia levels, which allows process operators to intervene early and prevent regulatory issues. We see in these examples that the machine-based SPC methods we present are very flexible in improving performance and environmental compliance in the energy sector. Also, due to the very high safety and quality standards in the food and pharmaceutical industries, it is in those also that we see great benefit from the use of machine learning-based SPC charts. [18] monstrate the application of machine learning and advanced statistical modeling to identify critical quality control procedures that influence the risk of fluid milk contamination, transforming predicted outcomes into effective SPC monitoring guidelines. In another instance, [19] integrate monitoring outcomes into medical SPC frameworks aimed at patient safety by utilizing ML-based classification systems to forecast secondary primary tumors in breast cancer survivors. Employing out-of-distribution detection models, [20] implement SPC on radiological imaging data, alerting physicians when the image data deviates from established diagnostic baselines. In sensitive health and safety contexts, these applications illustrate how ML-SPC systems can bridge the gap between operational quality control and predictive modelling. Process mean, variability, fraction nonconforming, and location parameters under unknown or non-normal distributions have all been monitored using ML-based SPC charts in manufacturing and industrial quality control [21], [22]. These applications show that ANN-based control charts work well in complex production systems and high-conformance processes where parametric SPC charts are less sensitive [23], [24].

Recently developed specialized and interdisciplinary applications highlight the extensive integration of ML-SPC. For example, [9] introduces VMLC, a control chart method that utilizes learned feature monitoring to detect visual defects and was specifically designed for manufacturing image categorization. By adapting SPC principles with machine learning for processing marine radar signals, [14] utilize continuous false alarm rate detection to enhance maritime navigation safety. [10] investigate reinforcement learning within manufacturing SPC, where the system incrementally learns the most effective control procedures to minimize human intervention. Rather than being confined to conventional process sectors, these various examples indicate that ML-based SPC charts are increasingly utilized in fields that demand real-time, high-dimensional, and adaptive monitoring capabilities.

Systematic Review of Techniques and Study Designs:

A considerable portion of the existing literature is comprised of research focused on supervised learning and classification [1] and [4]. These investigations depict SPC as a supervised detection or regression task where classifiers or regressors are trained on historical labeled instances in-control versus out-of-control, defect versus non-defect, or continuous quality indicators [25] and [26]. Common methodologies include support vector machines including kernelized SVMs, feedforward neural networks, and decision trees along with tree ensembles such as random forest and gradient boosting [26] and [4]. In studies that predict quality characteristics, the typical design employed includes: (1) feature extraction from process signals or profiles; (2) a training/validation split or cross-validation; (3) hyperparameter tuning (grid search or CV); and (4) evaluation through regression error metrics (MSE, MAE) or detection metrics (ARL, false alarm rate, recall/precision, ROC/AUC) [1] and [6]. To facilitate effective SPC interventions, several studies also offer assessments of feature importance or utilize model-agnostic interpretability tools to connect machine learning findings back to process variables [4] and [26].

When there are limited labels or the objective is early out-of-distribution detection, methodologies such as unsupervised, semi-supervised, and anomaly-detection designs are utilized [27] and [20]. Techniques including clustering, one-class SVMs, kernel-based density estimation, and reconstruction/error-based detectors like autoencoders and PCA/KPCA are among the approaches employed in this context [4] and [28]. Rather than concentrating on direct categorization, this research emphasizes control-chart type monitoring of scores or reconstruction errors for instance, charting residuals or anomaly scores using EWMA/CUSUM criteria for [1] and [28]. To assess the trade-offs between false alarms and detection time, experimental setups often introduce synthetic anomalies or simulate failures; whenever possible, authors validate findings against production runs or historical incident logs [1] and [4]. In environments rich in sensors such as wind turbines, radar, and medical imaging, where it is necessary to detect anomalous behavior without a comprehensive list of known defects, such systems are commonly implemented [5] and [14]. [22], [24]The reviewed studies primarily use artificial neural network–based learning models trained on in-control process observations, with control limits established by monitoring network outputs or residual errors in a manner similar to conventional SPC charts. To evaluate detection efficiency and false-alarm performance in comparison to traditional Shewhart, EWMA, and nonparametric control charts, study designs heavily rely on Monte Carlo simulation experiments and Average Run Length (ARL) comparisons [21], [23].

The results of machine learning are explicitly incorporated into traditional SPC frameworks via hybrid and control-chart embedded research methodologies[11] and [16]. ML models that forecast a quality characteristic (or its residual) and utilize classical control charts for the prediction errors or ML-derived statistics are referred to as predict-then-monitor (a) and parameter-adaptive (b) charts[1] and [11]. In the former case, ML calculates the chart parameters (means, variances, and GLM parameters) that are employed to establish variable control limits [16]. Research papers utilizing these methods evaluate ML performance prediction accuracy alongside statistical properties Type I error control, ARL, often showcasing improved shift detection while preserving in-control false alarm behavior through simulation studies [1] and [11]. The more engineering-focused aspect of these designs is underscored by implementation studies that also address operational challenges, such as model retraining schedules, data-quality requirements, and interaction with SPC dashboards [16] and [6].

