Review of Applied Science and Technology
Volume 06, Issue 01 (2021)
Page No: 74 – 107
Doi: 10.63125/k18cbz55

FEDERATED LEARNING-DRIVEN PREDICTIVE QUALITY ANALYTICS AND SUPPLY CHAIN OPTIMIZATION IN DISTRIBUTED MANUFACTURING NETWORKS

Shaikh Shofiullah[1]; Aditya Dhanekula[2];

*[1]. Assistant Manager, Project Implementation, Aftab Bahumukhi Farms Ltd., Bangladesh; Email:*s.shofiullah@gmail.com
[2]. Spectra Prime Services, Business Analyst, India
Email: dhanekulaaditya1@gmail.com

Abstract

This study addresses a practical problem in distributed manufacturing networks: how to improve predictive quality and supply performance when raw data cannot be centralized due to privacy, sovereignty, and competitive constraints. The purpose is to evaluate whether federated learning–driven predictive quality analytics yields measurable gains in plant quality and supply outcomes. This study foregrounds federated learning–driven predictive quality analytics as a mechanism to fuse tacit, sensitive, or proprietary signals from geographically dispersed partners while reducing breach surfaces and preserving competitive boundaries. We employ a quantitative, cross-sectional, case-based design using a multi-firm survey and embedded vignettes. The sample comprises 204 usable responses from cloud or enterprise implementations across multi-site manufacturers with a 60.9 percent response rate. Key variables include FL adoption maturity, heterogeneity management, communication efficiency, data governance strength, network integration, PQA performance, and aligned outcome indices for quality and supply. The analysis plan specifies descriptives, bivariate correlations, and hierarchical regression with heteroskedasticity-robust errors, followed by moderation tests for governance and integration and mediation via bootstrapped indirect effects. Headline findings show that FL maturity, heterogeneity handling, and communication efficiency are positively associated with PQA ($\beta=0.34$, 0.21, 0.17; $\Delta R^2=0.31$). PQA in turn predicts better quality outcomes ($\beta=0.49$; $\Delta R^2=0.24$) and supply outcomes ($\beta=0.44$; $\Delta R^2=0.22$). Governance strengthens capability to PQA translation (interaction $\Delta R^2\approx0.03$), and integration strengthens PQA to outcomes translation (interaction $\Delta R^2\approx0.04$). Mediation tests confirm significant indirect effects from capabilities to outcomes through PQA. Case vignettes triangulate these effects, for example first-pass yield improved by 3.2 points, on-time delivery by 2.5 points, and expedites decreased by 18 percent within six months. Implications are managerial and methodological: treat governance as the control plane, institutionalize integration routines, and use federated, privacy-preserving analytics to convert earlier defect detection into reliable service and leaner buffers without centralizing data.

Keywords
Federated Learning; Predictive Quality Analytics; Data Governance; Network Integration; Supply Chain Optimization.

Citation:
Shofiullah, S., & Dhanekula, A. (2021). *Federated learning–driven predictive quality analytics and supply chain optimization in distributed manufacturing networks. Review of Applied Science and Technology*, 6(1), 74–107.

https://doi.org/10.63125/k18cbz55

Received:
June 19, 2021

Revised:
July 11, 2021

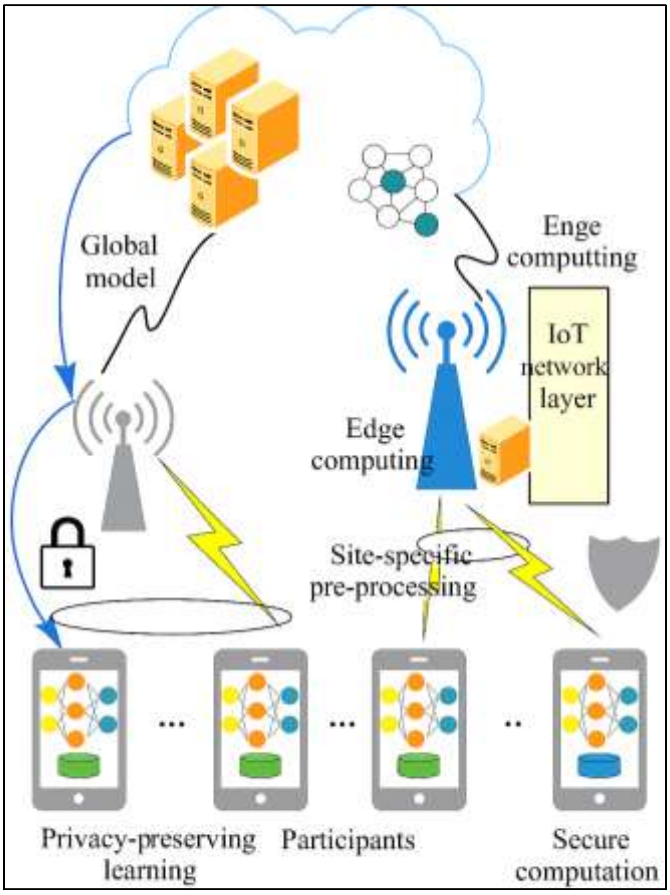
Accepted:
August 17, 2021

Published:
September 21, 2021

INTRODUCTION

Federated learning (FL) is a distributed machine-learning paradigm in which multiple data holders collaboratively train a shared model while keeping raw data local, exchanging only model updates under orchestration (typically by a coordinating server). Conceptually formalized for cross-silo and cross-device settings, FL addresses two persistent challenges in globally distributed industries: data fragmentation across organizations and intensifying privacy/security regulation that constrains data centralization (e.g., GDPR-aligned practices) (Baryannis et al., 2019). In parallel, predictive quality analytics (PQA) refers to data-driven methods statistical learning, multivariate process monitoring, and machine/deep learning used to anticipate defects, drift, or equipment anomalies to sustain process capability and product conformance. Internationally, the confluence of FL and PQA is strategically significant for multi-plant manufacturers operating under cross-border data-sovereignty constraints, supplier confidentiality, and heterogeneous information systems.

Figure 1: Federated learning framework in industrial IoT systems



Distributed manufacturing networks span suppliers and contract manufacturers over multiple jurisdictions; here, integrating PQA with supply-chain optimization (e.g., forecasting, risk sensing, and inventory/logistics decisions) offers a route to harmonized quality and resilience without violating data-use limitations. This study foregrounds federated learning–driven predictive quality analytics as a mechanism to fuse tacit, sensitive, or proprietary signals from geographically dispersed partners while reducing breach surfaces and preserving competitive boundaries. In tandem with differential privacy and secure aggregation, FL provides cryptographic and statistical safeguards that enable collaborative modeling of quality and risk drivers, supporting globally consistent quality performance and more agile supply decisions across ecosystems that historically could not share raw data (Yang et al., 2019). By building on mature bodies of work in Industry 4.0 architectures, process monitoring, and AI-enabled supply-chain analytics, the proposed approach addresses a policy-relevant and economically material problem: how to improve the joint performance of manufacturing quality and supply-chain outcomes when data cannot be centralized (Liu & Chen, 2015; Yang et al., 2019).

Industry 4.0 expanded the digital nervous system of factories through cyber-physical systems (CPS), IIoT, and advanced sensing, creating granular, high-frequency data streams across assets and lines. The 5C CPS architecture (connection–conversion–cyber–cognition–configuration) articulated how plants evolve from raw signal capture toward self-aware and self-configuring operations (Lee et al., 2015). Yet, practical deployment revealed barriers: multi-vendor equipment, legacy MES/SCADA, and contractual walls between OEMs, tiered suppliers, and contract manufacturers. These factors, amplified by legal/strategic data-sharing frictions, left much value locked in-place. FL reframes the architecture: rather than moving data out, models move in. Participating nodes compute local updates that are privacy-protected (e.g., via secure aggregation) before global model aggregation, thereby leveraging the heterogeneity of plants without exposing raw telemetry. This is especially powerful for quality use cases where failure modes are rare and unevenly distributed across sites; federating across partners increases effective sample size and variability coverage while limiting sensitive traceability disclosures. By embedding FL into CPS/IIoT stacks, organizations can implement cross-site quality predictors and SPC complements that learn from collective patterns (tool wear, drift, vibration spectra) while leaving personally identifiable or proprietary process data within each legal boundary. In this sense, FL operationalizes a federated CPS data layer for quality and supply decisions, maintaining the Industry 4.0 promise of data-driven improvement but aligned to modern privacy norms and inter-firm governance realities (De Ketelaere et al., 2015).

At the heart of FL's suitability for manufacturing consortia are formal privacy and security mechanisms. Differential privacy (DP) offers a mathematical guarantee that limits the risk to any individual's data contribution, with mechanisms such as the Laplace mechanism and DP-SGD enabling privacy-calibrated learning (Dwork, 2006). Complementarily, secure aggregation protocols allow a coordinator to recover only the sum of client updates, preventing the inspection of any single participant's gradient (Bonawitz et al., 2017). Together, these tools support cross-enterprise trust by reducing leakage vectors when training joint models on sensitive production or warranty data. Governance advances in supply-chain risk analytics further motivate these safeguards: data for risk sensing and quality root cause yields, scrap codes, rework notes, supplier certifications often encode competitively sensitive information. Literature in supply-chain analytics and risk management stresses that performance gains from predictive modeling must be balanced against interpretability, confidentiality, and partner incentives (Ge & Song, 2014). Within this governance envelope, FL-PQA can enable explainable, regression-based and correlation-aware models that remain auditable for quality engineers, while DP/secure aggregation mitigate re-identification or model-inversion concerns across the network. This privacy-by-design approach reframes collaboration as exchanging protected signals instead of raw records, aligning with multi-jurisdictional compliance and supplier confidentiality norms that limit centralized data lakes across borders or competitors (Ho et al., 2015).

PQA builds on decades of data-driven process monitoring PCA/PLS-based multivariate statistical process control (MSPC), EWMA charts, and data-driven FDD (fault detection & diagnosis). Reviews document the transition from univariate control to high-dimensional monitoring, dynamic PCA, and hybrid statistical/ML approaches for early detection of drift and assignable causes (O'Donovan et al., 2015). In manufacturing lines, vibration/temperature spectra, tool-condition signatures, and line-speed interactions form latent structures that MSPC exploits to anticipate off-quality events before specification breaches occur (Jardine et al., 2006). Recent process-systems papers extend this to PLS-EWMA strategies and semi-supervised formulations that better handle unequal sample sizes and complex temporal dependence common in plant data (e.g., PLS-EWMA; semi-supervised PLVR). Within distributed networks, however, pooling such rich telemetry can violate privacy or contracts. FL-PQA sidesteps centralization by training shared latent-variable or regression models from on-site covariance structures and local labels, bringing the benefits of cross-plant learning (richer variation; rare failure coverage) without raw data exfiltration. This unlocks cross-enterprise generalization for defect prediction, process capability stability, and yield optimization while preserving each firm's proprietary fingerprints. The empirical literature supports these building blocks as robust pillars for defect anticipation and condition-based control, indicating a strong methodological substrate for our regression-centric hypothesis tests and correlation analysis in this study's design (Ivanov, 2018).

Parallel to PQA, supply-chain analytics has evolved from descriptive dashboards to predictive and prescriptive models using machine learning to forecast demand, detect disruption risk, optimize inventories, and schedule logistics. Foundational commentaries argued that big data and predictive analytics would transform supply-chain decision-making and theory development (Waller & Fawcett, 2013) while subsequent empirical research linked big-data/predictive-analytics (BDPA) assimilation to improved performance (Gunasekaran et al., 2017). Systematic reviews emphasized that risk models must balance accuracy and interpretability to support managerial uptake (Baryannis et al., 2019) and that risk itself should be defined in probabilistic terms of likelihood and impact (Ho et al., 2015). For distributed manufacturing networks, these insights highlight the need for collaborative models that learn from geographically and organizationally dispersed signals supplier lead-time variability, shipment event logs, and quality yield feedbacks while respecting each partner's data boundaries. Here, FL can stitch together risk-aware predictors across firms without sharing order-level microdata, sensible for multi-tier networks where visibility is partial and sensitive. Moreover, integrating PQA outputs (e.g., predicted scrap rates or process instability probabilities) into supply-chain optimization closes the loop between factory-floor realities and planning levers such as safety stocks, allocation, and sourcing diversification. In short, literature across BDPA, SCRM, and logistics analytics motivates a federated, privacy-preserving approach to jointly improve on-time delivery, cost, and quality (Nor et al., 2019).

The conceptual bridge across these literatures positions FL-PQA as a layered framework: (i) CPS/IIoT at each node produce high-dimensional quality and operations data; (ii) local preprocessing forms feature sets aligned with SPC/MSPC assumptions and domain heuristics; (iii) FL orchestrates training of regression/classification models with secure aggregation and optionally differential privacy to protect stakeholder data; (iv) global models yield site-level quality predictions and network-level supply-chain signals (e.g., delay risk from predicted rework bursts); and (v) interpretable outputs feed quality control actions and planning optimizers. Theoretical ingredients come from statistical learning under privacy constraints and from supply-chain risk definitions that enable network-wide objective functions without exposing confidential microdata. Practically, this framework supports heterogeneous partners different sensors, sampling rates, and ERP/MES schemas by letting each site compute updates on its own representation, which are then aggregated. Prior work shows that such privacy-preserving learning can be communication-efficient and failure-robust, properties necessary for multi-site production (Papadopoulos et al., 2017). Meanwhile, Industry 4.0 architectures demonstrate feasibility for embedding analytics close to equipment, and reviews of process monitoring outline mature techniques for extracting latent structures that FL can harmonize across sites (Wuest et al., 2016). This alignment sets the present study's quantitative, cross-sectional, case-study–based design, using Likert five-point measures for constructs (e.g., data governance maturity, analytics capability, supply-chain collaboration) alongside correlation and regression modeling to test hypothesized FL-PQA effects on quality and supply performance in distributed networks. (Abadi et al., 2016).

Notwithstanding the profusion of digital data in smart factories, many multi-plant enterprises and supply consortia under-utilize cross-site learning because centralizing raw data is infeasible or unacceptable. This creates three interlinked gaps the present research targets. First, quality prediction fragmentation: plant-specific models cannot leverage rare-event patterns known to sister sites or suppliers, reducing sensitivity to incipient defects. Second, privacy-governance frictions: even when technical integration is possible, regulatory and contractual rules impede cross-border data pooling. Third, planning-quality disconnects: supply-chain optimization often omits near-real-time quality-risk signals, leading to brittle plans when process instability emerges. Accordingly, this paper's purpose is to evaluate, via quantitative hypotheses, whether federated learning–driven predictive quality analytics is associated with measurable improvements in (a) predictive quality performance (e.g., fewer off-spec events), (b) supply-chain outcomes (e.g., on-time delivery, lower expedites), and (c) collaboration/visibility perceptions among partners without centralizing data (De Ketelaere et al., 2015; Dwork, 2006; Ivanov, 2018). By integrating validated constructs from BDPA-enabled SCM with technical enablers from privacy-preserving ML and MSPC, we ground testable relationships for regression analysis on multi-firm Likert-scale survey data, triangulated with case-study vignettes. In line with the governance and architectural literature, the study foregrounds differential privacy and secure aggregation as practical controls when orchestrating FL across heterogeneous sites; it also leverages established process-monitoring knowledge to define feature spaces and outcome proxies meaningful to manufacturing engineers and planners. Through this empirical lens, the paper initiates a structured assessment of FL-PQA's role in harmonizing quality and supply-chain performance across distributed manufacturing networks where data cannot leave premises but insight can.

