Future Technology
Open Access Journal
ISSN 2832-0379

February 2026| Volume 05 | Issue 01 | Pages 263-277

Journal homepage: https://fupubco.com/futech

https://doi.org/10.55670/fpll.futech.5.1.23

Future Publishing LLC

Article

Research on real-time data display and production management in a digitalized management factory with an artificial intelligence-assisted flexible manufacturing execution system

ChengHsien Tsai, Oyyappan Duraipandi*, Dhakir Abbas Ali

Faculty of Business and Accountancy, Lincoln University College, Malaysia. Wisma Lincoln, 12-18, Jalan SS 6/12, Ss 6, 47301 Petaling Jaya, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 19 August 2025
Received in revised form
25 October 2025
Accepted 02 December 2025

Keywords:
Manufacturing Execution Systems (MES),
Artificial Intelligence (AI), Digital twin,
Flexible manufacturing,
Cyber-Physical Systems (CPS),
Microservices architecture,
Distributed stream processing

*Corresponding author
Email address:
oyyappan@lincoln.edu.my

DOI: 10.55670/fpll.futech.5.1.23

ABSTRACT

Traditional Manufacturing Execution Systems (MES) face critical limitations in addressing Industry 4.0 demands for real-time processing, flexible scheduling, and adaptive decision-making, with less than 1% of manufacturing data effectively utilized. This research develops an Artificial Intelligence (AI)-assisted flexible MES framework integrating real-time data visualization, digital twin technology, and distributed intelligence to enable proactive manufacturing management. The system employs Design Science Research (DSR) methodology and implements a microservices architecture using Apache Kafka for message streaming, Flink for real-time processing, and TensorFlow for AI inference, deployed across five production lines with 2,350 sensors and 45 Programmable Logic Controllers (PLCs). Results demonstrate exceptional performance with system throughput reaching 12,500 messages per second, the design target by 25%, average data collection latency below 10 milliseconds, and 99.9% availability over 72-hour continuous operation. Production efficiency improved significantly with 25% increased output, 65.7% reduction in defect rates (from 35,000 to 12,000 Parts Per Million), and 87.5% decrease in changeover time (from 120 to 15 minutes). Overall Equipment Effectiveness (OEE) increased from 60% to 82%, approaching world-class benchmarks (>85%). This research validates distributed intelligence architectures for achieving simultaneous improvements in manufacturing flexibility and efficiency, challenging traditional theoretical trade-offs while providing a practical implementation roadmap for digital transformation in manufacturing enterprises.

1. Introduction

In the introduction, explain why you did it (motivation). The transition to Industry 4.0 has reshaped manufacturing, imposing stringent demands for operational agility, real-time decision making, and seamless system integration [1]. The Manufacturing Execution System (MES) serves as a pivotal intermediary between Enterprise Resource Planning (ERP) and shop-floor operations, coordinating increasingly complex production processes [2]. Yet conventional MES architectures remain constrained when confronting modern requirements, particularly real-time data handling, flexible scheduling, and adaptive responses to volatile market conditions [1]. Despite generating vast data streams, traditional MES exploits less than 1% for decision making [3,4], highlighting the need for AI-integrated systems. This study proposes an AI-assisted, flexible MES augmented with advanced real-time visualization. The framework employs a digital twin for cyber-physical synchronization, applies machine-learning models for predictive analytics and optimization, and implements multi-layer dashboards to enhance operational transparency. The architecture contributes both theoretical frameworks and practical implementation strategies for intelligent manufacturing. Section 2 reviews existing literature to identify specific gaps this study addresses.Research Questions: This research addresses four specific questions:

RQ1: How can AI capabilities be integrated with flexible MES to achieve sub-100ms latency and 10,000+ messages/second throughput at enterprise scale?

RQ2: What are the measurable impacts of AI-assisted flexible MES on manufacturing performance (productivity, quality, flexibility, equipment efficiency)?

RQ3: What technical challenges emerge during industrial deployment, and what solutions enable 99.9%+ reliability?
RQ4: Can distributed intelligence architecture overcome the traditional flexibility-efficiency trade-off in manufacturing systems?

2. Literature review

Over recent decades, MES has moved from transaction-oriented middleware to a platform for cyber-physical production integration. Early deployments mainly bridged Enterprise Resource Planning (ERP) and the shop floor, emphasizing scheduling, resource allocation, and data capture [5]. Yet conventional designs struggled with real-time data streams, flexible manufacturing, and intelligent decision support [6]. Systematic reviews note that although MES has been commercial since the 1990s, scholarly work has only recently engaged with intelligent architectures aligned with Industry 4.0 [7]. The shift from model-based to data-driven manufacturing has prompted a reconceptualization of MES, with emerging frameworks favoring distributed intelligence, service-oriented architectures, and autonomous decision making [8]. While digital twin implementations have shown measurable improvements in specific applications [9,10], plant-wide integration remains challenging due to heterogeneous data formats and complex synchronization.