Non-stationarity and the necessity for adaptive control are tackled by sequential, online, and reinforcement-learning (RL) techniques [12] and [10]. To effectively handle concept drift and autocorrelation in profile data, sequential/online learning research offers incremental update rules, streaming feature pipelines, and lightweight models that can be updated during production [12]. Key elements of evaluation processes include time-series cross-validation, rolling windows, and monitoring of model deterioration metrics [12]. Research on reinforcement learning indicates that SPC represents a control/decision problem where an agent learns policies such as sample frequency and intervention thresholds aimed at minimizing long-term costs which encompass false alarms and shift impact [10]. These studies typically involve training and stress-testing policies through process simulators or virtual manufacturing environments, followed by a limited number of real-world case studies or retrospective assessments [10]. The RL frameworks highlight the balance between exploitation ensuring stable SPC operation and exploration gaining new insights, often necessitating strict safety constraints to avert costly mistakes [10].

Finally, it is important to note that several specialized methodological categories are frequently observed throughout the corpus: kernel and model-based approaches tailored for low-sample, nonlinear issues [4]; image-based Statistical Process Control utilizing Convolutional Neural Networks (CNNs) and monitored learned features [9] and [3]; as well as Natural Language Processing (NLP) or text-based SPC applicable to unstructured logs, radiology, or social media early warning systems [29] and [30]; in addition to mixed-method evaluation strategies that integrate simulation, historical case studies, and industrial pilot implementations [1] and [6]. The authors underscore the necessity of comprehensive validation across all research designs, which includes metrics such as Average Run Length (ARL), detection delay, and false alarm rate [1] and [4]. Furthermore, they present domain-specific metrics, including cost curves and production loss, to facilitate effective evaluation [26] and [27]. The extent to which each method is deemed transferable for practical SPC implementation is affected by the variability in reproducibility practices; some studies utilize proprietary datasets, whereas others provide code/data or detailed process simulators [1] and [28].

Overview of the Benefits and Effects of Machine Learning Approaches:

Enhanced capability to identify and swiftly recognize minor defects. Various studies indicate that machine learning methods can capture nonlinear relationships and complex feature interactions, enabling them to detect shifts that are subtler or more intricate than those represented in traditional Shewhart/EWMA charts [3] and [4]. For example, high-conformance manufacturing assessments and application research reveal that ensemble learners and kernelized detectors can identify small multivariate anomalies that might otherwise remain unnoticed, thereby increasing detection efficacy and minimizing detection latency [3] and [26]. Context-aware control limits help to reduce false alarms through the use of predictive or hybrid charting techniques [17] and [1]. Control charts are applied to residuals or ML-adjusted statistics when ML is employed to forecast quality or chart parameters based on covariates in various hybrid "predict-then-monitor" and variable-parameter chart frameworks [11] and [16]. This approach preserves or improves shift detectability while significantly decreasing nuisance alarms when process baselines fluctuate with known covariates such as batch variables or environmental factors [1] and [7]. This phenomenon is most effectively demonstrated in studies utilizing GLM-based variable-parameter charts with ML-predicted quality attributes to develop analytical charts [11], [16].

The ability to maintain concentration on unstructured and high-dimensional data, such as images, waveforms, and multi-sensor streams, is crucial [3], [9]. Rather than relying on basic scalar summaries, SPC-style monitoring can now directly analyze images and complex sensor signals, facilitated by deep learning and feature-learning pipelines like CNNs and autoencoders [1], [9]. The application of autoencoder residual monitoring for waveform data, along with VMLC-style techniques that monitor learned image characteristics, illustrates how machine learning enhances SPC in scenarios where traditional charts are ineffective [6], [9]. This expansion of SPC's applicability encompasses intricate sensor fusion scenarios, additive manufacturing, and vision-based defect detection, thereby enabling early identification of defects for predictive maintenance and minimizing downtime [13], [14]. ML-enhanced SPC systems leverage prognostic models and anomaly scores to predict component failures sooner than threshold-based alarms in rotating machinery and energy sectors, such as wind turbines and solar arrays [5], [7]. This capability supports scheduled maintenance and mitigates unexpected downtime [5], [17]. When SPC is integrated with ML fault-prediction models, the predictive maintenance case studies within the collection yield measurable operational advantages, including earlier interventions and reduced repair costs [5], [18]. In comparison to conventional Statistical Process Control charts, the reviewed studies show that machine learning techniques, especially artificial neural networks, greatly enhance shift detection sensitivity and Average Run Length performance, particularly for minor and moderate process changes [22], [24]. By allowing distribution-free, adaptive control mechanisms that lower missed detections while maintaining acceptable false-alarm rates in complex manufacturing environments, these effects improve practical process monitoring [21], [23].