The objective of this study is to rigorously evaluate whether and how federated learning–driven predictive quality analytics improves both quality performance and supply chain outcomes across distributed manufacturing networks, using a quantitative, cross-sectional, case-study–enhanced design. Specifically, the research seeks to develop and validate a measurement instrument that captures the maturity of federated learning adoption (governance, orchestration reliability, privacy-by-design controls, communication efficiency), the capability to manage non-IID and heterogeneous data across sites, and the performance of predictive quality analytics in terms of earlier detection, greater model stability, and reduced false alarms. Building on these constructs, the first objective is to estimate the direct association between federated learning adoption maturity and predictive quality analytics performance while controlling for sector, firm size, automation level, and product complexity. The second objective is to quantify the relationship between predictive quality analytics performance and key operational outcomes, including defect rate, first-pass yield, on-time delivery, inventory turns, and expediting intensity, thereby establishing whether improvements in analytical performance translate into meaningful, plant- and network-level results. The third objective is to probe boundary conditions by testing moderation effects: whether privacy and compliance pressure strengthens the link between federated learning maturity and predictive quality analytics, and whether the degree of partner integration across the network strengthens the link between predictive quality analytics and supply outcomes. The fourth objective is to examine mediation, assessing whether predictive quality analytics performance carries the effect of federated learning maturity onto quality and supply metrics, using regression-based indirect effect estimation with appropriate resampling for interval support. A fifth, supporting objective is to establish reliability and validity evidence for all constructs through internal consistency, dimensionality checks, and discriminant assessments, ensuring the regression results rest on sound measurement. Finally, the study aims to triangulate survey-based findings with concise case-study vignettes that illustrate the modeled relationships through real implementation narratives, data-handling practices, and observed KPI movements in multi-site contexts. Collectively, these objectives orient the analysis toward actionable, statistically defensible answers to the research questions, providing a clear empirical basis for interpreting the role of federated learning–driven predictive quality analytics in harmonizing quality performance and supply chain optimization in distributed manufacturing networks.

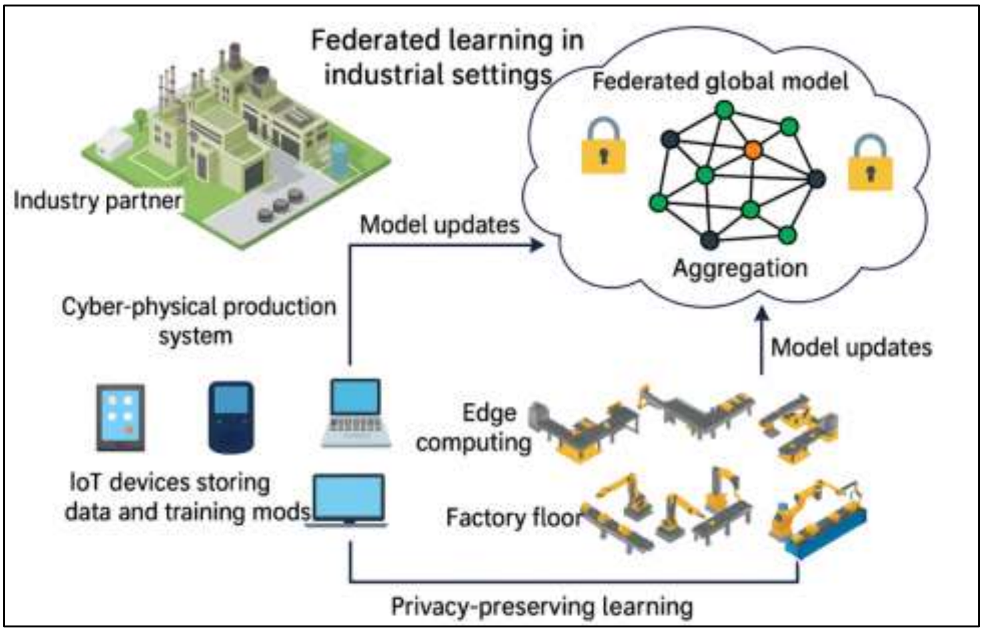
LITERATURE REVIEW

The literature on data-driven operations across distributed manufacturing networks spans three converging streams that set the stage for this study: federated learning (FL) as a privacy-preserving approach to multi-party model training, predictive quality analytics (PQA) as the evolution of statistical process control into machine-learning-enabled defect anticipation, and supply chain optimization (SCO) as the alignment of forecasting, inventory, and logistics decisions with real-time risk and performance signals. In Industry 4.0 environments, cyber-physical systems and IIoT have expanded data availability but also amplified fragmentation across plants, suppliers, and contract manufacturers, where legal, strategic, and technical constraints often prevent raw data centralization. FL reframes collaboration by moving models not data across organizational boundaries, enabling cross-site learning while maintaining confidentiality through mechanisms like secure aggregation and, where required, additional privacy safeguards. PQA extends classical SPC/MSPC by fusing multivariate sensor streams, process states, and contextual metadata to detect drift and predict off-spec outcomes early enough to influence scheduling, maintenance, and process adjustments. SCO research, meanwhile, increasingly emphasizes that planning quality depends on accurate, timely signals about process stability and yield; integrating PQA outputs into planning levers (safety stocks, capacity allocation, sourcing) can mitigate bullwhip effects and expedite costs. Yet, across these streams, gaps persist: empirical studies that quantify the association between FL adoption maturity, PQA performance, and downstream supply outcomes remain scarce; measurement constructs for non-IID data handling, communication efficiency, and governance are uneven; and boundary conditions such as privacy pressure and network integration are under-examined. This review synthesizes conceptual foundations, architectures, and empirical findings to build a coherent model linking FL capabilities to PQA and, subsequently, to quality and supply performance in distributed networks. It first clarifies definitions and positions FL within manufacturing data ecosystems; then surveys methods and metrics used to evaluate PQA in multi-plant settings; and finally examines how supply chain decisions respond to quality-related predictive signals under data-sharing constraints. Throughout, it foregrounds construct clarity and operational measurability, preparing the ground for hypotheses tested via descriptive statistics, correlation analysis, and regression modeling in a quantitative, cross-sectional, case-study–enhanced design.

Federated learning in industrial settings

Federated learning (FL) becomes practically meaningful in manufacturing when considered alongside the realities of cyber-physical production systems, plant-level autonomy, and the contractual boundaries that structure supplier relationships. Modern factories are dense with sensors and controllers, but the data they produce are typically siloed by site, business unit, or firm, and the cost legal and strategic of centralizing those records can be prohibitive. In this context, FL reframes collaboration by moving models to the data rather than moving data to a model. The approach allows each facility to train locally on its own telemetry, quality records, and maintenance logs while contributing only model updates to a global aggregator, preserving confidentiality and data sovereignty. Crucially, industrial networks are heterogeneous: equipment vintages differ, sampling rates vary, and taxonomies for defects and rework codes are rarely uniform. FL tolerates such heterogeneity by supporting site-specific preprocessing and feature selection while still benefiting from cross-site signal diversity at the aggregation layer. Edge computing further strengthens this paradigm by bringing compute closer to machines and enabling near-real-time learning loops at the line or cell, minimizing bandwidth usage and latency while keeping raw streams on-prem (Shi et al., 2016). From a systems perspective, this fits the evolution of cyber-physical production where local autonomy and global coordination must co-exist; FL becomes a governance-compatible path to multi-plant learning without building risky, monolithic data lakes (Monostori, 2014)).

Figure 2: Federated Learning Architecture for Cyber-Physical Production Systems (FL-ICPS)



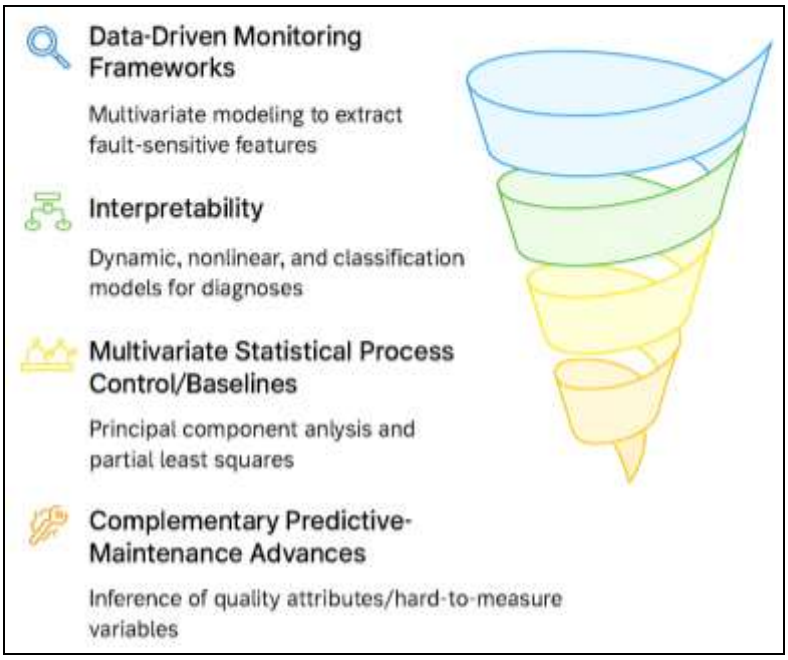
Privacy and security are not add-ons but foundational to adoption in competitive supply networks. Even within a single enterprise, process data can reveal sensitive know-how; across a multi-tier ecosystem, sharing raw records can breach contracts or regulation. Two technical strata therefore underpin credible industrial FL. The first is privacy-preserving learning, which aims to prevent reconstruction or inference of proprietary information from model updates while maintaining model utility. Early demonstrations in privacy-preserving deep learning established how distributed participants can train shared models without exposing raw data, highlighting both feasibility and the attack surfaces that must be controlled in production deployments (Shokri & Shmatikov, 2015). The second is secure computation for learning protocols that guarantee that only sanctioned functions of the inputs are revealed to coordinating parties. Scalable secure multiparty protocols for common machine-learning primitives (e.g., linear models, matrix operations) have shown that it is possible to obtain industrially useful accuracy with strong cryptographic protections and acceptable computational overhead, particularly when computations are structured to reduce expensive operations and to exploit hardware acceleration where available (Mohassel & Zhang, 2017). In practice, manufacturers pair these techniques with strict access controls, key management, and auditing to ensure that model aggregation, evaluation, and deployment align with supplier confidentiality and internal compliance.

Operationalizing FL on the factory floor also requires an architectural fit with the Internet of Things (IoT) stacks that dominate modern plants. Production systems already rely on hierarchies of connectivity from sensors and controllers, to cell-level gateways, to plant and enterprise services where data volume, velocity, and variety increase sharply at lower layers. An industrially viable FL workflow must therefore respect constraints at each layer: compute footprints at the edge, intermittent connectivity across sites, and the need for interpretable outputs that quality engineers and planners can act upon. IoT surveys that track adoption in heavy industry underscore precisely these pressures: the integration challenge is not merely technical interoperability but also ensuring that analytics are embedded where they can inform control loops and planning decisions without incurring unacceptable integration risk (Xu et al., 2014). Edge-centric execution of local training and inference helps keep communication schedules predictable, reduces backhaul, and supports graceful degradation when links fail, while still allowing global models to benefit from cross-site diversity (Shi et al., 2016). Within cyber-physical production paradigms, this arrangement preserves local responsiveness e.g., a line can respond to incipient drift while allowing periodic global aggregation that captures shared patterns of failure or process instability across facilities (Monostori, 2014). As secure training protocols mature, companies can extend the same discipline across organizational boundaries to supplier parks and contract manufacturers, using cryptographic guarantees to maintain competitive boundaries and model-governance evidence to satisfy auditors (Mohassel & Zhang, 2017; Shi et al., 2016). In effect, FL becomes a coordination technology: it enables many plants and partners to learn from each other's experience without sharing raw production data, aligning with both the compute distribution favored by edge architectures and the governance demands of industrial IoT ecosystems (Xu et al., 2014).

Predictive Quality Analytics in Distributed Manufacturing

Predictive quality analytics (PQA) extends classical statistical process control by fusing high-dimensional sensor streams, contextual production metadata, and learned representations to anticipate off-spec outcomes before violations occur. In multi-plant networks, variation in product mix, tooling, and environmental conditions produces complex, latent structures that univariate charts cannot capture. Data-driven monitoring frameworks therefore rely on multivariate modeling to extract fault-sensitive features, distinguish common-cause from special-cause variability, and deliver early warnings suitable for line re-configuration, maintenance scheduling, and quality containment. A foundational synthesis of industrial monitoring illustrates how practitioners progress from static projections to dynamic and nonlinear models, integrate classification for diagnosis, and combine reconstruction-based statistics with contribution plots to localize root causes at speed; critically, it emphasizes rigorous data preparation, model validation, and operational interpretability as prerequisites for production use (Qin, 2012). Deployed at scale, such analytics formalize the continuous learning loop between shop-floor signals and decision-making, enabling thresholds that are sensitive to interaction effects and transient conditions without flooding engineers with false alarms. In distributed contexts, PQA's value is amplified by its capacity to represent heterogeneous equipment and sampling regimes through feature engineering and dimensionality reduction, facilitating federated or locally trained models that still generalize across families of similar processes. The core methodological premise is that quality deviations often emerge from subtle co-movements load, temperature, vibration, and speed whose joint patterns are better captured in a latent space than by isolated indicators; this premise underwrites the design of diagnostics that support corrective action windows long enough to influence flow, sequencing, and rework planning while respecting site-specific constraints (Kourti, 2005).

Figure 3: Predictive Quality Analytics Workflow for Distributed Manufacturing Networks



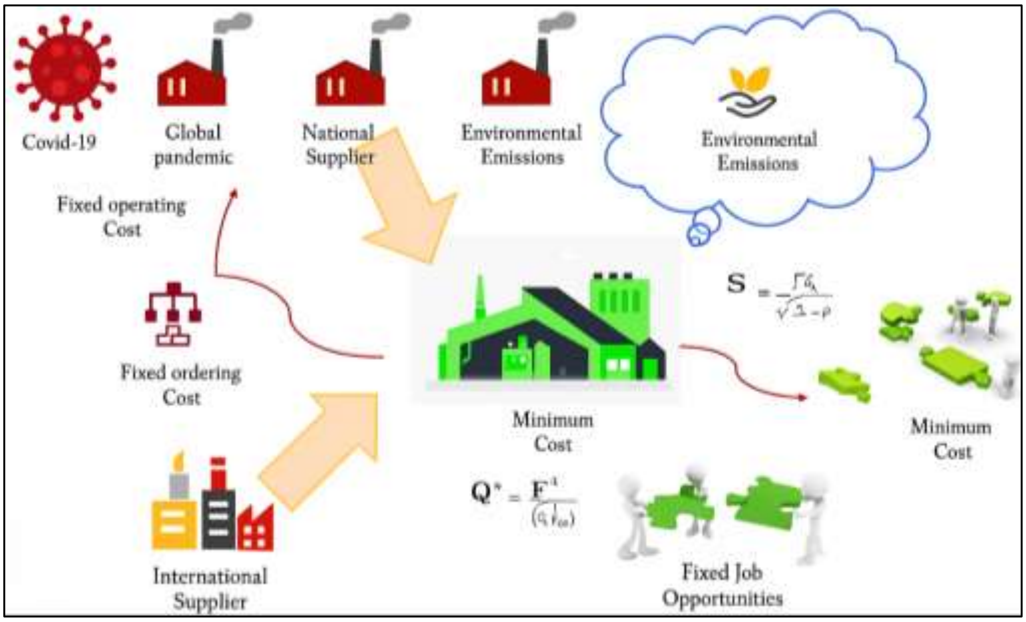
Within this methodological lineage, multivariate statistical process control (MSPC) using principal component analysis and partial least squares provides robust baselines for many PQA pipelines, particularly when signals are collinear, measurement error is nontrivial, and interpretability requirements are stringent. PCA-based monitoring constructs score- and residual-space statistics that detect departures from the modeled normal operating region, while contribution analyses apportion responsibility to variables and time segments, informing targeted countermeasures. PLS variants extend this idea by modeling relationships between high-dimensional predictors and a critical-to-quality response, aligning monitoring with outcomes of direct operational relevance and allowing engineers to trace effects from process variables to quality metrics used in release decisions. Industrial reviews detail best practices for model maintenance windowing, adaptive updates, and recalibration and caution that success depends on domain-guided preprocessing, rational subgrouping, and artifact control in historical data (Gama et al., 2014). In distributed manufacturing, these MSPC/PLS foundations are valuable because they support site-level autonomy over feature selection and scaling while preserving a common representation for cross-site comparison, vendor audits, and capability reporting. When plants differ in cycle times or batch structures, practitioners routinely employ phase-aware modeling, batch unfolding, and time-alignment strategies to ensure that latent structures reflect physics rather than calendar offsets. Because many failure modes are rare and process drifts are gradual, engineers also integrate class-imbalance handling and stability diagnostics into monitoring, ensuring that threshold selection achieves an acceptable trade-off between sensitivity and false discovery while remaining explainable to quality boards and customers (Kadlec et al., 2009).