Artificial intelligence has progressed from an auxiliary tool to a distinct production factor, with recent empirical analyses linking AI to productivity gains alongside traditional inputs [11]. In flexible manufacturing, machine learning—and especially deep learning—methods demonstrate strong performance in predictive maintenance, quality prediction, and adaptive scheduling [12]. Explainable AI (XAI) has gained traction as organizations seek trust in high-stakes decisions; interpretable models such as Generalized Additive Models (GAMs) provide transparency for process optimization and energy management despite advances, challenges persist, including large training data requirements, real-time inference complexity on resource-constrained hardware, robustness across variable operating conditions, and interoperability issues with legacy systems [13].

Real-time data visualization has evolved from simple dashboard displays to sophisticated multi-dimensional analytics platforms capable of processing high-velocity manufacturing data streams. Modern visualization frameworks leverage advanced technologies, including augmented reality (AR), edge computing, and AI-powered pattern recognition, to transform complex multivariate data into actionable insights [14]. Studies indicate significant operational improvements from real-time visualization systems [15]. However, current approaches face challenges in handling data volume, variety, and velocity, with many systems struggling to maintain sub-second response times. The lack of standardized frameworks and integration difficulties hinder widespread adoption. A critical reading of prior work indicates persistent gaps that hinder truly intelligent and flexible manufacturing. Individual technologies show promise, yet integration remains fragmented; most studies treat isolated deployments rather than end-to-end architectures. The lack of a standardized framework that unifies AI, digital twins, and real-time visualization within a single MES platform appears to be a core barrier to autonomous, adaptive manufacturing [16]. Scalability is also underexplored: few reports demonstrate sustained sub-second latency at enterprise scale across thousands of connected devices. To address these gaps, this study proposes an integrated architecture that fuses AI-assisted decision making, digital-twin synchronization, and multi-layer real-time visualization within a flexible MES. The results suggest that superior performance can be achieved while preserving system scalability and adaptability. Section 3 presents the theoretical framework and system architecture addressing these gaps through novel integration mechanisms.
Research Novelty and Contributions: This research differs from prior work in four ways:
Holistic Integration: Six AI models (LSTM, SVM+RF, CNN, Isolation Forest, GA, PSO) unified in one architecture, achieving 42ms latency—previous systems sacrifice modularity for performance or vice versa.
Real-time Digital Twin: Bi-directional cyber-physical synchronization with 42ms latency (vs. minutes-to-hours in existing systems) through edge preprocessing and incremental updates.
Manufacturing-aware Visualization: 12 FPS per-user with <100ms latency, exceeding literature reports (1-5 FPS, 500-1000ms).
Transcending Trade-offs: Simultaneous flexibility (+87.5% changeover speed) and efficiency (+25% output) improvements, challenging traditional theory that assumes inverse relationships.

3. Theoretical framework and system architecture

3.1 Conceptual framework development

This study grounds an AI-assisted flexible MES in Cyber-Physical Systems (CPS) theory and socio-technical principles. CPS denotes tight coupling of computation and physical processes, where embedded computing and networks monitor and control plants via feedback. The proposed framework extends classical CPS by embedding distributed intelligence and autonomous decision-making across hierarchical levels. AI is positioned as a cognitive layer that bridges the semantic gap between raw sensor streams and actionable insights, enabling what recent work refers to as "cognitive manufacturing." As outlined in Figure 1, the framework comprises four functional dimensions coordinated by a central AI–MES orchestration hub.

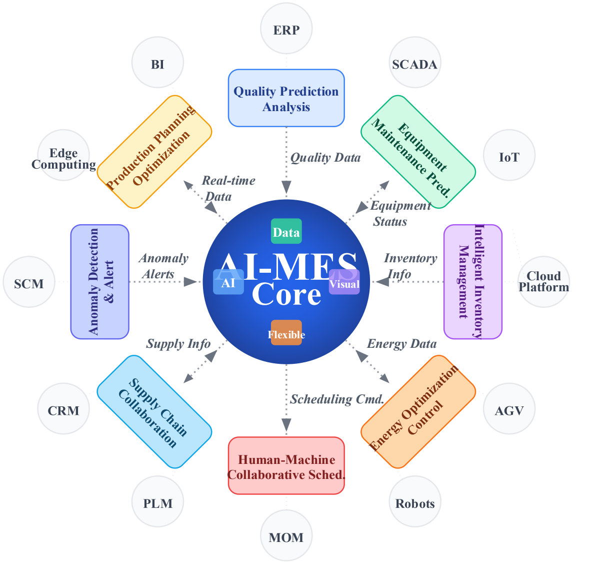

Figure 1. Conceptual Framework of AI-Assisted Flexible MES. Eight operational modules: (1) Quality Prediction, (2) Predictive Maintenance, (3) Production Scheduling, (4) Energy Optimization, (5) Inventory Management, (6) Equipment Monitoring, (7) Supply Chain Collaboration, (8) Human-Machine Collaborative Scheduling