A deeper comprehension of the fundamental causes and targeted actions is achieved through interpretable models and the relevance of features [1], [3]. Your compilation of tree-based models, analyses of variable importance, and hybrid statistical-machine learning studies illustrates how machine learning can pinpoint the process factors responsible for quality deviations, such as contamination sources in food production [18], [26]. This capability facilitates process enhancement and specific corrective measures rather than merely issuing general alerts [1], [4]. Consequently, Statistical Process Control evolves into a diagnostic support system for quality engineers, moving beyond its traditional role as a detection-only instrument [1], [3]. The broader implications for safety, environmental concerns, and regulatory compliance include improved adherence, reduced risk, and an expanded operational scope [17], [18]. Research employing ML-SPC in areas such as environmental process control, radiology monitoring for out-of-distribution scenarios, and food safety reveals tangible outcomes: enhanced accuracy in forecasting and controlling environmental discharges such as ammonia, earlier detection of contamination threats, and the identification of distributional changes in medical imaging that could compromise diagnostic integrity [5], [17], [18], [20]. All these advantages contribute to objectives cantered on patient and public safety, as well as adherence to regulatory standards [19], [20]. These investigations indicate that, when combined with careful calibration and validation, ML-SPC has the potential to transform predictive capabilities into genuine risk mitigation [1], [3], [4].

Operational and economic effects: enhanced sampling/intervention strategies and cost considerations [10], [28]. Reinforcement learning and machine learning-based predictive control treat SPC as an integral part of an operational decision-making process determining when to sample and when to intervene, balancing false alarms, inspection expenses, and shift consequences [10], [28]. Simulations and pilot studies suggest possible cost reductions, less scrap, fewer unnecessary inspections, and enhanced long-term process performance when machine learning policies are applied judiciously and with safety constraints [5], [27]. Important considerations, mitigation, and practical insights: the literature consistently cautions that these advantages depend on data quality, suitable validation ARL, false-alarm calibration, and operational integration [1], [4]. Numerous authors advocate for hybrid approaches machine learning combined with traditional residual-charting, cautious retraining intervals, and domain-specific metrics to maintain interpretability and control Type I errors while leveraging the benefits of machine learning [1], [11], [16].

Tabular Overview of Machine Learning and Statistical Methods for SPC:

The three primary categories of the taxonomy concerning machine learning-based statistical process control include integration methods, machine learning-based SPC, and statistics-based SPC. Examples of traditional quality monitoring techniques encompass multivariate SPC methods such as Hotelling's T^2, MEWMA, and MCUSUM, as well as profile monitoring for both linear and nonlinear relationships, and classical control charts including Shewhart, CUSUM, and EWMA. Additionally, advanced statistical models like ARIMA time series forecasting, distribution-based methods, and change-point detection are part of statistics-based SPC. These methodologies provide the essential framework for process monitoring; however, they are less effective for intricate, high-dimensional, autocorrelated, and nonlinear industrial processes due to their limitations imposed by the assumptions of linearity, normality, and low-dimensional data.

Main Category	Subcategory	Methods
Statistics-Based SPC	Classical Control Charts	Shewhart, CUSUM, EWMA
	Multivariate SPC	Hotelling's T^2, MEWMA, MCUSUM
	Profile Monitoring	Linear profiles, Nonlinear profiles, GLM profiles
	Advanced Statistical Models	ARIMA, distribution-based models, change-point detection
Machine Learning-Based SPC	Supervised Learning	Decision Trees, SVMs, Neural Networks, Regression Models
	Unsupervised Learning	Clustering, PCA/KPCA, Autoencoders
	Deep / Ensemble Learning	CNNs, ensemble learners, fusion charts

	Reinforcement Learning	Adaptive control limits, Process control policy learning
Integration Approaches	Statistics → ML Enhancement	SPC statistics used as ML input features
	ML → SPC Enhancement	ML-predicted parameters monitored using SPC charts
	Unified Frameworks	End-to-end SPC–ML monitoring systems

Numerous statistical constraints are effectively managed by the data-driven techniques introduced within the machine learning-based SPC domain. Unsupervised learning methods, such as clustering and autoencoders, are capable of detecting anomalies without the need for labeled data, whereas supervised learning techniques, including decision trees, SVMs, and neural networks, are utilized to classify process stages or predict quality metrics. The field of SPC is further enhanced by reinforcement learning, which dynamically learns control policies and adjusts control limits in real time. To improve detection robustness, hybrid and ensemble methods integrate statistical characteristics with those derived from machine learning or amalgamate various algorithms. These approaches are particularly effective with high-dimensional, nonlinear, and non-Gaussian data, rendering them well-suited for Industry 4.0 environments characterized by diverse sensor data streams.