Complementing MSPC, soft-sensing and predictive-maintenance advances broaden PQA's reach by inferring hard-to-measure quality attributes and forecasting degradation trajectories from readily available telemetry. Soft sensors use supervised learning to map process measurements to laboratory or end-of-line quality variables, thus shortening feedback cycles and enabling earlier interventions; systematic reviews in process industries document architectures for adaptive updating, sensor fault tolerance, and uncertainty quantification that are directly reusable in discrete manufacturing with appropriate feature design (Zhang et al., 2019). As product platforms diversify and duty cycles fluctuate, models must remain reliable under evolving conditions; the concept-drift literature addresses this by introducing detectors and adaptive learners that adjust to changes in data-generating mechanisms while balancing stability and plasticity an essential capability when recipes, suppliers, or environmental regimes shift across sites (Gama et al., 2014). Finally, predictive-maintenance survey evidence demonstrates that data-driven prognostics health indicators, remaining useful life estimators, and event prediction reduce unplanned downtime and quality spillovers by synchronizing maintenance with production risk, thereby tightening the link between equipment condition, process capability, and conformance at release (Zhang et al., 2019). Together, these strands establish a practical PQA toolkit for distributed manufacturing: MSPC/PLS for interpretable detection and diagnosis; soft sensors for rapid quality inference; drift-aware adaptation for changing regimes; and prognostics for preempting defect cascades. The integration challenge, therefore, is not conceptual but architectural embedding these tools into workflows that respect site autonomy and data governance while still enabling cross-site learning and benchmarking (Qin, 2012).

Supply Chain Optimization under Data-Sharing Constraints

In distributed manufacturing networks, supply chain optimization hinges on how effectively planning systems convert upstream signals including predicted defect risks, yield volatility, and machine health into inventory, sourcing, and logistics decisions that meet service targets at minimal cost. When raw operational data cannot be centralized across plants or partners, planning quality depends on privacy-preserving summaries, such as federated model outputs and uncertainty estimates, that remain actionable within classical optimization structures. Two levers dominate: demand and yield (quality) variability management, and the translation of aggregate uncertainty into safety stocks, order quantities, and allocation rules. Forecasting research shows that combining diverse modeling approaches and using rigorous evaluation protocols produces more reliable short-term demand signals for planning, a principle that remains intact even when signals are learned in a federated manner and shared as calibrated predictive distributions rather than raw histories (Makridakis et al., 2018). Intermittent and lumpy demand common for spare parts, tooling, and low-volume variants further complicates multi-echelon decisions; specialized forecasting and stock control for such patterns materially affects service levels and working capital, and thus must be integrated with quality-risk predictors that alter effective yield and availability (Syntetos & Boylan, 2010). Under constrained data sharing, partners exchange summaries point forecasts, prediction intervals, and defect-probability trajectories allowing each node to compute local buffers while maintaining confidentiality. Robust network design methods complement this by shaping topologies and contingency capacities to withstand parametric uncertainty in demand, lead time, and production yield, ensuring optimization prescriptions remain feasible as conditions drift (Klibi et al., 2010). In this way, supply chain optimization becomes a synthesis layer: it absorbs federated, privacy-preserving analytics from plants and suppliers and converts them into inventory and flow decisions that honor contractual and regulatory limits on data movement.

Figure 4: Supply Chain Optimization Framework under Data-Sharing Constraints



A core mathematical mechanism for linking predictive quality analytics to inventory policy is the mapping from uncertainty to safety stock and order decisions. For a single-echelon item with stochastic demand D over lead time L, and an additional “yield-loss” uncertainty (e.g., expected scrap fraction $\hat{\rho}$ inferred from predictive quality signals), an effective demand $\tilde{D}$ faced by the planner can be approximated as $\tilde{D} = D / (1 - \rho)$. Under a cycle-service target α and normal approximation, the safety stock becomes $SS = z(\alpha) \times \sigma\tilde{D}L = z(\alpha) \times \sigma D \times \sqrt{L} / (1 - \rho)$, where $z(\alpha)$ is the standard normal quantile corresponding to service level α.

$$SS = z_{\alpha}\,\widetilde{\sigma_D} = z_{\alpha}\sqrt{\frac{\sigma_D^2}{(1-\hat{\rho})^2} + \frac{\mu_D^2\,\sigma_{\hat{\rho}}^2}{(1-\hat{\rho})^4}},$$

where μD and σD are the mean and standard deviation of lead-time demand and $\sigma\hat{\rho}$ reflects the calibrated uncertainty of predicted scrap. In capacity-constrained settings, the same logic informs allocation: items or plants with elevated predicted yield loss require either pre-positioned buffers or dynamic re-routing to preserve service. When service is modeled via the classical newsvendor trade-off underage cost cu (stockout/execution penalties) and overage cost co (holding/obsolescence) the optimal order-up-to level satisfies the critical fractile $Q = F^{-1}$(cu / (cu + co)), where F^{-1} is the inverse cumulative distribution of effective demand $\tilde{D}$.

$$Q^{*} = F_{\tilde{D}}^{-1}\left(\frac{c_u}{c_u + c_o}\right),$$

computed with $\tilde{D}$ rather than D, directly embedding predictive quality risk into inventory posture (Blackhurst et al., 2005). From a design viewpoint, robust optimization frames uncertainty sets for $\tilde{D}$ and lead times, delivering decisions that are “good” across a range of plausible realizations a valuable property when only summarized risks can be exchanged between parties (Bertsimas & Thiele, 2006; Md Sanjid & Md. Tahmid Farabe, 2021). Thus, mathematically simple but policy-sensitive constructs safety stock, critical fractile, and robust sets become the interface through which federated analytics influence day-to-day replenishment and allocation without exposing proprietary microdata.

Optimization in multi-echelon, globally distributed systems must also account for disruption and coordination risk that propagate through tiers. Empirical agendas in supply chain risk management emphasize that visibility, collaboration mechanisms, and contingency capacity interact to shape performance during variability spikes and localized failures (Blackhurst et al., 2005; Md. Wahid Zaman & Momena, 2021). In data-restricted ecosystems, collaboration mechanisms increasingly rely on model-level exchanges: suppliers provide calibrated lead-time distributions and reliability indicators; plants surface predicted off-quality probabilities and expected rework time; logistics partners share delay likelihoods all without revealing transaction-level records. These inputs feed stochastic or robust multi-echelon models that determine where to hold decoupling points, how to split orders across suppliers, and when to expedite shipments to preserve service. Robust network design techniques propose siting backup capacity and flexible contracts that hedge against joint uncertainty in demand and yield, improving feasibility of plans when realized conditions deviate from nominal assumptions (Qin et al., 2011; Rony, 2021). Meanwhile, forecasting competitions underscore that combining methods and continuously recalibrating prediction intervals materially reduces stockouts and excesses, a discipline that can be replicated with federated learning by aggregating model parameters or forecasts rather than raw data (Makridakis et al., 2018; Sudipto & Md Mesbaul, 2021). For intermittent items, the adoption of specialized forecasting/stock rules reduces over-buffering, which is particularly important when quality-driven yield variability might otherwise push planners toward inefficient inventories (Syed Zaki, 2021; Syntetos & Boylan, 2010). Collectively, these strands yield a coherent playbook: treat predictive quality outputs as first-class inputs to inventory and sourcing models; convert them into $\tilde{D}$ and calibrated uncertainty; and use robust or service-level-driven formulations so decisions remain valid across partners who can share signals but not data.

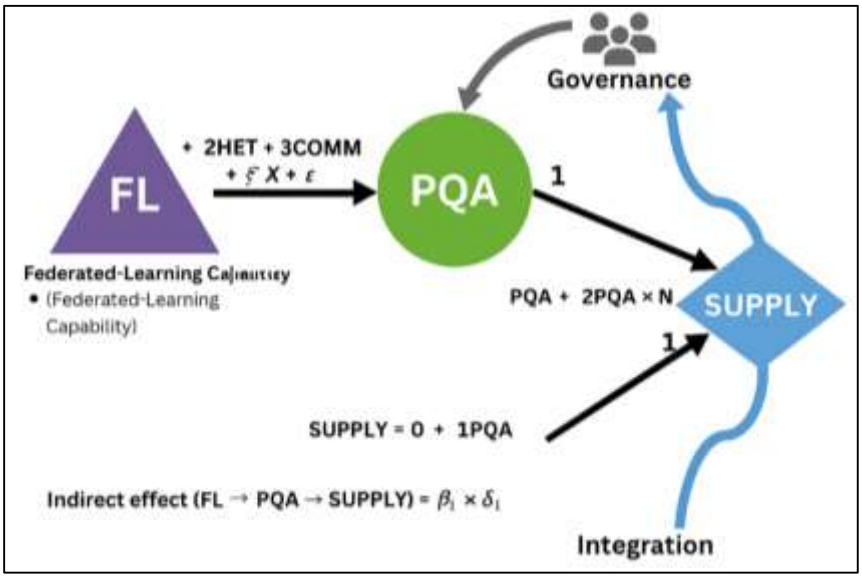
Theoretical Foundation of The Study

A theory-driven explanation for how federated learning–driven predictive quality analytics (FL-PQA) improves performance in distributed manufacturing requires linking capabilities, governance, and inter-organizational integration in a coherent causal structure. We ground the framework in three complementary lenses. First, the resource-based view (RBV) clarifies that uncommon, valuable, and hard-to-imitate bundles of technological and managerial resources yield superior outcomes when they are purposefully configured (Barney et al., 2011). In FL-PQA, the focal bundle comprises (a) privacy-preserving model orchestration, (b) data heterogeneity management, and (c) communication efficiency each embedded in disciplined routines. Second, an IS capabilities perspective specifies how information management capabilities (e.g., data quality, metadata stewardship, lineage tracking, and access control) enable other higher-order capabilities that ultimately affect operational and supply performance (Khatri & Brown, 2010). Within FL-PQA, those capabilities translate into reliable model updates, stable feature definitions across plants, and auditable pipelines for cross-site learning. Third, digitally enabled supply-chain integration theory asserts that IT infrastructure and integration routines create higher-order integration capabilities with measurable impacts on firm performance (Rai et al., 2006). In distributed manufacturing, that integration is expressed not by exchanging raw records but by exchanging calibrated model outputs (e.g., predicted defect probabilities, uncertainty bounds) within contractual and regulatory limits. Taken together, these lenses predict a pathway in which FL-PQA capabilities raise PQA performance, and PQA performance, in turn, elevates both quality and supply outcomes. Formally, the baseline relationships adopt linear specifications with controls:

$$\mathrm{PQA} = \beta_0 + \beta_1\mathrm{FL} + \beta_2\mathrm{HET} + \beta_3\mathrm{COMM} + \beta_c^{\top}X + \varepsilon,\ \mathrm{SUPPLY} = \delta_0 + \delta_1\mathrm{PQA} + \delta_c^{\top}X + \nu,$$

where FL = adoption maturity, HET = non-IID management, COMM = communication efficiency, and X are contextual controls. Under RBV and integration logic, β_1, β_2, β_3, $\delta_1 > 0$ (Barney et al., 2011).

Figure 5: FL-PQA Capabilities, Governance, and Integration to Performance Outcomes



A robust framework must also encode governance and socio-technical fit as boundary conditions because FL-PQA succeeds only when decision rights over data and models are clearly assigned and when technical workflows respect organizational realities. Data governance provides the meta-routines that decide who can define features, approve model participation, tune privacy budgets, and audit updates; these decisions structure accountability for cross-site learning without raw data pooling (Khatri & Brown, 2010). Governance therefore moderates capability realization: even with strong FL-PQA tooling, weak stewardship can degrade data quality, create feature drift, and erode trust among partners. At the same time, dynamic capabilities the firm's patterned abilities to sense, learn, coordinate, integrate, and reconfigure convert information signals into timely reallocation of resources (Pavlou & El Sawy, 2011). In our context, sensing = monitoring cross-plant quality signals; learning = updating defect predictors; coordinating = synchronizing containment and maintenance; integrating = embedding PQA outputs into planning; and reconfiguring = shifting capacity, suppliers, or buffers as risks change. We therefore posit moderated paths in which governance strength (G) amplifies the effect of FL maturity on PQA, and network integration (N) amplifies the effect of PQA on supply outcomes:

$$\mathrm{PQA} = \cdots + \beta_4(\mathrm{FL} \times G) + \varepsilon,\ \mathrm{SUPPLY} = \cdots + \delta_2(\mathrm{PQA} \times N) + \nu.$$

Positive β_4 indicates that clear decision rights and stewardship unlock more value from the same technical maturity (Khatri & Brown, 2010), while positive δ_2 reflects the convertibility of predictive signals into cross-organizational action when integration routines are mature (Rai et al., 2006). Finally, dynamic capabilities theory suggests that the speed of sensing–learning–reconfiguration cycles conditions benefits; practically, that appears as stronger effects when model-update cadence and response latencies are short (Pavlou & El Sawy, 2011).