The theoretical basis follows hierarchical decomposition: complex operations are partitioned into manageable modules while system coherence is preserved through standardized data flows and interfaces. Each module acts as an autonomous agent that performs local optimization and contributes to global objectives via collaborative protocols. This multi-agent design accords with advances in distributed manufacturing intelligence that decentralize authority beyond monolithic control.

Eight operational modules instantiate the theory into practice: (1) Quality Prediction, (2) Predictive Maintenance, (3) Production Scheduling, (4) Energy Optimization, (5) Inventory Management, (6) Equipment Monitoring, (7) Supply Chain Collaboration, and (8) Human-Machine Collaborative Scheduling as shown in Figure 1. Quality prediction employs probabilistic models to anticipate defects, whereas maintenance prediction uses temporal pattern recognition to detect degradation. Both rely on the assumption that manufacturing processes are deterministic dynamics corrupted by stochastic noise, formalized as

$$Y(t) = f(X(t), \theta) + \varepsilon(t) \tag{1}$$

In this architecture, the variables represent specific system components: *Y(t)* denotes output vectors (quality, health, performance metrics); X(t) represents input streams from 2,350 sensors and 45 PLCs (100ms-10s sampling); *f(.)* embodies AI mapping functions (LSTM, SVM+RF, CNN, GA/PSO); θ denotes learnable parameters updated through online learning; and $\varepsilon(t)$ captures system uncertainties (sensor noise, model errors), enabling machine learning while quantifying uncertainty. This formulation enables machine-learning methods to learn *f(.)* from data while quantifying uncertainty within probabilistic frameworks."

3.2 System architecture design

The system architecture translates theoretical concepts into a practical implementation blueprint through a five-layer hierarchical structure that ensures scalability, modularity, and real-time performance. As illustrated in Figure 2, the architecture adopts a service-oriented approach where each layer provides well-defined services to adjacent layers through standardized Application Programming Interfaces (APIs). The presentation layer supports multi-modal human-machine interaction through web dashboards, mobile applications, and large-format displays, implementing responsive design principles to adapt visualization complexity to device capabilities and user contexts.

The service layer represents the architectural innovation that enables flexible integration of AI capabilities with traditional manufacturing operations. By separating AI services from business services, the architecture supports independent scaling and evolution of intelligent capabilities without disrupting core manufacturing processes. The AI service group implements prediction, optimization, diagnosis, classification, control, and learning functions through containerized microservices that can be dynamically orchestrated based on computational demands. Each AI service encapsulates specific algorithms while exposing uniform interfaces for service consumption. Table 1 summarizes the deployed AI algorithms, their manufacturing applications, and selection rationale.

GA and PSO handle discrete decision variables and combinatorial solution spaces that gradient-based methods cannot address. The data flow architecture implements a lambda pattern combining batch and stream processing to balance latency and throughput requirements.