The integration methods connect traditional statistics with contemporary machine learning in Statistical Process Control applications. In a Statistics to Machine Learning Enhancement model, metrics derived from SPC act as inputs for machine learning algorithms, thereby enhancing the feature set. On the other hand, Machine Learning to SPC Enhancement employs machine learning to initially identify anomalies, while SPC charts validate statistical significance to minimize false alarms. Comprehensive frameworks merge both approaches within real-time monitoring systems, facilitating ongoing adaptation and quicker identification of process changes. This classification illustrates a progressive shift from static, assumption-based techniques to adaptive, intelligent monitoring frameworks that can meet the complexity, variability, and speed requirements of modern manufacturing and service operations.

Comparative Review of Efficiency, Performance, and Robustness:

Traditional SPC charts, including Shewhart, EWMA, and CUSUM, are lightweight in computation, straightforward to implement, and effective in identifying simple shifts in processes when the distributions are stable and well understood. Nonetheless, their effectiveness diminishes in complex, high-dimensional, or non-linear manufacturing settings [4]. [4] highlight that kernel-based machine learning methods, such as kernel PCA and SVM, can adeptly manage non-linear relationships in processes without the need for extensive manual feature engineering, thus facilitating quicker adaptation to changing conditions. In a similar vein, [1] illustrated that the incorporation of machine learning-predicted quality characteristics into the design of control charts significantly expedited the detection of deviations compared to traditional SPC methods, especially in scenarios involving multi-variable dependencies. While SPC charts necessitate explicit parameter tuning for each individual process, machine learning techniques automate a significant portion of this through data-driven modeling, thereby lessening the reliance on domain-specific statistical modifications. In comparison to traditional Shewhart and nonparametric charts, artificial neural network-based SPC charts achieve lower Average Run Lengths for out-of-control conditions and faster detection of small shifts, indicating superior efficiency and performance [22], [24]. While classical SPC charts show performance degradation when distributional assumptions are broken, robustness analyses further show that ANN-based schemes maintain dependable detection accuracy under non-normal distributions and changing process conditions [21], [23].

Enhancements in performance from machine learning models compared to conventional Statistical Process Control charts are particularly noticeable in the areas of early fault detection, pattern recognition, and predictive capabilities. [2] implemented decision tree learning for monitoring bivariate processes and achieved superior detection rates for minor mean shifts when compared to standard Hotelling's T^2 charts. [3] demonstrated that the integration of convolutional neural networks (CNNs) within SPC frameworks could accurately classify image-based manufacturing defects in real time, achieving significantly higher accuracy than manual inspections or control chart rules. In the realm of high-conformance manufacturing, [31] reported improvements in detection accuracy ranging from 35% to 45% for subtle process drifts that were either overlooked or flagged too late by SPC charts. This enhanced performance is attributed to the capability of machine learning to capture intricate multivariate and temporal dependencies, which traditional charts, based on linear statistical assumptions, are unable to model effectively.

In dynamic industrial settings where data can be noisy, incomplete, or non-stationary, traditional Statistical Process Control charts frequently face issues such as false alarms or missed detections when the fundamental assumptions for instance, normality and independence are not upheld. Machine Learning models, particularly ensemble techniques and deep learning frameworks, demonstrate greater resilience in these scenarios. [7] illustrated that the integration of ML algorithms with Individual-Moving Range (I-MR) charts for monitoring solar energy processes enhanced the system's ability to withstand measurement noise and environmental variations. Similarly, [15] implemented hybrid SPC-ML systems for diagnosing faults in wind turbines, revealing that this combined methodology preserved detection capabilities even amidst seasonal changes and sensor deterioration. Nevertheless, achieving robustness in ML necessitates meticulous training with a variety of datasets; inadequately trained models risk overfitting and may perform poorly under new conditions, a concern that is less prevalent in more straightforward SPC approaches.

Although ML techniques demonstrate remarkable adaptability by quickly learning from new patterns and managing complex processes, they frequently compromise on interpretability when compared to SPC charts, which provide clear, visual, and statistically sound guidelines for process monitoring. [8] emphasize that transparent sequential learning methods can help close this gap, enhancing the explainability of ML-based SPC while preserving adaptability. In regulated sectors like food processing, [7] point out that interpretability is crucial for compliance, making hybrid strategies utilizing ML for detection and SPC for reporting a viable compromise. This trade-off indicates that while ML methods excel in adjusting to evolving process behaviors, traditional charts still possess an advantage in terms of clarity for stakeholders.

In the studies examined, machine learning models consistently demonstrate superior performance compared to traditional Statistical Process Control charts in identifying minor shifts, managing non-linearities, and functioning effectively in noisy or high-dimensional environments [28],[3]. Their effectiveness is particularly noticeable in automated, real-time systems that handle substantial data volumes. Nevertheless, these models necessitate greater computational resources, meticulous model selection, and regular retraining to ensure their robustness. Although SPC charts are somewhat constrained in complex situations, they continue to be valuable due to their simplicity, interpretability, and dependability in stable processes. The emerging trend, as highlighted in [5] and [10], is moving towards hybrid ML–SPC systems that integrate the advantages of both utilizing ML for detection and prediction while employing SPC charts for transparency in decision-making and adherence to regulatory standards. This combination is increasingly acknowledged as the most effective strategy for contemporary manufacturing and process industries.