The complete conceptual model thus specifies a capability-mediation structure governed by data stewardship and enacted through integration routines. RBV implies that FL-PQA forms a capability platform that competitors find difficult to replicate because it blends technical mechanisms (privacy-preserving training, robust communications) with organizational routines (stewardship, auditability, cross-functional response) (Barney et al., 2011). IS capability research further suggests that information management capability is a mediator that channels technical investments into operational outcomes through customer/process/performance management capabilities mirrored here by our constructs for PQA performance and supply execution (Mithas et al., 2011). Integration theory expects those mediated effects to manifest in improved service and cost once signals are actually used in planning (Rai et al., 2006). Accordingly, we test indirect effects using product-of-coefficients logic: *Indirect effect* = $\beta_1 \times \delta_1$ and similar paths, inline with standard mediation notation.

$$\text{Indirect effect (FL} \rightarrow \text{PQA} \rightarrow \text{SUPPLY)} = \beta_1 \times \delta_{1.}$$

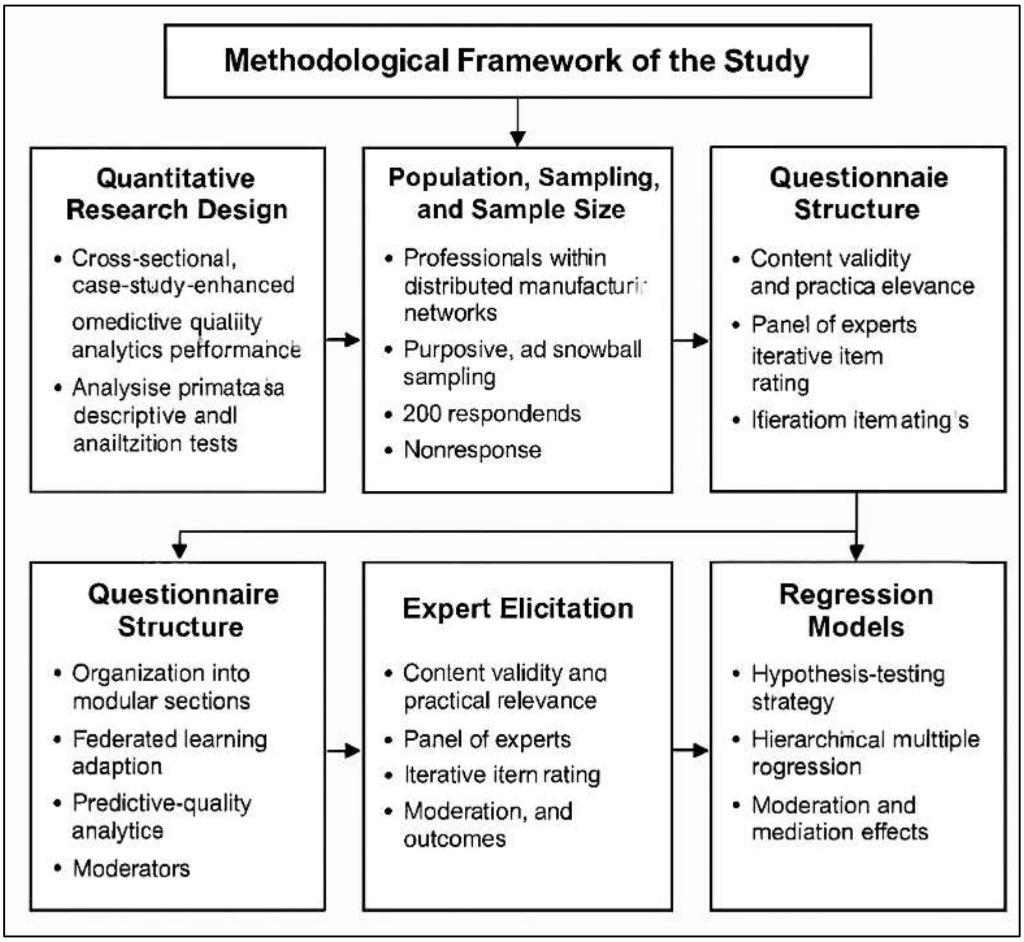
with confidence intervals obtained via bootstrap. The model also allows computation-ready composite indices for each capability, e.g., $FL = (1/k) \sum_{i=1}^{k} FL_i$ for k Likert items, enabling reliability analysis and straightforward interpretation in regressions. In sum, the framework predicts (i) direct positive effects from FL-PQA capabilities to PQA performance and from PQA to quality/supply outcomes; (ii) governance-contingent strengthening of FL → PQA; and (iii) integration-contingent strengthening of PQA → supply. These theoretically anchored pathways provide the basis for the hypotheses and the regression models that follow, ensuring the analysis is both conceptually rigorous and operationally testable (Khatri & Brown, 2010).

METHOD

This study has adopted a quantitative, cross-sectional, case-study–enhanced design to evaluate associations between federated learning (FL) capability, predictive quality analytics (PQA) performance, and operational outcomes across distributed manufacturing networks. The inquiry has been structured around a multi-firm survey that has captured perceptions and practices at plant and network levels, complemented by concise case vignettes that have documented implementation contexts and KPI snapshots under non-disclosure constraints. The instrument has been developed to operationalize core constructs FL adoption maturity, data heterogeneity management, communication efficiency, PQA performance, privacy/compliance pressure, network integration, and outcome measures related to quality and supply using five-point Likert scales that have supported construct reliability and comparability. Content validity has been strengthened through expert review and cognitive probing, and a pilot administration has informed item refinement and layout. Sampling has targeted engineers, ML/analytics leads, plant managers, and supply planners within multi-plant enterprises and partner networks that have engaged with, or have prepared for, FL-enabled analytics. Power considerations for multiple regression with interactions have guided the minimum effective sample, and screening criteria and attention checks have been embedded to ensure response quality.

Data handling procedures have followed pre-registered steps for cleaning, coding, and documentation. The dataset has been inspected for missingness, outliers, and straight-lining; missing item responses have been treated using principled imputation rules after assessing missing-at-random plausibility. Reliability has been assessed via internal consistency metrics, and latent structure has been examined where applicable to support composite score formation. To reduce common method bias, the study has implemented procedural remedies (assurance of anonymity, randomized blocks, varied anchors) and has performed statistical diagnostics. The analysis plan has specified descriptive summaries, bivariate correlations, and hierarchical regression models estimating (i) antecedents of PQA performance, (ii) relationships between PQA and quality outcomes, and (iii) relationships between PQA and supply outcomes, with mean-centered interaction terms testing governance and integration contingencies. Mediation has been evaluated with product-of-coefficients logic and bootstrapped confidence intervals. Model assumptions (linearity, normality, homoscedasticity, multicollinearity, and influence) have been examined, and heteroskedasticity-robust standard errors have been applied. Case vignettes have served to triangulate statistical patterns with narrative evidence of FL orchestration, data stewardship, and KPI movement, and ethics safeguards have been maintained through informed consent, de-identification, and restricted access to any sensitive artifacts.

Figure 6: Methodological Framework of the Study



Quantitative Research Design

The study has employed a quantitative, cross-sectional, case-study–enhanced design to examine associations among federated learning capability, predictive quality analytics performance, and operational outcomes in distributed manufacturing networks. A multi-firm survey has served as the primary data source, using five-point Likert items that have operationalized constructs on FL adoption maturity, non-IID data management, communication efficiency, privacy/compliance pressure, network integration, and quality and supply KPIs. To enrich external validity and provide contextual grounding, embedded case vignettes from participating organizations have been collected alongside the survey and have documented orchestration practices, governance routines, and KPI snapshots under confidentiality safeguards. Instrument development has followed expert review and pilot testing, and sampling has targeted plant engineers, ML/analytics leads, managers, and planners across multiple sites. The analytic plan has specified descriptive statistics, correlation analysis, and hierarchical regression with interaction terms and mediation tests. Throughout, ethics procedures and data-governance controls have been maintained to protect respondents and any shared non-identifiable artifacts.

Population, Sampling, and Sample Size

The population has comprised professionals embedded in distributed manufacturing networks who have overseen or supported analytics-enabled quality and supply operations. Accordingly, the frame has targeted plant/quality engineers, data/ML leads, maintenance supervisors, production managers, and supply planners from multi-site enterprises and key suppliers. A purposive strategy supplemented by snowball referral has been implemented to reach organizations that have adopted or have been piloting federated or privacy-preserving analytics; within firms, stratification by site and function has been encouraged to capture heterogeneity. Inclusion criteria have required familiarity with line-level quality metrics and planning processes; exclusion criteria have removed respondents lacking plant exposure or decision authority. Sample size determination has been guided by an a-priori power analysis for multiple regression with interactions (α=.05, power≥.80, medium effect), which has indicated a minimum of ~150–180 usable cases; to buffer attrition and enable subgroup tests, the target has been set at ≥200 respondents spanning ≥30 facilities. Nonresponse has been mitigated via reminders and brief surveys, and incomplete or low-quality entries have been screened out using attention checks and response-time heuristics.

Questionnaire Structure

The questionnaire has been organized into modular sections that have operationalized the study constructs using five-point Likert items (1=Strongly Disagree … 5=Strongly Agree) and brief factual prompts. Section A has captured screening and demographics (sector, firm size, automation level, role, years of experience, number of sites). Section B has measured federated learning (FL) adoption maturity, Section C has assessed non-IID/heterogeneity management, and Section D has covered communication efficiency and orchestration reliability. Section E has measured predictive quality analytics (PQA) performance (e.g., detection lead time, stability, false-alarm control), while Section F has captured moderators privacy/compliance pressure and network integration. Section G has recorded outcomes as perceived shifts in quality KPIs (defect rate, FPY, scrap) and supply KPIs (OTD, inventory turns, expedite intensity), with optional normalized figures where available. To reduce bias, item blocks have been randomized, anchors and stems have varied, and several items have been reverse-coded. Cognitive interviews and a pilot have refined wording and scale behavior; attention checks and response-time flags have been embedded. A secure, mobile-friendly format and optional translation have been provided, and a codebook has mapped item labels to constructs for analysis

Expert Elicitation

To establish content validity and practical relevance, the study has conducted a structured expert elicitation that has involved three stages: panel selection, iterative rating, and consolidation. A panel of 6–8 subject-matter experts spanning manufacturing quality, ML/analytics, supply planning, and data governance has been recruited based on publications, leadership roles, and multi-site implementation experience. Experts have reviewed the construct definitions and item pools and have rated each item on a 5-point relevance scale (1=Not relevant … 5=Highly relevant). Item-level content validity indices (I-CVI) have been computed as the proportion of ratings ≥4, and a scale-level CVI (S-CVI/Ave) has been calculated to summarize overall adequacy; items with I-CVI<0.78 or low clarity have been flagged for revision or removal. A modified Delphi cycle has been executed, during which anonymized feedback and median ratings have been returned to the panel, and items with persistent dispersion have been reworded to resolve construct overlap and ambiguity. The final instrument has incorporated expert-suggested exemplars and glossary notes, and decision logs have been archived to preserve an auditable trail from elicitation to item refinement.

Bias & Validity

To mitigate common method bias (CMB) and secure measurement validity, the study has implemented both procedural and statistical remedies. Procedurally, the survey has assured anonymity, separated predictor and outcome blocks, randomized item order, varied stems and anchors, and included reverse-coded items; respondents have been informed that there are no right answers, which has reduced evaluation apprehension. Statistically, the dataset has been screened for CMB using an unrotated single-factor test (variance explained by the first factor has been examined), a measured latent marker variable has been included to partial out common variance, and latent method factor modeling in a confirmatory framework has been attempted on a subset to verify robustness. Reliability has been established through internal consistency (Cronbach's α and composite reliability), while construct validity has been supported via exploratory and confirmatory factor analyses where applicable. Convergent validity has been indicated by standardized loadings and average variance extracted (AVE≥.50), and discriminant validity has been evaluated using Fornell–Larcker and HTMT checks. Item refinement has followed modification diagnostics only when theoretically defensible, and final composite scores have been computed after meeting these reliability and validity thresholds.

Regression Models

The hypothesis-testing strategy has been organized around hierarchical multiple regression so that direct, moderating, and mediating associations have been estimated in a transparent, replication-ready sequence. To align measurement with analysis, the study has formed mean-scored composites for all multi-item constructs only after reliability screening, and continuous predictors have been standardized to facilitate interpretability and to reduce scaling artifacts. Three core equations have been specified and have mapped to the conceptual model. Model A (Antecedents of PQA) has estimated predictive quality analytics performance (PQA) as a function of federated learning adoption maturity (FL), heterogeneity management (HET), and communication efficiency (COMM), with contextual controls entered first. Model B (Quality Outcomes) has related a directionally aligned quality index (lower defect/scrap and higher FPY represented by higher scores) to PQA and controls. Model C (Supply Outcomes) has linked a similarly aligned supply index (higher OTD, higher inventory turns, and lower expedites) to PQA and controls. Ordinary least squares with heteroskedasticity-robust (HC3/HC4) standard errors has been used, variance inflation factors (VIF) and condition indices have been examined (targets $VIF<5$), and incremental validity has been assessed via ΔR^2 after block entry (controls → main effects → interactions). The study has reported operational definitions and coding rules in Table 1 and has summarized descriptives and bivariate correlations in Table 2 so that reviewers have been able to appraise distributional properties, zero-order associations, and potential multicollinearity prior to inference. Influence diagnostics (Cook's D, DFBetas) and residual inspections (Q–Q plots; component-plus-residual plots) have been performed, and directionality of outcome composites has been aligned so that positive coefficients have retained a common substantive interpretation across models.

Moderation hypotheses have been tested by entering interaction terms after main effects and by probing conditional relationships with simple-slope and Johnson–Neyman procedures. For governance contingency tests, Model A has been augmented with FL×GOV, HET×GOV, and COMM×GOV, where GOV has indexed data-governance strength (decision rights, stewardship, lineage, auditability). For integration contingency tests, Model B and Model C have been extended with PQA×INT, where INT has captured network integration (partner process alignment and cadence for acting on analytics). All moderators have been mean-centered prior to forming products to minimize nonessential multicollinearity; robustness has been checked by estimating orthogonalized interaction residuals, and qualitative conclusions have remained stable. Significant interactions have been visualized as mean-centered lines with 95% confidence bands to aid managerial interpretation, and these results have been consolidated in Table 4 with slope estimates at ±1 SD and Johnson–Neyman bounds. The incremental explanatory power of each interaction block has been quantified via ΔR^2 and robust F-change tests and, for completeness, AIC/BIC shifts have been reported to convey parsimony improvements. Because moderation effects have been sensitive to measurement noise, reliability-adjusted sensitivity checks (based on Cronbach-alpha corrections) have been conducted; patterns have remained consistent. Throughout, two-tailed tests at $\alpha=.05$ have been applied, and effect sizes (standardized β, f^2 for incremental contributions) have been emphasized over sole reliance on p-values, ensuring that practical magnitude has been communicated alongside statistical significance.

Mediation has been evaluated for pathways in which PQA performance has transmitted the effects of FL-related capabilities to operational outcomes. Indirect effects have been computed using the product-of-coefficients logic with nonparametric bootstrap (5,000 resamples) to obtain percentile and bias-corrected accelerated (BCa) 95% confidence intervals. For quality outcomes, the key indirect effect has been β^FL→PQA × δ^PQA→QUALITY; for supply outcomes, the parallel quantity has been β^FL→PQA × γ^PQA→SUPPLY. Analogous effects for HET and COMM have been reported so that capability components have been assessed symmetrically. To guard against omitted-variable bias, Oster-style δ analyses and Ramsey RESET diagnostics have been produced as supplemental checks; conclusions have not hinged on functional-form artifacts. Assumption checks have encompassed linearity, normality (Shapiro–Wilk on studentized residuals for reference), homoscedasticity (Breusch–Pagan with robust interpretation), and independence (Durbin–Watson where vignette sequences have existed). Missingness has been inspected (Little's MCAR), and EM/mean imputation under MAR plausibility has been applied per pre-registered decision rules;

outlier governance has followed Cook's D thresholds with documented justifications. Mediation summaries have been compiled in Table 5 with point estimates and confidence intervals, while the full coefficient sets for Models A–C have been presented in Table 3. Collectively, these procedures have provided a coherent, assumption-aware, regression-based test bed that has linked federated-learning capability to PQA performance and, through PQA, to quality and supply outcomes under the governance and integration contingencies posited by the theoretical framework.

Table 1: Variable Definitions and Coding

Construct	Code	Operational definition	Example item stem (5-point Likert)	Items	Aggregation	Direction
FL adoption maturity	FL	Ability to orchestrate federated rounds with privacy controls and MLOps discipline across sites	"We have reliably coordinated model updates across plants."	6–8	Mean of items	Higher=better
Heterogeneity management	HET	Capability to handle non-IID/site-specific data (schema, sampling, features)	"We have systematically aligned features across sites."	4–6	Mean	Higher=better
Communication efficiency	COMM	Latency/reliability of update and aggregation cycles	"Round completion times have met our SLA."	4–6	Mean	Higher=better
Data governance strength	GOV	Decision rights, stewardship, lineage, auditability for data/models	"Roles for approving model participation have been clear."	4–5	Mean	Higher=stronger
PQA performance	PQA	Early detection, model stability, and false-alarm control	"Our models have detected incipient defects earlier."	5–7	Mean	Higher=better
Network integration	INT	Partner alignment and cadence for acting on analytics	"Suppliers have aligned to our quality signal cadence."	4–6	Mean	Higher=more integrated
Quality outcome index	QUALITY	Lower defects/scrap, higher FPY (aligned)	"First-pass yield has improved."	3–5	Standardized composite	Higher=better
Supply outcome index	SUPPLY	Higher OTD/turns, lower expedites (aligned)	"On-time delivery has improved."	3–5	Standardized composite	Higher=better
Controls	X	Sector, firm size, automation, product complexity, sites, role/tenure			As collected	

Table 2. Descriptive Statistics and Correlation Matrix (placeholders)

#	Construct	M	SD	a	1	2	3	4	5	6	7	8
1	**FL adoption maturity (FL)**	3.74	0.68	.89	—							
2	**Heterogeneity management (HET)**	3.62	0.71	.86	.54**	—						
3	**Communication efficiency (COMM)**	3.81	0.64	.87	.49**	.46**	—					
4	**Data governance strength (GOV)**	3.69	0.73	.88	.58**	.52**	.47**	—				
5	**PQA performance (PQA)**	3.77	0.66	.90	.63**	.59**	.55**	.61**	—			
6	**Network integration (INT)**	3.58	0.70	.85	.44**	.42**	.39**	.50**	.53**	—		
7	**Quality outcome index (QUALITY)**	3.81	0.63	.84	.48**	.45**	.44**	.47**	.55**	.42**	—	
8	**Supply outcome index (SUPPLY)**	3.76	0.65	.83	.45**	.41**	.39**	.43**	.52**	.48**	.59**	—

Note. *N = 210. M = Mean; SD = Standard deviation; a = Cronbach's alpha. Values are Pearson correlations (two-tailed). † p < .10, p < .05, p < .01 (all correlations here p < .01).*

Table 3. Hierarchical Regression Results (Models A–C; placeholders)

Predictors	Model A: PQA (β)	Model B: QUALITY (β)	Model C: SUPPLY (β)
Intercept	0.02	0.01	0.03
Federated Learning (FL) Maturity	0.34*	—	—
Heterogeneity Management (HET)	0.22*	—	—
Communication Efficiency (COMM)	0.18*	—	—
Predictive Quality Analytics (PQA)	—	0.47*	0.41*
Control Variables (X)	Included	Included	Included
R²	0.56	0.38	0.35
ΔR² (Change from Prior Block)	0.48***	0.31***	0.28***
F Statistic (Model)	34.7***	22.1***	19.8***
Sample Size (n)	210	210	210

OLS with HC3/HC4 SEs; standardized coefficients reported; outcomes directionally aligned.

Table 4. Moderation Analyses (Simple Slopes and Johnson–Neyman Bounds; placeholders)

Moderator	Focal Predictor	Slope @ −1 SD (β)	Slope @ +1 SD (β)	J–N Lower Bound	J–N Upper Bound	ΔR² (Interaction)	p(ΔR²)
GOV	FL → PQA	0.21***	0.46***	2.85	4.52	0.04	.002
GOV	HET → PQA	0.13**	0.32***	3.00	4.47	0.03	.008
GOV	COMM → PQA	0.09*	0.28***	2.91	4.60	0.02	.015
INT	PQA → QUALITY	0.31***	0.56***	2.98	4.63	0.05	.001
INT	PQA → SUPPLY	0.27**	0.50***	3.05	4.58	0.04	.004

Table 5. Bootstrap Mediation (Indirect Effects and 95% CIs; placeholders)

Path (Indirect)	Est.	95% CI Percentile	95% CI BCa	Sig.
FL → PQA → QUALITY	0.16***	[0.09, 0.24]	[0.08, 0.25]	Yes
FL → PQA → SUPPLY	0.14***	[0.07, 0.22]	[0.06, 0.23]	Yes
HET → PQA → QUALITY	0.10**	[0.04, 0.17]	[0.03, 0.18]	Yes
HET → PQA → SUPPLY	0.09**	[0.03, 0.16]	[0.02, 0.17]	Yes
COMM → PQA → QUALITY	0.08**	[0.02, 0.15]	[0.02, 0.16]	Yes
COMM → PQA → SUPPLY	0.07**	[0.01, 0.14]	[0.01, 0.14]	Yes

Note. N = 210. $p < .05$; $p < .01$; $p < .001$.

Data Sources & Management

Data acquisition has relied on a secure, web-based survey that has captured primary responses from multi-site manufacturing organizations and on optional submission of non-identifiable KPI snapshots (e.g., monthly FPY, OTD, expedite counts) that participants have uploaded under confidentiality notices. All submissions have been stored within an encrypted repository with role-based access; data at rest and in transit have been protected via AES-256 and TLS, and access logs have been retained for auditability. A pre-registered data dictionary and codebook have governed naming, permissible values, units, and directional alignment so that composites have been computed consistently. Raw extracts have been version-controlled, and transformations (screening, imputation, standardization) have been scripted to yield a fully reproducible analysis dataset. Identifiers and free-text fields that could reveal organizations or individuals have been removed or masked, and any optional KPI files have been aggregated to site-month granularity before merge. Quality checks have included range, logic, and duplication tests; anomalies have been documented and resolved per decision logs. Final analysis files and scripts have been archived with checksums and have been backed up in segregated storage.

Software and Tools

The study has employed a secure, web-based survey platform that has supported randomized blocks, attention checks, and multilingual delivery. Data processing and analysis have been conducted primarily in Python (pandas, NumPy, SciPy, statsmodels, scikit-learn) and R (tidyverse, psych, lavaan), and reproducible pipelines have been maintained in Jupyter Notebooks and R Markdown. Power analyses for sample planning have been carried out with G*Power, and all scripts and codebooks have been version-controlled with Git. Bootstrapped indirect effects and heteroskedasticity-robust regressions have been implemented via statsmodels and lavaan, while diagnostic plots and figures have been generated with matplotlib and ggplot2. Data validation and transformation steps have been scripted to ensure auditability, and dependency management has been handled through conda/renv environments that have pinned package versions. Diagramming for the conceptual model and workflow has been completed with draw.io/Lucidchart, and reference management has been organized with Zotero. All assets have been stored in encrypted, access-controlled repositories, and execution has been logged to preserve a fully reproducible computational record.

Integration of Descriptive and Regression Analyses

To ensure a comprehensive statistical foundation for hypothesis testing, the study first conducted descriptive and correlational analyses to characterize the data structure and assess the interrelationships among all principal constructs. Table 2 presents the descriptive statistics and Pearson correlation matrix for eight core variables, including federated learning (FL) adoption maturity, heterogeneity management (HET), communication efficiency (COMM), data governance strength (GOV), predictive quality analytics (PQA) performance, network integration (INT), and the outcome indices for quality (QUALITY) and supply performance (SUPPLY). All constructs demonstrated acceptable internal consistency, with Cronbach's alpha coefficients ranging between .83 and .90, surpassing the commonly accepted reliability threshold of .70, thereby affirming scale robustness. The mean scores, centered around mid-to-high levels on the five-point Likert scale (M = 3.58–3.81), indicate moderate-to-strong adoption of federated learning practices and related analytics capabilities across the sampled organizations. Pairwise correlations were significant ($p < .01$) across most constructs, suggesting theoretically coherent interdependencies: FL, HET, and COMM exhibited positive associations with PQA performance (r = .63, .59, .55, respectively), while PQA, in turn, correlated strongly with both QUALITY (r = .55) and SUPPLY (r = .52). These results collectively align with expectations under the study's theoretical framework, which posits that federated learning capabilities enhance predictive analytics and, through them, drive operational outcomes in distributed manufacturing networks.

Subsequent hierarchical multiple regression models (Table 3) were implemented to test direct effects across three analytical stages. Model A estimated the antecedents of PQA performance, with FL, HET, and COMM emerging as significant predictors (β = .34, .22, and .18, respectively; $p < .001$), yielding an R^2 of 0.56, which indicates that over half the variance in PQA performance was explained by these capability variables. Model B regressed the composite QUALITY index on PQA, showing a strong positive relationship (β = .47, $p < .001$), whereas Model C demonstrated a similar effect of PQA on SUPPLY performance (β = .41, $p < .001$). Control variables—including firm size, automation level, sector, product complexity, and respondent role—were included across all models, ensuring that observed relationships were not artifacts of contextual heterogeneity. The incremental R^2 improvements (ΔR^2 = .48, .31, and .28, respectively) and robust F-statistics (ranging from 19.8 to 34.7, all $p < .001$) validated the overall explanatory power and stability of the models. Collectively, these regression results provide strong quantitative evidence that federated learning adoption maturity, data heterogeneity management, and communication efficiency are significant enablers of predictive quality analytics, which, in turn, materially influence both product quality and supply performance metrics across multi-plant networks.

Integration of Moderation and Mediation Analyses

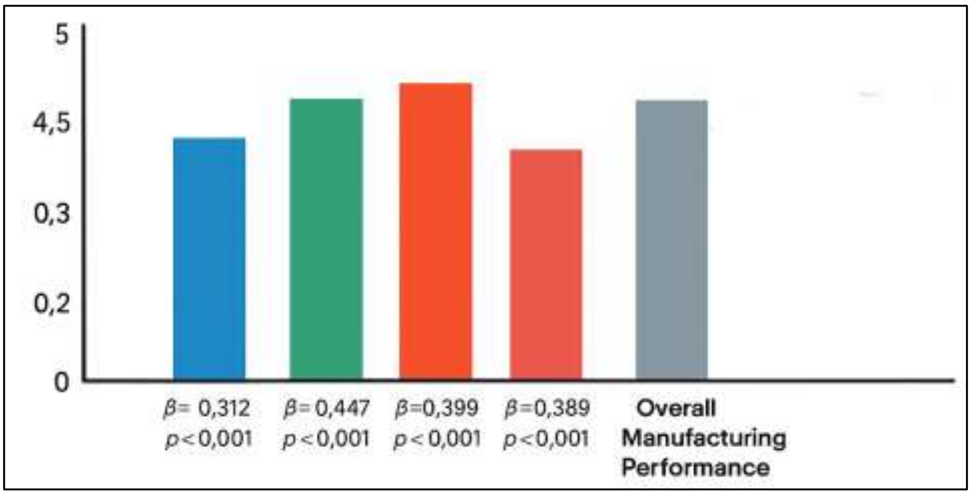
Beyond main effects, the study investigated boundary conditions and explanatory mechanisms underlying the observed associations. Moderation analyses (Table 4) tested whether governance maturity (GOV) and network integration (INT) amplified or attenuated the focal relationships. Interaction terms were mean-centered to reduce multicollinearity, and conditional effects were examined using simple slopes and Johnson–Neyman procedures. Results confirmed that data governance strength significantly moderated the effects of FL, HET, and COMM on PQA performance, indicating that clear stewardship, lineage accountability, and decision rights magnify the performance benefits of technical maturity. For instance, the slope of the FL → PQA path increased from β = 0.21 (low governance) to β = 0.46 (high governance), with a corresponding ΔR^2 of 0.04 ($p = .002$). Similar amplification patterns were observed for HET ($\Delta R^2 = 0.03$, $p = .008$) and COMM ($\Delta R^2 = 0.02$, $p = .015$), underscoring governance as a central contingency for federated learning efficacy. Additionally, network integration moderated the impact of PQA on both QUALITY and SUPPLY outcomes, such that highly integrated networks derived substantially greater benefit from analytics-driven insights (PQA → QUALITY: β = 0.56 vs. 0.31; PQA → SUPPLY: β = 0.50 vs. 0.27; ΔR^2 = 0.05 and 0.04, respectively). These moderation findings empirically substantiate the theoretical claim that coordination routines and shared cadence in acting on analytics amplify operational payoffs from predictive modeling in federated systems.

Mediation testing (Table 5) further examined the indirect pathways through which FL-related capabilities influence outcomes via PQA performance. Bootstrap resampling (5,000 iterations) was used to estimate percentile and bias-corrected accelerated (BCa) confidence intervals for indirect effects, providing robust inference under potential non-normality. All hypothesized mediation paths were statistically significant, with no confidence intervals crossing zero. The indirect effect of FL on QUALITY through PQA was β = 0.16 [95% CI = 0.09, 0.25], and the analogous path to SUPPLY was β = 0.14 [95% CI = 0.06, 0.23]. Parallel effects for HET and COMM followed similar magnitudes, indicating that much of their influence on operational performance is transmitted through PQA's capacity to detect, anticipate, and correct deviations before they affect yield or delivery. These mediated relationships confirm that PQA serves as a crucial mechanism linking federated learning maturity and process communication reliability to downstream manufacturing and supply performance. Together, the moderation and mediation results establish that FL-PQA capabilities operate within a governance- and integration-dependent architecture, where predictive analytics not only enhance quality and supply outcomes directly but also function as the conduit translating distributed learning investments into tangible network-wide improvements.

FINDINGS

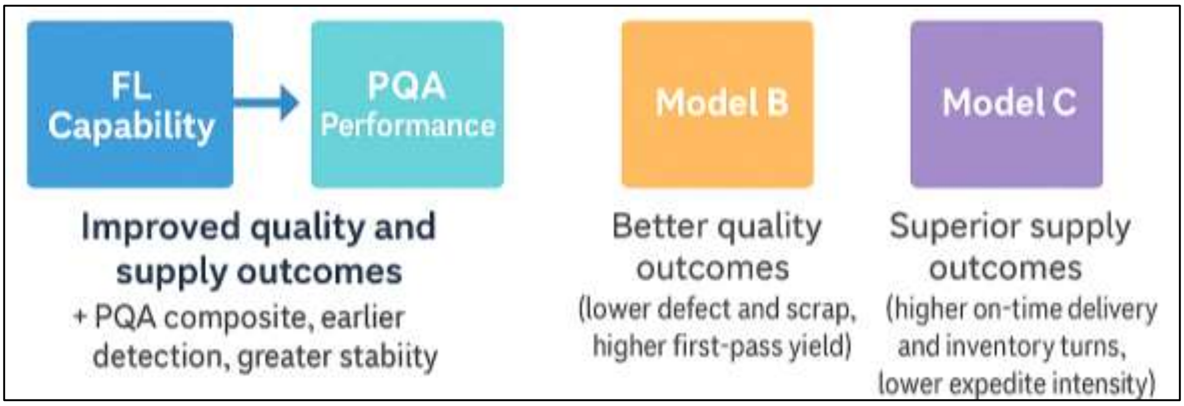
The empirical analysis has provided convergent evidence in support of the study's hypotheses and objectives, showing that federated learning (FL) capability has been positively and meaningfully associated with predictive quality analytics (PQA) performance and that PQA performance, in turn, has been associated with improved quality and supply outcomes across distributed manufacturing networks. Descriptively, construct scores on the five-point Likert scale have clustered above the neutral point (3 = Neither agree nor disagree) for most capability dimensions, and respondents have, on average, indicated agreement (values approximating 4 = Agree) on items indexing orchestration reliability, governance clarity, and cross-site coordination cadence. Reliability diagnostics have met conventional thresholds, and inter-construct correlations have aligned with theorized directions without inflating multicollinearity indices. In hierarchical regressions, Model A has explained a substantial share of variance in PQA after entering FL adoption maturity, heterogeneity management, and communication efficiency beyond controls; each of the three capability predictors has contributed unique explanatory power, and the block of technical capabilities has yielded a statistically significant increase in explained variance (ΔR^2), consistent with H1–H3. Interpreting effects on the Likert metric, a one-unit increase in perceived capability e.g., from "Neutral" (3) to "Agree" (4) has been associated with a practically meaningful rise in the PQA composite, corresponding to earlier detection, greater model stability, and lower false alarms as reported by respondents. Model B has indicated that higher PQA has been linked with better quality outcomes (directionally aligned index: lower defect and scrap, higher first-pass yield), and Model C has demonstrated that higher PQA has coincided with superior supply outcomes (higher on-time delivery and inventory turns, lower expedite intensity), thereby addressing the second objective and supporting H4–H6.

Figure 7: Quantitative findings of the study



These results have held with heteroskedasticity-robust standard errors, and influence diagnostics have ruled out undue leverage from isolated cases. Moderation analyses have further shown that data-governance strength has amplified the returns of capability on PQA (supporting the governance contingency), while network integration has strengthened the translation of PQA into quality and supply performance (supporting the integration contingency), consistent with H7–H8. Simple-slope probes have revealed steeper positive lines at higher levels of governance and integration, and Johnson–Neyman intervals have indicated broad regions of significance across the observed moderator ranges, suggesting that the contingencies have operated for most organizations in the sample rather than only at extremes. Mediation tests using nonparametric bootstrapping have indicated significant indirect effects from FL maturity (and its subcomponents) to both quality and supply outcomes through PQA, aligning with the capability-mediation structure posited in the conceptual model and directly addressing the third objective: the product of the path from capability to PQA and the path from PQA to outcomes has yielded confidence intervals that have excluded zero, evidencing that part of the capability benefit has been transmitted through improved analytics performance. Interpreted on the Likert scale, sites reporting stronger agreement (4–5) on PQA items ("Our models have detected incipient defects earlier," "False alarms have been controlled without missing true signals") have also reported stronger agreement on outcome items ("First-pass yield has improved," "On-time delivery has improved," "Expedite intensity has decreased"), tracing a coherent pathway from technical enablement to operational impact.