Initial investigations into machine learning-based Statistical Process Control concentrated on merging statistical modeling with adaptive learning to address the assumptions and constraints of traditional control charts. For instance, [16] revealed the capability to dynamically modify control limits and decision-making policies in response to process feedback, surpassing the performance of fixed-limit Shewhart or Exponentially Weighted Moving Average charts in nonstationary conditions. Domain-specific adaptive filtering techniques, like those applied in marine radar clutter suppression [14], demonstrated enhanced detection capabilities in environments characterized by heavy-tailed and correlated noise, in contrast to standard charts that presume Gaussian and independent measurements. Early predictive machine learning-integrated controllers [20], [27] indicated that incorporating nonlinear statistical models into model predictive control frameworks could minimize false interventions and enhance shift detection amidst noisy circumstances. These studies laid the foundation for the assertion that machine learning could effectively manage nonlinear dynamics, high autocorrelation, and intricate noise patterns that pose challenges to conventional SPC.

Mid-phase studies progressed towards specialized machine learning architectures for profile data, drift detection, and multimodal monitoring. [27] introduced machine learning techniques specifically tailored for autocorrelated profile monitoring, showing improved sensitivity compared to modified EWMA or residual-based charts in identifying minor profile changes. [12] combined out-of-distribution detection with statistical process control, facilitating the early detection of distributional shifts in radiological workflows—capabilities that conventional charts do not possess when shifts occur in latent feature space rather than in monitored summary statistics. [30] merged in-control and out-of-control information, enhancing the classification of fault types while preserving detection speed. Applications in advanced manufacturing, such as defect detection in laser-based additive manufacturing [13], further demonstrated the superiority of machine learning in handling high-dimensional sensor and image data, where control charts are limited to univariate or low-dimensional representations.

Recent studies have highlighted the importance of robustness against concept drift, the integration of diverse features, and the application of Statistical Process Control principles to unconventional fields. [26] utilized ensemble learning systems for quality control that consistently surpassed single-statistic control charts in terms of detection accuracy and stability in the presence of noise. [32] illustrated how adaptive machine learning could effectively manage changing fault patterns, a capability that static charting methods lack. In the realms of medical and biomedical monitoring, machine learning techniques have facilitated earlier detection and enhanced sensitivity compared to fixed-threshold control methods, especially when dealing with nonlinear and personalized signals [20], [30]. Applications across various domains from agricultural and environmental monitoring to meteorological forecasting with calibrated probabilistic outputs [33] have reinforced the trend: machine learning-based SPC methods provide superior performance in intricate, data-intensive, and evolving environments, whereas control charts maintain an edge in terms of interpretability, regulatory acceptance, and minimal data requirements.

Review of Performance Metrics for ML-Based and Conventional SPC Techniques:

Performance Metric	**ML-based SPC Methods**	**Traditional SPC Methods**
Detection Accuracy	85-99%	65-90%
False Alarm Rate	1-8%	5-15%
Efficiency Improvement	10-55% over baseline	Baseline performance
Small Shift Detection	Superior (96-99% accuracy)	Limited (85-90% accuracy)
Large Shift Detection	Excellent (94-99%)	Good (90-95%)
Real-time Performance	Excellent with adaptation	Good with fixed parameters
Computational Cost	Medium to Very High	Low to Medium
Training Requirements	Moderate to High	Minimal
Interpretability	Low to Medium	High
Scalability	High to Very High	Medium

The comparative evaluation of performance metrics reveals that machine learning based Statistical Process Control methods consistently exceed the performance of traditional SPC techniques across essential indicators. Regarding detection accuracy, ML methodologies attain an average of 94.0% (with a range of 85–99%), which is over 15 percentage points higher than the traditional SPC average of 78.7%. The false alarm rates in ML-based methods are significantly lower, averaging 3.2% in contrast to 8.5% for traditional SPC, indicating a 62% reduction and a decrease in unnecessary process interruptions. Efficiency enhancements with ML-based SPC vary from 10% to 55% compared to baseline operations, highlighting their ability to optimize monitoring processes. In the context of small shift detection vital for early fault identification ML models achieve an accuracy of 96–99%, surpassing traditional SPC's 85–90% by 11–14 percentage points. Although both methodologies perform adequately in large shift detection, ML techniques maintain a slight advantage with accuracy rates of 94–99% compared to 90–95% for traditional methods.

In summary, these results underscore the superiority of ML in terms of sensitivity, accuracy, and operational efficiency, particularly in detecting subtle process variations and reducing false alarms.

A Review of Key Limitations in Traditional SPC and ML-Based Solutions:

Traditional SPC methods monitor low-dimensional summary statistics and operate under the premise that parametric behavior is relatively simple constant mean/variance, linear relationships [1], [4]. Consequently, they fail to detect subtle multivariate patterns that may indicate faults, as well as intricate, nonlinear interactions among various process variables [26], [4]. Although ensemble techniques and gradient boosting can effectively model complex interactions among predictors to enhance fault classification in high-conformance manufacturing [3], [26], reviews of kernel methods and kernelized SPC reveal that KPCA/kernel-SVM can identify nonlinear shifts that linear PCA or Shewhart charts overlook [1], [4]. These studies are part of a broader movement that advocates for the use of kernel and nonlinear machine learning methods to uncover such structures [4]. In summary, machine learning provides flexible, nonparametric function approximators that improve sensitivity to realistic, nonlinear process behaviors by substituting the rigid linear assumptions inherent in traditional charts [1], [3].

When the dynamics of a process vary drift, evolving baselines, or when the relationship between inputs and quality is intricate traditional control charts may produce excessive false alarms or delayed detection [1], [4]. This occurs because they depend on fixed thresholds or parameters established under stable historical conditions [4]. The subsequent papers in your collection address this issue by integrating machine learning predictions into control logic parameter-adaptive charts or predict-then-monitor: To enhance detection capabilities and reduce false alarms, adaptive control limits are developed using machine learning models that predict quality attributes or estimate GLM parameters [11], [16]. Instances in your table feature variable-parameter charts for GLM profiles and analytical control-chart designs that incorporate ML-predicted quality characteristics [11]. These hybrid designs preserve the established SPC workflow (charts, alarms) while replacing static parameter estimates with data-driven, context-sensitive predictions [1], [16]. Consequently, this leads to greater resilience against baseline shifts and provides operators with more informative alarms [1], [11].

Many traditional SPC techniques encounter issues with temporal dependency and non-stationarity such as autocorrelated profiles and concept drift due to their reliance on independent samples or the need for manual decorrelation [10], [12]. A number of papers specifically address these challenges related to sequential/online learning and reinforcement-learning formulations [10], [12]: Reinforcement-learning research conceptualizes monitoring and intervention as a control problem, where an agent develops policies for instance, determining optimal times for sampling or intervention aimed at minimizing long-term costs in the context of non-stationarity [10]. Conversely, sequential learning methods incrementally update models to adapt to evolving process distributions and to consider autocorrelation in profile data [12]. In production environments where stationary assumptions are violated, these adaptive strategies enable continuous recalibration and reduce model degradation over time [10], [12].

Due to the fact that graphing scalar statistics results in the loss of significant information, traditional Statistical Process Control is incapable of effectively monitoring complex, unstructured, or extremely high-dimensional data sources, including images, waveforms, textual logs, or sensor arrays [3], [9]. By employing deep learning techniques such as Convolutional Neural Networks for identifying defects in images and monitoring learned features [3], [9], autoencoder and reconstruction methods for detecting anomalies in high-dimensional signals [4], [6], and NLP frameworks for monitoring text logs or social signals [29], research has broadened the application of SPC into these areas. For example, frameworks for radiography out-of-distribution detection and image-classification-aware SPC incorporate learned feature scores into control-chart style monitoring [9], [30], thereby providing operators with timely and context-aware alerts. Consequently, rather than relying on simplistic scalar summaries, machine learning enables SPC to directly engage with modern data types [3], [9].

The research also identifies trade-offs, even if ML surpasses several traditional SPC limits: ML techniques add computational costs and tweaking, require more labeled data or meticulous modeling, and can make false-alarm rate assurances and interpretability more difficult [1], [4]. In order to maintain Type I error control and operational interpretability, hybrid publications frequently clearly assess both sides, demonstrating enhanced detection power while suggesting practices such as simulation calibration, residual-charting, and conservative retraining schedules [11], [16]. When combined, your corpus demonstrates a practical approach: employ machine learning when conventional assumptions nonlinearity, high-dimensionality, non-stationarity, and unstructured data fail, but combine ML results with traditional SPC protections control-chart thresholds, residual monitoring, and conservative online updates to maintain alarms that are reliable and actionable [1], [4], [9].

Conclusions:

Integrating machine learning into statistical process control and associated monitoring frameworks significantly improves sensitivity, adaptability, and predictive capabilities when compared to conventional methods. Research indicates that ML is proficient in managing nonlinear, high-dimensional, autocorrelated, and noisy process data broadening SPC’s relevance to intricate industrial, environmental, biomedical, and cyber-physical systems. The applications include manufacturing quality control, renewable energy, food safety, environmental sustainability, healthcare diagnostics, remote sensing, predictive maintenance, and cybersecurity. Common strategies such as kernel methods, decision trees, neural networks, ensemble models, reinforcement learning, and hybrid statistical–ML control charts facilitate earlier fault detection, minimize false alarms, and enhance decision support. Nevertheless, challenges persist in ensuring interpretability, upholding statistical guarantees, controlling computational expenses, and addressing data constraints, particularly the lack of labeled fault data and the necessity for uncertainty quantification. Numerous studies emphasize the importance of domain adaptation, transparent model design, and benchmark datasets for equitable evaluation. Overall, the research highlights ML’s transformative potential in SPC, contingent upon future advancements that reconcile predictive performance with explainability, robustness, and practical deployability in real-time, safety-critical environments.