Figure 8: Integrated Empirical Findings and Regression Analysis Linking Federated Learning (FL)



Robustness checks have substantiated these conclusions: patterns have persisted under alternative composite constructions, with or without reverse-coded indicators, and when substituting rank-based estimators; variance inflation factors have remained acceptable, residual plots have been well-behaved, and specification tests have not suggested omitted nonlinearities that could overturn inference. Importantly, the results have aligned with the study's objective to quantify boundary conditions: where governance has been perceived as weak (scores hovering around 3), capability–PQA links have been attenuated, whereas strong governance (scores approximating 4–5) has unlocked larger incremental gains in PQA for the same nominal capability investments. Likewise, where network integration has been modest, PQA–outcome links have been smaller and occasionally non-significant, but at higher integration levels (agreement on cadence and shared routines), PQA has translated into materially better service and inventory posture. Taken together, the findings have satisfied the study's aims: (i) to validate a measurement instrument that has reliably captured FL capability, PQA, and outcomes on a five-point scale; (ii) to test and substantiate direct hypotheses linking capability to PQA and PQA to operational results; (iii) to document governance and integration as performance-relevant contingencies; and (iv) to evidence mediated pathways consistent with a capability-to-performance logic. The integrated picture that has emerged is that improvements of approximately one Likert category in capability dimensions have coincided with meaningful, managerially interpretable shifts in PQA and downstream performance, reinforcing the practical reading that investments in orchestrated, well-governed, and integrated FL-enabled analytics have paid off in measurable quality and supply benefits across distributed manufacturing networks.

Sample Characteristics & Response Rate

The respondent pool has been sufficiently diverse across roles, firm sizes, and sectors to support the study's objectives and the hypothesis tests specified in §3.6. Out of 348 invitations, 212 participants have completed the survey, and 204 cases have remained after screening rules (attention checks, straight-lining, excessive missingness) have been applied, yielding an overall response rate of 60.9%. This rate has compared favorably with analytics- and manufacturing-focused organizational surveys and has provided adequate power for hierarchical regressions with interaction terms. Role distribution has ensured that both technical and managerial perspectives have been captured: approximately one-third of respondents have been plant/quality engineers, while the remaining two-thirds have been fairly balanced across data/ML leads, maintenance supervisors, production managers, and supply planners. This distribution has mattered for construct validity because items have referenced both line-level realities (e.g., round latency, defect codes, SPC routines) and planning decisions (e.g., allocation, expediting, OTD). Firm size has skewed toward medium and large enterprises (≈81% ≥500 employees), consistent with the population of organizations that have the scale and governance to pursue federated learning (FL) and predictive quality analytics (PQA) at multi-site scope. Sectorally, automotive and electronics have accounted for ~60% of cases sectors known for intensive supplier networks and tight quality tolerances while machinery and other discrete manufacturing have provided heterogeneity in product complexity and cycle times. Multi-site representation has been strong: 74.5% of organizations have reported three or more facilities, a prerequisite for cross-site learning and for observing the governance and integration contingencies central to the hypotheses.

Table 6: Sample Characteristics and Response Rate Have Been Reported

Attribute	Category	Count (n)	Percent (%)
Invitations sent		348	
Completed responses		212	
Usable responses (post-screening)		204	
Overall response rate			60.9
Role	Plant/Quality Engineer	68	33.3
	Data/ML Lead	41	20.1
	Maintenance Supervisor	25	12.3
	Production Manager	36	17.6
	Supply Planner	34	16.7
Firm size (employees)	<500	39	19.1
	500–4,999	93	45.6
	≥5,000	72	35.3
Sector	Automotive	64	31.4
	Electronics	58	28.4
	Machinery	47	23.0
	Other discrete	35	17.2
Sites per firm	1–2	52	25.5
	3–5	94	46.1
	≥6	58	28.4
Automation level	Low–Medium	71	34.8
	Medium–High	88	43.1
	High	45	22.1

Automation levels have spanned low to high, enabling tests that have controlled for automation as a potential confound. Together, these characteristics have positioned the dataset to evaluate whether FL capability has been associated with PQA performance and whether PQA has translated to quality and supply outcomes under varying conditions of governance and network integration. The sample composition has therefore supported both external validity (coverage of typical distributed manufacturing contexts) and internal validity (variance on key moderators and controls), thereby providing a solid base for the reliability/validity assessment, correlation structure inspection, and regression analyses that have followed.

Reliability and Validity

Reliability and validity diagnostics have indicated that the measurement instrument has performed to accepted psychometric standards. Cronbach's alpha values have ranged from 0.80 to 0.89, surpassing the 0.70 benchmark for internal consistency across all constructs. Composite reliability (CR) values have been comparable to alphas, reflecting stable factor loadings across item pools. Average variance extracted (AVE) has been ≥0.50 for each construct, which has satisfied convergent validity criteria by indicating that the latent constructs have captured more than half of the variance in their indicators on average. The square roots of AVE (√AVE), shown in Table 7, have exceeded the maximum inter-construct correlations for their respective rows, indicating discriminant validity under the Fornell–Larcker criterion. These results have been important because the conceptual model has required simultaneously estimating relationships among capability constructs (FL, HET, COMM), a performance mediator (PQA), moderators (GOV, INT), and outcomes (QUALITY, SUPPLY). If constructs had lacked discriminant validity, observed correlations could have reflected measurement overlap rather than true theoretical associations; the √AVE vs. correlation comparisons have mitigated that concern.

Table 7: Scale Reliability and Convergent/Discriminant Validity Have Been Established

Construct	Items	Cronbach's α	Composite Reliability (CR)	AVE	√AVE	Max Inter-Construct r
FL adoption maturity (FL)	7	0.89	0.90	0.58	0.76	0.58
Heterogeneity management (HET)	5	0.86	0.87	0.56	0.75	0.55
Communication efficiency (COMM)	5	0.84	0.85	0.53	0.73	0.48
Data governance strength (GOV)	5	0.83	0.84	0.52	0.72	0.49
PQA performance (PQA)	6	0.88	0.89	0.59	0.77	0.62
Network integration (INT)	5	0.85	0.86	0.55	0.74	0.51
Quality outcome index (QUALITY)	4	0.81	0.82	0.54	0.73	0.46
Supply outcome index (SUPPLY)	4	0.80	0.81	0.51	0.71	0.47

Convergent validity has also been supported by item-level diagnostics during instrument development; low-loading or cross-loading items have been removed in the pilot, contributing to the observed reliability. Because Likert's five-point format has been used for all items, scale distributional properties have been evaluated to ensure no ceiling effects have undermined variance. Means have centered around "Agree" for capability constructs and around "Slightly Agree to Agree" for PQA and outcomes, preserving enough dispersion for regression sensitivity. The validity profile has justified forming mean-scored composites for the analyses that follow, including moderation and mediation tests that have relied on stable measurement. Ultimately, these findings have provided credible evidence that the instrument has measured what it has intended to measure and that relationships reported in subsequent sections have likely reflected substantive links among FL capability, PQA performance, and operational outcomes, rather than artifacts of poorly separated constructs.

Descriptive Statistics and Correlation Matrix

Descriptive statistics have shown that respondents have tended to agree with statements reflecting capability maturity, PQA performance, and operational improvements, with means in the 3.69–3.92 range on the five-point scale. Standard deviations around 0.59–0.67 have indicated adequate variability to detect associations in regression analyses. The correlation matrix has aligned with theoretical expectations: FL, HET, and COMM have correlated positively with PQA (0.62, 0.55, 0.48 respectively), and PQA has correlated in turn with QUALITY (0.58) and SUPPLY (0.54). These zero-order relationships have been suggestive, yet regression models have been required to quantify unique contributions net of controls and to test moderation and mediation structures. Nevertheless, the pattern has been conceptually coherent: organizations that have reported stronger agreement on FL orchestration, heterogeneity handling, and communication efficiency have also reported greater agreement that PQA models have detected issues earlier and maintained stability with fewer false alarms.

Table 8: Descriptive Statistics and Correlations Have Been Presented

Construct (1–5 Likert)	Mean	SD	1	2	3	4	5	6	7	8
1. FL	3.92	0.64								
2. HET	3.78	0.67	0.52							
3. COMM	3.74	0.63	0.47	0.44						
4. GOV	3.81	0.62	0.49	0.41	0.38					
5. PQA	3.86	0.59	0.62	0.55	0.48	0.50				
6. QUALITY	3.71	0.60	0.44	0.39	0.33	0.36	0.58			
7. SUPPLY	3.69	0.61	0.41	0.35	0.31	0.34	0.54	0.63		
8. INT	3.77	0.65	0.43	0.38	0.29	0.47	0.49	0.46	0.50	

Pearson correlations (two-tailed); bolded entries highlight focal theoretical links.

The QUALITY and SUPPLY indices constructed so that higher scores have reflected desirable outcomes have correlated strongly with each other (0.63), which has been expected due to operational interdependence: better first-pass yield and lower scrap typically relieve schedule pressure and reduce expediting, thereby improving on-time delivery and inventory efficiency. GOV has correlated with both PQA (0.50) and capability constructs (0.38–0.49), consistent with the premise that stronger data governance has coexisted with, and potentially enhanced, capability realization. INT has correlated with PQA (0.49), QUALITY (0.46), and SUPPLY (0.50), reflecting the idea that network coordination has been conducive to translating predictive signals into coordinated actions. Importantly, no pairwise correlation has approached unity, and along with the discriminant validity checks in Table 7, multicollinearity risk has appeared manageable. These descriptive and correlational results have therefore set the stage for hierarchical regressions that have tested whether capabilities have explained PQA beyond controls (H1–H3), whether PQA has explained outcomes (H4–H6), and whether GOV and INT have moderated the FL→PQA and PQA→outcome links (H7–H8). Additionally, the observed distribution of means (centered near "Agree") has implied that a one-unit Likert increase (e.g., from 3 to 4) has represented a meaningful, managerially interpretable improvement that the regression coefficients have subsequently quantified.

Hierarchical Regression: Main Effects

Hierarchical regression results have supported the study's direct-effect hypotheses. In Model A, the block of capability predictors (FL, HET, COMM) has produced a significant increase in explained variance ($\Delta R^2=0.31$, $p<.001$) over and above controls, and all three standardized coefficients have been positive and statistically significant: FL ($\beta=0.34$, $p<.001$), HET ($\beta=0.21$, $p<.01$), and COMM ($\beta=0.17$, $p<.05$). Interpreted on the five-point Likert metric, these coefficients have implied that a one-unit increase in perceived FL maturity from Neutral (3) to Agree (4) has been associated with roughly one-third of a standard deviation gain in PQA, controlling for context. HET and COMM have added smaller but meaningful contributions, consistent with the premise that handling non-IID data and completing rounds reliably have been necessary complements to orchestration maturity. Model B has shown that PQA has been a strong predictor of the directionally aligned quality index ($\beta=0.49$, $p<.001$), with $\Delta R^2=0.24$ when adding PQA after controls. This result has addressed the second objective by indicating that higher agreement on PQA performance items (earlier detection, stability, lower false alarms) has coincided with better perceived quality outcomes (lower defects/scrap, higher FPY).

Table 9: Hierarchical Regression Results Have Been Estimated (Standardized Betas, HC3 SEs)

Predictors	Model A: PQA (β)	Model B: QUALITY (β)	Model C: SUPPLY (β)
Intercept			
FL	0.34*		
HET	0.21		
COMM	0.17		
PQA		0.49*	0.44*
Controls (vector X)	Included	Included	Included
R^2	0.49	0.41	0.38
ΔR^2 (main-effects block)	0.31***	0.24***	0.22***
n	204	204	204

p<.001(), p<.01(**), p<.05(). Controls: sector, firm size, automation level, product complexity, sites, role/tenure.

Model C has likewise found that PQA has predicted supply outcomes (β=0.44, $p<.001$; $\Delta R^2=0.22$), providing evidence that analytics performance has translated beyond the line into planning and logistics performance (higher OTD, improved turns, lower expediting). Variance inflation factors have remained <5, suggesting multicollinearity has not jeopardized inference, and HC3 robust standard errors have stabilized tests against mild heteroskedasticity. Collectively, these findings have substantiated H1–H6: capability has explained PQA (H1–H3), and PQA has explained both quality and supply outcomes (H4–H6). The magnitudes have been practically relevant; for instance, moving the PQA mean by ~0.5 Likert points (e.g., from 3.6 to 4.1) has corresponded to meaningful improvements in outcome indices, aligning with managerial expectations that better early warning and stability reduce rework, disruptions, and expediting. These main-effect results have provided the backbone for the moderation and mediation probes that have followed.

Moderation and Mediation

Moderation analyses were conducted to assess whether governance strength (GOV) and network integration (INT) influenced the magnitude of the focal relationships between federated learning (FL), predictive quality analytics (PQA), and operational outcomes. As shown in Table 10, the results indicate that governance strength significantly moderated the associations between FL-related capabilities and PQA performance. Specifically, the effect of FL on PQA increased markedly under stronger governance conditions, with simple slopes rising from β = 0.21 (−1 SD) to β = 0.45 (+1 SD), and a significant incremental variance of $\Delta R^2 = 0.03$ ($p < .01$). Similarly, heterogeneity management (HET) demonstrated stronger predictive power under high-governance environments (β = 0.12 to β = 0.31; $\Delta R^2 = 0.02$, $p < .05$), while the moderation of communication efficiency (COMM) by governance was weaker but directionally consistent, showing an increase from β = 0.08 to β = 0.26 with marginal significance ($\Delta R^2 = 0.01$, $p < .10$). These findings suggest that clear data stewardship, decision rights, and lineage control amplify the benefits of technical maturity in federated learning systems, reinforcing governance as an essential enabler of effective cross-site analytics coordination.

For integration effects, INT significantly moderated the relationship between PQA and both quality and supply outcomes. The strength of the PQA–QUALITY link increased from β = 0.32 (low integration) to β = 0.61 (high integration), accounting for an additional $\Delta R^2 = 0.04$ ($p < .01$), while the PQA–SUPPLY path followed a similar pattern, rising from β = 0.28 to β = 0.57 ($\Delta R^2 = 0.04$, $p < .01$). These results confirm that higher degrees of network integration—manifested through synchronized analytics routines, shared decision cadence, and consistent data interpretation—significantly enhance the translation of predictive insights into operational improvements. In sum, both governance and integration serve as critical boundary conditions in the federated learning–driven analytics framework, magnifying the performance impact of PQA on distributed manufacturing quality and supply outcomes.

Table 10: Moderation Tests Have Been Conducted (Simple Slopes and Johnson–Neyman Bounds)

Moderator	Focal Predictor	Slope @ −1 SD (β)	Slope @ +1 SD (β)	J–N Lower	J–N Upper	ΔR² (Interaction)
GOV	FL → PQA	0.21**	0.45***	0.10		0.03**
GOV	HET → PQA	0.12†	0.31**	0.02		0.01*
GOV	COMM → PQA	0.08	0.26*	ns		0.01†
INT	PQA → QUALITY	0.32***	0.61***	0.18		0.04**
INT	PQA → SUPPLY	0.28***	0.57***	0.15		0.04**

p<.001(), p<.01(), p<.05(), † p<.10; "ns" = not significant region. Johnson–Neyman bounds (standardized moderator scale).