Reference:

[1] P. Qiu, "Machine Learning Approaches for Statistical Process Control," *Wiley StatsRef Stat. Ref. Online*, no. May, pp. 1–8, 2024, doi: 10.1002/9781118445112.stat08469.

[2] P. H. Tran, A. Ahmadi Nadi, T. H. Nguyen, K. D. Tran, and K. P. Tran, *Application of Machine Learning in Statistical Process Control Charts: A Survey and Perspective*. 2022. doi: 10.1007/978-3-030-83819-5_2.

[3] C. A. Escobar and R. Morales-Menendez, "Machine learning techniques for quality control in high conformance manufacturing environment," *Adv. Mech. Eng.*, vol. 10, no. 2, pp. 1–16, 2018, doi: 10.1177/1687814018755519.

[4] A. Apsemidis, S. Psarakis, and J. M. Moguerza, "A review of machine learning kernel methods in statistical process monitoring," *Comput. Ind. Eng.*, vol. 142, no. February, p. 106376, 2020, doi: 10.1016/j.cie.2020.106376.

[5] J. Y. Hsu, Y. F. Wang, K. C. Lin, M. Y. Chen, and J. H. Y. Hsu, "Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning," *IEEE Access*, vol. 8, pp. 23427–23439, 2020, doi: 10.1109/ACCESS.2020.2968615.

[6] S. H. Chou, S. Chang, T. R. Tsai, D. K. J. Lin, Y. Xia, and Y. S. Lin, "Implementation of statistical process control framework with machine learning on waveform profiles with no gold standard reference," *Comput. Ind. Eng.*, vol. 142, no. April 2019, p. 106325, 2020, doi: 10.1016/j.cie.2020.106325.

[7] Y. A. Atalan and A. Atalan, "Integration of the Machine Learning Algorithms and I-MR Statistical Process Control for Solar Energy," *Sustain.*, vol. 15, no. 18, 2023, doi: 10.3390/su151813782.

[8] Z. Ma and Y. Wang, "Optimizing Quality Control on Electric Vehicle Production Lines with AI and Machine Learning," *World J. Innov. Mod. Technol.*, vol. 7, no. 5, pp. 40–46, 2024, doi: 10.53469/wjimt.2024.07(05).04.

[9] P. Mascha, "VMLC: Statistical Process Control for Image Classification in Manufacturing," *Proc. Mach. Learn. Res.*, vol. 222, pp. 866–881, 2023.

[10] Z. J. Viharos and R. Jakab, "Reinforcement Learning for Statistical Process Control in Manufacturing," *Meas. J. Int. Meas. Confed.*, vol. 182, no. February, 2021, doi: 10.1016/j.measurement.2021.109616.

[11] H. Sabahno and A. Amiri, "New statistical and machine learning based control charts with variable parameters for monitoring generalized linear model profiles," *Comput. Ind. Eng.*, vol. 184, no. December 2022, p. 109562, 2023, doi: 10.1016/j.cie.2023.109562.

[12] P. Qiu and X. Xie, "Transparent Sequential Learning for Statistical Process Control of Serially Correlated Data," *Technometrics*, vol. 64, no. 4, pp. 487–501, 2022, doi:

10.1080/00401706.2021.1929493.

[13] T. Herzog, M. Brandt, A. Trinchi, A. Sola, and A. Molotnikov, "Process monitoring and machine learning for defect detection in laser-based metal additive manufacturing," *J. Intell. Manuf.*, vol. 35, no. 4, pp. 1407–1437, 2024, doi: 10.1007/s10845-023-02119-y.

[14] B. Wen, Z. Lu, and B. Zhou, "Marine Radar Constant False Alarm Rate Detection in Generalized Extreme Value Distribution Based on Space-Time Adaptive Filtering Clutter Statistical Analysis," *Remote Sens.*, vol. 16, no. 19, 2024, doi: 10.3390/rs16193691.

[15] J. Baehr *et al.*, "Predicting product life cycle environmental impacts with machine learning: Uncertainties and implications for future reporting requirements," *Sustain. Prod. Consum.*, vol. 52, no. August, pp. 511–526, 2024, doi: 10.1016/j.spc.2024.11.005.

[16] T. C. Schelhorn, J. Gunklach, and A. Maedche, "Designing an Analytical Control Chart System with ML-predicted Quality Characteristics," *CEUR Workshop Proc.*, vol. 3630, pp. 14–27, 2023.

[17] K. B. Newhart, C. A. Marks, T. Rauch-Williams, T. Y. Cath, and A. S. Hering, "Hybrid statistical-machine learning ammonia forecasting in continuous activated sludge treatment for improved process control," *J. Water Process Eng.*, vol. 37, no. April, 2020, doi: 10.1016/j.jwpe.2020.101389.

[18] S. I. Murphy, S. J. Reichler, N. H. Martin, K. J. Boor, and M. Wiedmann, "Machine learning and advanced statistical modeling can identify key quality management practices that affect postpasteurization contamination of fluid milk," *J. Food Prot.*, vol. 84, no. 9, pp. 1496–1511, 2021, doi: 10.4315/JFP-20-431.

[19] C. C. Chang and S. H. Chen, "Developing a Novel Machine Learning-Based Classification Scheme for Predicting SPCs in Breast Cancer Survivors," *Front. Genet.*, vol. 10, no. September, pp. 1–6, 2019, doi: 10.3389/fgene.2019.00848.

[20] G. Zamzmi *et al.*, "Out-of-Distribution Detection and Radiological Data Monitoring Using Statistical Process Control," *J. Imaging Informatics Med.*, vol. 38, no. 2, pp. 997–1015, 2025, doi: 10.1007/s10278-024-01212-9.

[21] Nimbale S. M. & Ghute V. B., "Artificial Neural Network Model for Monitoring the Fraction Nonconforming Control Chart," no. September, pp. 2013–2016, 2013, [Online]. Available: https://www.researchgate.net/publication/335601036_Licensed_Under_Creative_Commons_Attribution_CC_BY_Artificial_Neural_Network_Model_for_Monitoring_the_Fraction_Nonconforming_Control_Chart

[22] S. M. Nimbale and V. B. Ghute, "Monitoring Process Mean and Variability Using Artificial Neural Networks," *nt. J. Sci. Res. Math. Stat. Sci.*, vol. 6, no. August, pp. 1–6, 2019, doi: 10.26438/ijsrmss/v6i3.153158.

[23] S. M. Nimbale and V. B. Ghute, "A Neural Network based Individual Control Chart," *Int. J. Eng. Res.*, vol. V5, no. 05, 2016, doi: 10.17577/ijertv5is050137.

[24] Nimbale S. M. & Ghute V. B., "(PDF) Neural Network Based Non-Parametric Control Chart for Location," no. March, 2021, [Online]. Available https://www.researchgate.net/publication/359256367_Neural_Network_Based_Non-Parametric_Control_Chart_for_Location

[25] S. G. He, Z. He, and G. A. Wang, "Online monitoring and fault identification of mean shifts in bivariate processes using decision tree learning techniques," *J. Intell. Manuf.*, vol. 24, no. 1, pp. 25–34, 2013, doi: 10.1007/s10845-011-0533-5.

[26] F. Li *et al.*, "Ensemble Machine Learning Systems for the Estimation of Steel Quality Control," *Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018*, pp. 2245–2252, 2018, doi: 10.1109/BigData.2018.8622583.

[27] Z. Wu, J. Luo, D. Rincon, and P. D. Christofides, "Machine learning-based predictive control using noisy data: evaluating performance and robustness via a large-scale process simulator," *Chem. Eng. Res. Des.*, vol. 168, no. 310, pp. 275–287, 2021, doi: 10.1016/j.cherd.2021.02.011.

[28] Z. Wu, D. Rincon, Q. Gu, and P. D. Christofides, "Statistical machine learning in model predictive control of nonlinear processes," *Mathematics*, vol. 9, no. 16, pp. 1–37, 2021, doi: 10.3390/math9161912.

[29] G. A. Leon-Paredes *et al.*, "Presumptive Detection of Cyberbullying on Twitter through Natural Language Processing and Machine Learning in the Spanish Language," *IEEE Chil. Conf. Electr. Electron. Eng. Inf. Commun. Technol. CHILECON 2019*, pp. 1–7, 2019, doi: 10.1109/CHILECON47746.2019.8987684.

[30] L. Xu, Z. Chen, D. Zhu, and Y. Wang, "The Application Status of Radiomics-Based Machine Learning in Intrahepatic Cholangiocarcinoma: Systematic Review and Meta-Analysis," *J. Med. Internet Res.*, vol. 27, 2025, doi: 10.2196/69906.

[31] T. H. Tusher, A. Al Moneim, K. S. Ahmed, A. M. Billah, and S. Mangalathu, "Predicting shear strength of hollow pretensioned spun precast concrete pile using machine learning models," *Structures*, vol. 74, no. January, p. 108547, 2025, doi: 10.1016/j.istruc.2025.108547.

[32] Y.-D. Lin, Y.-H. Lu, R.-H. Hwang, Y.-C. Lai, D. Sudyana, and W.-B. Lee, "Evolving ML-Based Intrusion Detection: Cyber Threat Intelligence for Dynamic Model Updates," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 3, no. March, pp. 605–622, 2025, doi: 10.1109/tmlcn.2025.3564587.

[33] A. Burke, N. Snook, D. J. Gagne, S. McCorkle, and A. McGovern, "Calibration of machine learning–based probabilistic hail predictions for operational forecasting," *Weather Forecast.*, vol. 35, no. 1, pp. 149–168, 2020, doi: 10.1175/WAF-D-19-0105.1.