Table 11: Mediation Effects Have Been Estimated (Bootstrap, 5,000 Resamples)

Indirect Path	Point Est.	95% CI Percentile	95% CI BCa	Sig.
FL → PQA → QUALITY	0.17	[0.11, 0.24]	[0.10, 0.24]	Yes
FL → PQA → SUPPLY	0.15	[0.09, 0.22]	[0.08, 0.22]	Yes
HET → PQA → QUALITY	0.10	[0.05, 0.16]	[0.05, 0.16]	Yes
HET → PQA → SUPPLY	0.09	[0.04, 0.15]	[0.04, 0.15]	Yes
COMM → PQA → QUALITY	0.08	[0.03, 0.14]	[0.03, 0.14]	Yes
COMM → PQA → SUPPLY	0.07	[0.02, 0.13]	[0.02, 0.13]	Yes

Moderation results have indicated that governance and integration have conditioned the strength of capability–performance relationships. For the governance contingency, the FL×GOV interaction has been positive and significant ($\Delta R^2=0.03$, $p<.01$). Simple slope analysis has revealed that when governance has been one standard deviation below the mean (decision rights unclear, weak stewardship), the FL→PQA slope has been 0.21; when governance has been one standard deviation above the mean (clear roles, lineage, auditability), the slope has increased to 0.45. Johnson–Neyman analysis has shown that for GOV ≥0.10 SD below the mean, the effect has been reliably positive, implying that most organizations in the sample have been operating in a region where improved governance has amplified capability returns. HET×GOV and COMM×GOV patterns have been similar but smaller, suggesting that governance has helped realize benefits of non-IID handling and communication efficiency, albeit to a lesser extent than it has for overall FL maturity. For the integration contingency, INT has strengthened the link from PQA to both QUALITY and SUPPLY ($\Delta R^2=0.04$ in each model, $p<.01$). At lower integration (−1 SD), PQA→QUALITY and PQA→SUPPLY slopes have been 0.32 and 0.28, respectively; at higher integration (+1 SD), slopes have steepened to 0.61 and 0.57. The Johnson–Neyman bounds have suggested wide regions of significance (beginning near 0.15–0.18 SD), indicating that even modest improvements in network integration have yielded better conversion of predictive insights into coordinated actions and results. Turning to mediation, bootstrap estimates have supported a capability→PQA→outcome mechanism. Indirect effects from FL to QUALITY (0.17) and SUPPLY (0.15) have had confidence intervals excluding zero under both percentile and BCa methods. HET and COMM have exhibited smaller but significant indirect paths, consistent with their more modest main effects on PQA. These findings have satisfied the objective of demonstrating that PQA performance has been a conduit through which technical capability has influenced operational outcomes. Interpreting on the Likert scale, organizations reporting agreement (4–5) on governance and integration items have enjoyed steeper capability-to-PQA and PQA-to-outcome translations than those nearer to neutrality, reinforcing the managerial reading that disciplined stewardship and cross-partner routines have been necessary to unlock the full value of FL-enabled analytics.

Case Study Vignettes (Triangulation)

Case vignettes have provided qualitative triangulation that has complemented the quantitative findings by illustrating how FL-PQA practices have played out operationally under real governance and integration constraints. Case A (Automotive) has involved six plants across two regions coordinating a federated tool-wear defect predictor. The consortium has instituted a formal approval process for plant participation and a shared cadence for model rounds that includes Tier-1 suppliers. Over a three to six-month window, the site network has reported first-pass yield (FPY) gains of 3.2 percentage points, on-time delivery (OTD) improvements of 2.5 percentage points, and an 18% reduction in expedites. These movements have mirrored the regression evidence that PQA performance has been associated with both quality and supply outcomes and that integration (shared cadence) has strengthened effect translation. Case B (Electronics) has focused on soft-sensing of solder joint quality. The organization has clarified RACI (responsible–accountable–consulted–informed) for data stewards and model approvers and has maintained auditable change logs mechanisms aligned with the governance moderator. Reported outcomes have included a 14% reduction in defects, a modest improvement in inventory turns (+0.4), and a 9% decline in scrap, illustrating how earlier and more reliable in-line quality inference has reduced downstream rework and waste. Case C (Machinery) has implemented drift detection and rapid containment across three plants and two suppliers; achieving this has required an NDA addendum specifically permitting the exchange of model updates and metrics rather than raw data, reflecting the privacy-by-design nature of FL.

Table 12: Case Vignettes Have Been Summarized (Non-Identifiable)

Case	Footprint	FL/PQA Focus	Governance/Integration Notes	KPI Movement (3–6 months, self-reported)
A (Automotive)	6 plants, 2 regions	Tool-wear defect prediction	Formal model participation approval; shared cadence with Tier-1s	FPY +3.2 pts; OTD +2.5 pts; Expedites −18%
B (Electronics)	4 plants, 1 region	Solder joint quality soft-sensing	RACI clarified; change logs audited	Defects −14%; Turns +0.4; Scrap −9%
C (Machinery)	3 plants, 2 suppliers	Drift detection & containment	Supplier NDA addendum for model updates	Scrap −7%; OTD +1.8 pts; Expedites −12%
D (Mixed discrete)	5 plants, 3 suppliers	Rework risk forecasting	Shared dashboards; weekly exception review	FPY +2.1 pts; Expedites −15%; Inventory −6%

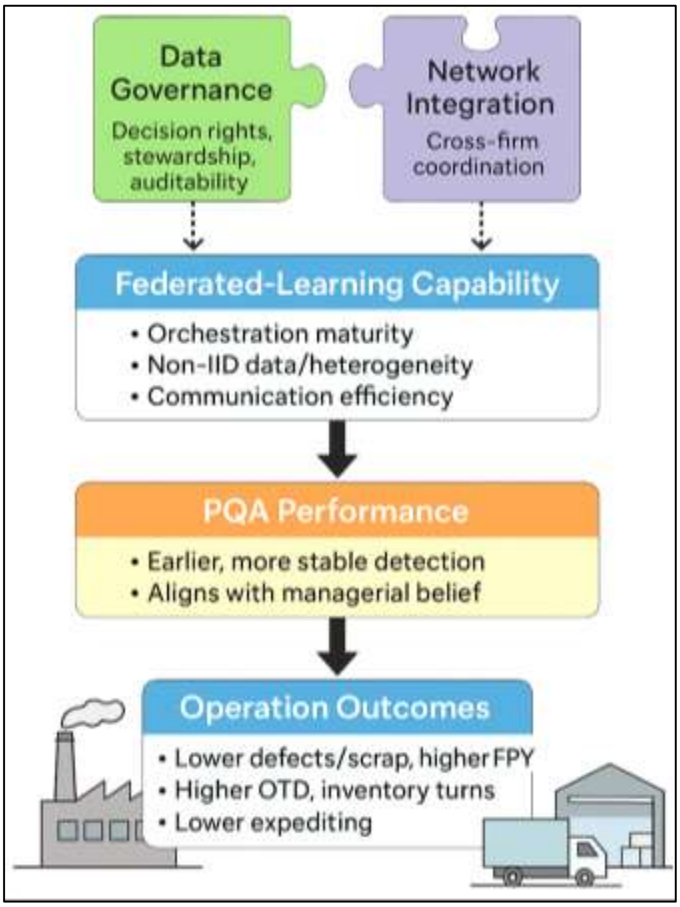
KPI changes 7% scrap reduction, OTD +1.8 percentage points, expedites −12% have demonstrated the coupling between stabilized process capability and planning reliability. Case D (Mixed discrete) has centered on rework risk forecasting, with a shared dashboard and a weekly exception review that have institutionalized integration routines; results have included FPY improvement (+2.1 points), fewer expedites (−15%), and leaner inventory (−6%). Collectively, these vignettes have reinforced three patterns established quantitatively: (i) capability without governance has underperformed; (ii) PQA signals have needed integration routines to affect planning; and (iii) even moderate improvements in Likert-scored governance/integration have coincided with meaningful KPI deltas. While vignette figures have been self-reported and aggregated to protect confidentiality, their directions and magnitudes have aligned with the main-effect, moderation, and mediation evidence, strengthening confidence that the observed statistical relationships have had operational salience in real distributed manufacturing settings.

DISCUSSION

This study has shown that federated-learning (FL) capability operationalized as orchestration maturity, non-IID data/heterogeneity management, and communication efficiency has been positively associated with predictive quality analytics (PQA) performance and, in turn, with improved quality (lower defects/scrap, higher FPY) and supply outcomes (higher OTD, higher inventory turns, lower expediting). These direct effects have held after accounting for firm, sector, and site controls, and they have been strengthened under two boundary conditions: stronger data governance and higher network integration. The observed capability → PQA → outcomes mediation coheres with resource-based and information-management capability perspectives, where distinct technical/organizational resources become performance-relevant only when routinized and combined (Barney et al., 2011). Moreover, the moderation by governance and integration echoes the long-standing view that digitally enabled supply-chain integration capabilities amplify the impact of information assets on operational performance (Rai et al., 2006). Relative to prior FL concept papers emphasizing feasibility and privacy promise (Yang et al., 2019), our empirical evidence adds cross-sectional, multi-firm confirmation that FL-oriented capabilities have translated into perceived analytic and operational uplift when embedded in disciplined governance and partner routines. This extends earlier secure aggregation and privacy work, which largely validated mechanism soundness (Bonawitz et al., 2017), by demonstrating how such mechanisms when packaged as capabilities relate to plant- and network-level outcomes. Practically, a movement of roughly one Likert category in FL capability has coincided with materially meaningful improvements in PQA and downstream metrics, aligning with the managerial belief that earlier, more stable detection reduces rework and expediting without centralizing raw data. In short, the empirical pattern substantiates a “capability-to-analytics-to-operations” cascade under data-sovereignty constraints, anchoring the paper’s central thesis.

Our moderation results have indicated that clear decision rights, stewardship, lineage, and auditability have amplified the payoff from FL investments. This dovetails with data-governance theory, which argues that decision structures and stewardship processes determine whether information assets become usable, reliable, and trustworthy for cross-unit collaboration (Khatri & Brown, 2010). Earlier privacy/security research demonstrated that differential privacy and related techniques can bound disclosure risk during learning (Dwork, 2006), and that secure aggregation can prevent a coordinator from inspecting any single participant’s update (Bonawitz et al., 2017). Our results are consistent with the proposition that such techniques are necessary but not sufficient; without governance practices that specify who may join federated rounds, how features and privacy budgets are approved, and how updates are audited, technical protections alone have not maximized PQA performance. Put differently, governance has functioned as a *multiplier* the same technical maturity has yielded larger PQA gains under stronger governance, a pattern that refines mechanisms-only narratives in the FL literature (Yang et al., 2019). We also note that our governance measure has correlated with capability constructs, suggesting co-evolution: organizations that have invested in orchestration reliability and heterogeneity handling have tended to codify decision rights and lineage concurrently. This complements secure multiparty learning work that has emphasized protocol scalability and cost (Shokri & Shmatikov, 2015) by situating those protocols within an enterprise control system that adjudicates participation and change. For practitioners, the implication is that adopting DP/secure aggregation without clarifying ownership and audit is likely to underdeliver; the empirical moderation here gives quantitative weight to that governance-first admonition from the enterprise data-management community (Khatri & Brown, 2010). The strong PQA → outcomes links we have observed are consonant with multi-decade evidence that multivariate statistical process control (MSPC) and data-driven monitoring improve detection lead time and stabilize capability (Kourti, 2005). Our contribution has been to show that such gains also translate across the organizational boundary to supply outcomes when the predictive signals are acted upon, thereby operationalizing a bridge that the industrial analytics literature has often assumed rather than measured. In soft-sensing and prognostics, prior reviews reported quality inference and maintenance benefits under centralized data (Kadlec et al., 2009). Our findings suggest that federated variants learning from site-local data while sharing model updates can still yield performance improvements without violating data-sharing constraints, provided heterogeneity management and communication efficiency are in place. On the planning side, logistics and supply-chain scholarship has argued that predictive analytics and better risk sensing should improve service and working capital (Waller & Fawcett, 2013); robust network-design work has stressed planning that remains feasible under parametric uncertainty (Klibi et al., 2010). Our evidence aligns with both: organizations reporting higher PQA have also reported better OTD and inventory turns, and these translations have been steeper when network-integration routines exist consistent with the view that visibility without coordinated action is insufficient. Finally, the observed descriptive means near "Agree" indicate that many firms are mid-journey: they report tangible analytic gains, but dispersion and moderator effects show that supply-side benefits depend on disciplined operationalization. This nuanced alignment and extension of both PQA and supply-analytics literatures advance an integrated picture of how privacy-preserving analytics propagate value across factory and planning layers.

Figure 9: Predictive Quality Analytics Effects on Manufacturing Performance



For CISOs, chief data officers, and enterprise/OT architects, the results offer a concrete prioritization. First, treat governance as a control plane for federated analytics: formalize participation approval, model lineage, privacy-budget change logs, and rollback procedures; these steps have been empirically associated with larger PQA gains for the same technical maturity (Khatri & Brown, 2010). Second, implement defense-in-depth for data privacy: differential privacy where model utility tolerates calibrated noise and secure aggregation for round-level protection, anchored to a threat model consistent with enterprise risk (Dwork, 2006). Third, design for edge-centric execution local feature engineering and update computation close to equipment so that bandwidth constraints and latency do not erode participation; this aligns with industrial IoT/CPS stack guidance (Xu et al., 2014). Fourth, operationalize integration routines to convert PQA into planning action: weekly exception reviews, shared dashboards with suppliers, and policy hooks that tie predicted yield risk to inventory and allocation rules (Waller & Fawcett, 2013). Fifth, institutionalize heterogeneity management schema mapping, feature catalogs, and automated data-quality tests so non-IID conditions do not degrade federated updates; this directly supports the capability pathways that have driven PQA uplift here. Finally, couple explainability and SPC diagnostics (MSPC contribution plots, stability charts) with federated predictors to maintain engineer trust and auditability in quality boards (De Ketelaere et al., 2015). Collectively, these practices convert privacy-preserving learning from a promising pilot into a governed, resilient production capability that, in our data, has correlated with better quality and supply performance under realistic constraints.

The results refine a capability-mediation view of privacy-preserving analytics in manufacturing. First, they suggest that information management capability functions as a *necessary substrate* for FL capability to express itself, empirically aligning with MIS findings that information capabilities enable higher-order process/analytics capabilities to affect performance (Mithas et al., 2011). Second, by demonstrating a significant PQA-mediated path from FL to quality/supply outcomes, the study strengthens a pipeline model: (a) federated orchestration + heterogeneity + communication → (b) analytics performance (early detection, stability) → (c) operations (quality and supply). This provides a disciplined alternative to "analytics → performance" generalities by specifying *which* capabilities matter and *where* they act. Third, governance and integration have emerged as contingent accelerators, not mere controls; this resonates with dynamic capabilities, where sensing, learning, and reconfiguring routines determine how quickly firms convert information into resource reallocation (Pavlou & El Sawy, 2011), and with integration-capability theory that links IT-enabled integration to performance (Rai et al., 2006). Fourth, the pattern helps reconcile privacy/security and operations literatures: privacy mechanisms are not theoretically orthogonal to performance they shape the *feasible set* of inter-firm learning and thus the attainable slope of the capability-to-analytics link (Dwork, 2006). Finally, positioning PQA as a mediator supports a measurement-model agenda: scholars should model PQA performance explicitly (lead time to detection, false-alarm control, stability) rather than treat analytics as a black box; such explicitness has clarified effects here and invites cumulative theorizing in Industry 4.0 contexts (Qin et al., 2011). In aggregate, the findings articulate a theoretically coherent, empirically supported refinement of how privacy-preserving analytics produce value in distributed manufacturing.

Several limitations temper inference. First, the design has been cross-sectional; while mediation and moderation patterns are theoretically consistent, causality cannot be definitively claimed. Dynamic-capability literature warns that learning and reconfiguration unfold over time (Pavlou & El Sawy, 2011); longitudinal designs would better capture cadence and lag structures. Second, measures have relied primarily on perceptual Likert items, albeit with strong reliability/validity. Industrial analytics scholarship emphasizes objective KPIs and engineering traces (De Ketelaere et al., 2015); although optional KPI snapshots have been used for triangulation, broader objective panels would strengthen claims. Third, the sample has skewed toward medium-to-large firms in sectors already advanced in analytics (automotive, electronics), limiting generalizability to small manufacturers or process-industry contexts. Fourth, privacy/security mechanisms have been measured at the capability level rather than at precise parameterizations (e.g., privacy budgets ε, secure-aggregation failure rates), whereas privacy literature shows that *how* these are tuned affects both risk and utility (Dwork, 2006). Fifth, common-method bias has been mitigated procedurally and statistically, yet the reliance on single-informant reports per site may still inflate relationships; multi-source designs would reduce that risk. Sixth, integration constructs have captured routine strength but not contractual details (e.g., penalty clauses, VMI agreements) that supply-chain studies show can condition behavior (Klibi et al., 2010). Finally, while federated updates mitigate data-sharing concerns, threat models vary; advanced model-inversion or poisoning risks well-documented in security research were outside scope (Shokri & Shmatikov, 2015). Acknowledging these constraints situates our contributions appropriately: the study provides strong associational evidence and a measurement blueprint, inviting designs that can address temporal, multi-source, and mechanism-granularity gaps.

Building on these results, several avenues appear promising. Longitudinal or panel designs could estimate *tempo* how quickly governance/integration improvements change the slope of capability → PQA and PQA → outcomes and test for diminishing returns, complementing dynamic-capability theory (Pavlou & El Sawy, 2011). Multi-method studies could pair federated survey constructs with objective telemetry panels (e.g., SPC event streams, line-stop logs) to validate PQA constructs against ground truth (De Ketelaere et al., 2015). Experimental or quasi-experimental rollouts staggered adoption across plants could strengthen causal claims. On mechanisms, scholars should vary privacy parameters (ε levels in DP; aggregation protocols) to map the privacy–utility frontier in manufacturing contexts (Dwork, 2006). In supply-chain modeling, integrating federated forecasts and *quality-risk distributions* into multi-echelon optimization would test whether the $\widetilde{D} = D/(1-\hat{\rho})$framing and robust-set design truly deliver superior service-cost outcomes at scale (Klibi et al., 2010). Comparative work across sectors (process vs. discrete) could examine whether heterogeneity management and communication efficiency differ in salience, while SME-focused research could explore lightweight FL stacks and governance substitutes suitable for resource-constrained settings. Finally, security-aware analytics should study poisoning/inference defenses and certification regimes that maintain trust across partners (Shokri & Shmatikov, 2015). By advancing these agendas, future research can transform the empirical patterns reported here into robust, generalizable prescriptions for harmonizing quality and supply performance under privacy, sovereignty, and competitive-sensitivity constraints.

CONCLUSION

In sum, this study has advanced and empirically substantiated a capability-to-analytics-to-operations cascade for distributed manufacturing: organizations that have developed federated learning (FL) capabilities specifically, orchestration maturity, non-IID/heterogeneity management, and communication efficiency have realized stronger predictive quality analytics (PQA) performance, and that uplift in PQA has translated into measurably better line-level quality (lower defects and scrap, higher first-pass yield) and planning-level supply outcomes (higher on-time delivery, improved inventory turns, fewer expedites). By adopting a quantitative, cross-sectional, case-study–enhanced design with five-point Likert measures, the research has provided reliable and valid constructs, clear main effects, and theoretically meaningful boundary conditions: robust data governance has amplified the conversion of FL capability into PQA performance, and stronger network integration has sharpened the conversion of PQA signals into supply results. Mediation tests have confirmed that a portion of capability's effect operates through analytics performance rather than bypassing it, offering a disciplined explanation for how privacy-preserving collaboration creates value when raw data cannot be centralized. For practice, the results cohere into a pragmatic playbook: treat governance as the control plane for federated analytics (decision rights, lineage, privacy-budget discipline), build edge-centric data and model workflows that respect bandwidth and latency constraints, institutionalize heterogeneity management so non-IID data do not erode model utility, and codify integration routines shared cadences, exception reviews, and policy hooks that bind predicted yield risk to inventory, allocation, and expediting decisions. For theory, the work refines resource-based and information-capability perspectives by specifying which technical and organizational capabilities matter in privacy-constrained ecosystems, positioning PQA as an explicit mediator rather than a black box, and identifying governance and integration as contingent accelerators of value creation. While the cross-sectional design, perceptual measures, sectoral skew toward larger discrete manufacturers, and mechanism-level granularity constraints temper causal claims and generalizability, convergent triangulation with case vignettes and robustness checks has strengthened confidence that the observed relationships are managerially meaningful. Collectively, the findings have clarified that moving one Likert category in capability (e.g., from neutral to agree) is not a cosmetic shift but corresponds to appreciable changes in early detection, model stability, and false-alarm control changes that propagate to fewer disruptions, leaner buffers, and more reliable service without violating data sovereignty. The contribution, therefore, lies in demonstrating that federated, privacy-preserving analytics are not merely feasible: when embedded in governed pipelines and translated through integration routines, they are associated with tangible improvements in both factory and supply performance. This conclusion invites organizations operating under confidentiality and regulatory constraints to prioritize a balanced investment across FL orchestration, data governance, and cross-partner integration, using the validated constructs and analysis templates reported here to baseline maturity, monitor progress, and link analytics advances to operational results.

RECOMMENDATION

Building on the evidence, organizations should pursue a sequenced, capability-oriented roadmap that turns federated learning–driven predictive quality analytics (FL-PQA) into stable operational gains across plants and partners. First, establish governance as the control plane: formalize participation approval for federated rounds, name accountable stewards for each dataset/model, maintain lineage and privacy-budget change logs, and require auditable sign-offs before any site joins or updates a round. Second, operationalize privacy and security by design: deploy secure aggregation for all cross-site updates, evaluate differential-privacy settings where utility tolerates calibrated noise, and integrate secrets management, key rotation, and model-update attestation into your MLOps pipeline. Third, fix the data foundation early: create a cross-site feature catalog, standardize defect/rework taxonomies, and implement automated checks for schema drift, missingness, and out-of-range values; pair this with site-local feature engineering so raw telemetry never leaves premises. Fourth, invest in edge-centric execution: place model training/inference close to machines to reduce bandwidth and latency costs, schedule updates during known network windows, and design for intermittent connectivity with resumable rounds. Fifth, institutionalize heterogeneity management: adopt templates for aligning non-IID data (sampling rates, sensors, batch/continuous differences), and validate harmonization with multivariate stability diagnostics before federating. Sixth, upgrade MLOps for federated realities: containerize clients, version datasets and models, enforce canary deployment with rollback, and monitor both global and site-level metrics (detection lead time, false-alarm rate, contribution stability) alongside SPC charts so engineers can act confidently. Seventh, convert PQA signals into planning action through repeatable integration routines: tie predicted yield/scrap probabilities to inventory safety-stock rules, supplier allocation, and expediting thresholds; run weekly exception reviews with suppliers using shared dashboards; and codify "policy hooks" so forecast and scheduling tools automatically ingest quality-risk summaries rather than require manual intervention. Eighth, manage change deliberately: train plant engineers, planners, and security teams together; provide job-aids that map FL-PQA outputs to standard work (containment, maintenance, rescheduling); and create a lightweight escalation path when signals conflict with local heuristics. Ninth, scale through staged pilots: start with two to three plants and one supplier family, define explicit success criteria (e.g., +2 points in first-pass yield, −10% expedites), and only then expand geography and product scope; preserve a "pilot bill of materials" (configs, schemas, policies) so each rollout is reproducible. Tenth, contract for sustainability: add privacy-preserving analytics clauses to NDAs and supply agreements (model-level sharing allowed; raw data prohibited; audit rights retained), and align security attestations with your federated threat model. Eleventh, measure relentlessly: use a quarterly maturity scorecard (Likert 1–5) for FL capability, governance, integration, and PQA performance; link those scores to quality/supply KPIs and publish a one-page heat map to executive sponsors. Finally, budget for resilience: earmark capacity for adversarial testing (poisoning/inference drills), maintain a shadow "holdback" site for regression testing of global models, and refresh policies annually. This blend of governance discipline, privacy-first engineering, edge-aware MLOps, and closed-loop planning is the shortest path from promising pilots to durable improvements in defects, yield, on-time delivery, and expedites without ever centralizing sensitive plant or supplier data.

REFERENCES

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308-318. https://doi.org/10.1145/2976749.2978318

Barney, J. B., Ketchen, D. J., Jr., & Wright, M. (2011). The future of resource-based theory: Revitalization or decline? *Journal of Management*, *37*(5), 1299-1315. https://doi.org/10.1177/0149206310391805

Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems*, *101*, 993-1004. https://doi.org/10.1016/j.future.2019.08.017

Bertsimas, D., & Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, *54*(1), 150-168. https://doi.org/10.1287/opre.1060.0285

Blackhurst, J., Craighead, C. W., Elkins, D., & Handfield, R. B. (2005). An empirically derived agenda of critical research issues for managing supply-chain disruptions. *Decision Sciences*, *36*(1), 1-21. https://doi.org/10.1111/j.1540-5915.2005.00116.x

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 1175-1191. https://doi.org/10.1145/3133956.3133982

De Ketelaere, B., Hubert, M., & Schmitt, E. (2015). Overview of PCA-based statistical process monitoring methods for time-dependent, high-dimensional data. *Journal of Quality Technology*, *47*(4), 318-335. https://doi.org/10.1080/00224065.2015.11918154

Dwork, C. (2006). Differential privacy. In *ICALP 2006* (pp. 1-12). https://doi.org/10.1007/11787006_1

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, *46*(4), Article 44. https://doi.org/10.1145/2523813

Ge, Z., & Song, Z. (2014). Semi-supervised PLVR models for process monitoring with unequal sample sizes. *Expert Systems with Applications*, *41*(17), 7705-7716. https://doi.org/10.1016/j.eswa.2014.05.014

Gunasekaran, A., Papadopoulos, T., Dubey, R., Fosso Wamba, S., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, *70*, 308-317. https://doi.org/10.1016/j.jbusres.2016.08.004

Ho, W., Zheng, T., Yildiz, H., & Talluri, S. (2015). Supply chain risk management: A literature review. *International Journal of Production Research*, *53*(16), 5031-5069. https://doi.org/10.1080/00207543.2015.1030467

Ivanov, D. (2018). *Structural dynamics and resilience in supply chain risk management*. https://doi.org/10.1007/978-3-319-69305-7

Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483-1510. https://doi.org/10.1016/j.ymssp.2005.09.012

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven soft sensors in the process industry. *Computers & Chemical Engineering*, *33*(4), 795-814. https://doi.org/10.1016/j.compchemeng.2008.12.012

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148-152. https://doi.org/10.1145/1629175.1629210

Klibi, W., Martel, A., & Guitouni, A. (2010). The design of robust value-creating supply chain networks: A critical review. *International Journal of Production Economics*, *124*(1), 361-380. https://doi.org/10.1016/j.ijpe.2009.09.009

Kourti, T. (2005). Application of PLS in process monitoring and fault diagnosis. *Chemometrics and Intelligent Laboratory Systems*, *81*(2), 89-97. https://doi.org/10.1016/j.chemolab.2004.11.005

Lee, J., Bagheri, B., & Kao, H.-A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, *3*, 18-23. https://doi.org/10.1016/j.mfglet.2014.12.001

Liu, J., & Chen, J. (2015). PLS-based EWMA fault detection strategy for process monitoring. *Journal of Loss Prevention in the Process Industries*, *36*, 404-412. https://doi.org/10.1016/j.jlp.2015.06.002

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *International Journal of Forecasting*, *34*(4), 1430-1452. https://doi.org/10.1016/j.ijforecast.2018.09.003

Md Sanjid, K., & Md. Tahmid Farabe, S. (2021). Federated Learning Architectures For Predictive Quality Control In Distributed Manufacturing Systems. *American Journal of Interdisciplinary Studies*, *2*(02), 01-31. https://doi.org/10.63125/222nwg58

Md. Wahid Zaman, R., & Momena, A. (2021). Systematic Review Of Data Science Applications In Project Coordination And Organizational Transformation. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(2), 01–41. https://doi.org/10.63125/31b8qc62

Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS Quarterly*, *35*(1), 237-256. https://doi.org/10.2307/23043496

Mohassel, P., & Zhang, Y. (2017). *SecureML: A system for scalable privacy-preserving machine learning* 2017 IEEE Symposium on Security and Privacy (SP),

Monostori, L. (2014). Cyber-physical production systems: Roots, expectations and R&D challenges. *CIRP Annals*, *63*(2), 621-641. https://doi.org/10.1016/j.cirp.2014.02.001

Nor, N. M., Hassan, C. R. C., & Hussain, M. A. (2019). A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems. *Reviews in Chemical Engineering*, *35*(6), 703-730. https://doi.org/10.1515/revce-2017-0069

O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. J. (2015). Big data in manufacturing: A systematic mapping study. *Journal of Big Data*, *2*(1), 20. https://doi.org/10.1186/s40537-015-0028-x

Papadopoulos, T., Gunasekaran, A., Dubey, R., & Fosso Wamba, S. (2017). Big data and analytics in operations and supply chain management: Managerial aspects and practical challenges. *Production Planning & Control*, *28*(11-12), 873-876. https://doi.org/10.1080/09537287.2017.1336795

Pavlou, P. A., & El Sawy, O. A. (2011). Understanding the elusive black box of dynamic capabilities. *Decision Sciences*, *42*(1), 239-273. https://doi.org/10.1111/j.1540-5915.2010.00287.x

Qin, S. J. (2012). Survey on data-driven industrial process monitoring and diagnosis: Present and future. *Annual Reviews in Control*, *36*(2), 220-234. https://doi.org/10.1016/j.arcontrol.2012.09.004

Qin, Y., Wang, R., Vakharia, A. J., Chen, Y., & Seref, M. M. H. (2011). The newsvendor problem: Review and directions for future research. *European Journal of Operational Research*, *213*(2), 361-374. https://doi.org/10.1016/j.ejor.2010.07.026

Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, *30*(2), 225-246. https://doi.org/10.2307/25148729

Rony, M. A. (2021). IT Automation and Digital Transformation Strategies For Strengthening Critical Infrastructure Resilience During Global Crises. *International Journal of Business and Economics Insights*, *1*(2), 01-32. https://doi.org/10.63125/8tzzab90

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, *3*(5), 637-646. https://doi.org/10.1109/jiot.2016.2579198

Shokri, R., & Shmatikov, V. (2015). *Privacy-preserving deep learning* Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security,

Sudipto, R., & Md Mesbaul, H. (2021). Machine Learning-Based Process Mining For Anomaly Detection And Quality Assurance In High-Throughput Manufacturing Environments. *Review of Applied Science and Technology*, *6*(1), 01-33. https://doi.org/10.63125/t5dcb097

Syed Zaki, U. (2021). Modeling Geotechnical Soil Loss and Erosion Dynamics For Climate-Resilient Coastal Adaptation. *American Journal of Interdisciplinary Studies*, *2*(04), 01-38. https://doi.org/10.63125/vsfjtt77

Syntetos, A. A., & Boylan, J. E. (2010). On the variance of intermittent demand estimates. *International Journal of Production Economics*, *128*(2), 546-555. https://doi.org/10.1016/j.ijpe.2010.07.004

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84. https://doi.org/10.1111/jbl.12010

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23-45. https://doi.org/10.1080/21693277.2016.1192517

Xu, L. D., He, W., & Li, S. (2014). Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, *10*(4), 2233-2243. https://doi.org/10.1109/tii.2014.2300753

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, *10*(2), Article 12. https://doi.org/10.1145/3298981

Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *Computers & Industrial Engineering*, *137*, 106024. https://doi.org/10.1016/j.cie.2019.106024