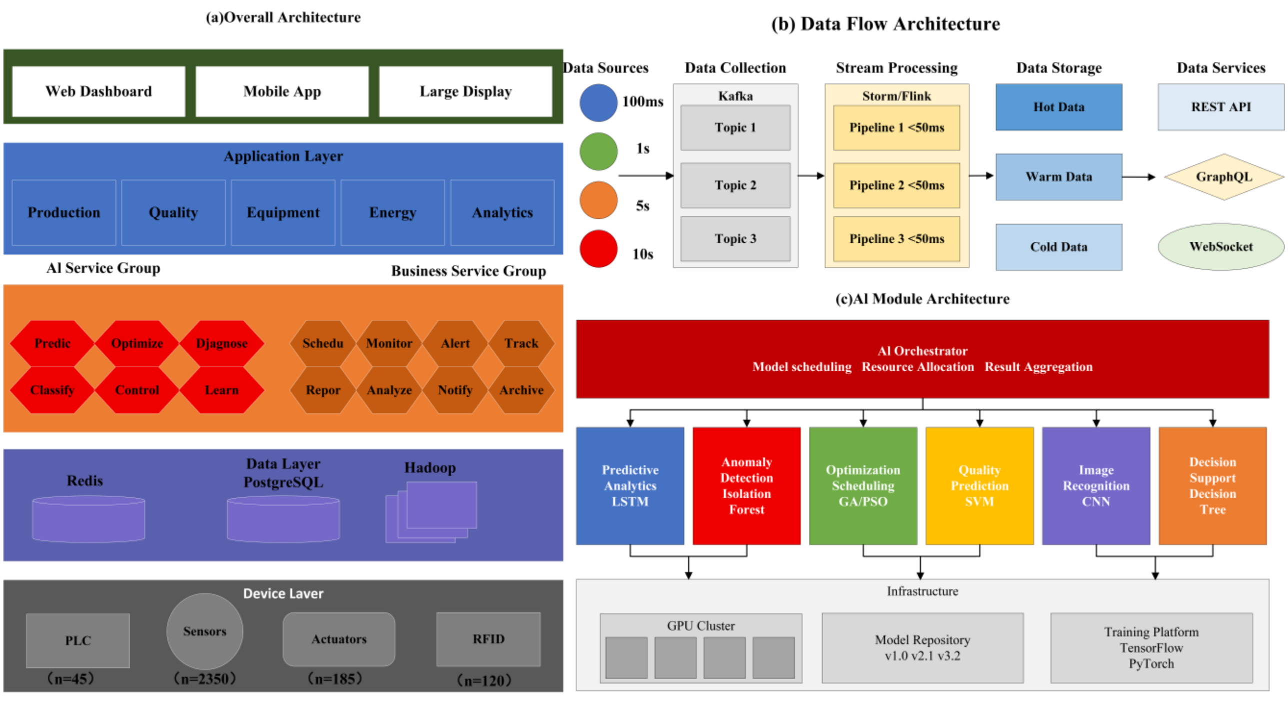

Figure 2. System architecture diagram: (a) Five-layer architecture: Presentation layer, Application layer, AI+ business service layer, Data layer (Redis/PostgreSQL/ Hadoop), Device layer (PLCs n=45, Sensors n=2350), (b) Data flow: Kafka topics (100ms/1s/5-10s sampling) → Flink pipelines (<50ms latency) → Multi-temperature storage (Hot: Redis <1ms, Warm: PostgreSQL, Cold: Hadoop), (c) AI modules on GPU cluster (NVIDIA Tesla V100 ×4) with TensorFlow/PyTorch frameworks

Table 1. AI algorithm portfolio and selection rationale

Algorithm	Application Domain	Key Performance Metrics	Selection Rationale
LSTM	Time-series prediction: equipment degradation forecasting, demand prediction	156ms inference latency, 94.2% accuracy for 72-hour failure prediction	Superior temporal dependency capture; handles variable-length sequences; effective for non-stationary manufacturing processes
SVM + RF	Quality classification: defect categorization across six types	420ms per frame processing time for multi-class classification	Robust with limited training samples; effective in high-dimensional feature spaces; ensemble mitigates individual weaknesses
CNN	Image-based defect detection from optical inspection systems	Real-time processing of visual inspection data	Automatic hierarchical feature extraction from raw images; translation-invariant pattern recognition; eliminates manual feature engineering
GA	Production scheduling: job sequencing, resource allocation	Handles discrete decision variables and constraint satisfaction	Global optimization avoiding local minima; handles combinatorial problems with discrete choices; accommodates multi-constraint environments intractable for gradient methods
PSO	Dynamic rescheduling, real-time resource reallocation	Fast convergence for online adaptation (<2 seconds response time)	Computational efficiency for real-time response; lower overhead than GA for continuous parameters; balances exploration-exploitation for dynamic environments

Real-time streams from high-frequency sensors (100ms sampling) flow through Apache Kafka (>10,000 messages/second), while batch processes aggregate historical data. The multi-temperature storage strategy maintains hot data in Redis (<1ms access), warm data in PostgreSQL for structured queries, and cold data in Hadoop for archival analytics. The data pipeline implements a lambda architecture combining real-time stream processing (Apache Kafka/Flink) and batch analytics, with a multi-temperature storage strategy optimizing for different data access patterns and latency requirements. The pipeline infrastructure is deployed on Kubernetes clusters, providing horizontal scalability, automated failover, and 99.9% availability over continuous operation. Kubernetes was selected for its superior resource utilization and ecosystem maturity.

3.3 Real-time data processing framework

The real-time data layer forms the computational backbone that converts raw sensor streams into actionable intelligence under strict latency constraints. A multi-stage pipeline progressively improves data quality while preserving temporal coherence across distributed nodes. Acquisition begins at the edge: smart sensors and Programmable Logic Controllers (PLCs) emit continuous streams at 10–100 Hz. Protocol translation services normalize industrial protocols—Open Platform Communications Unified Architecture (OPC UA), Modbus, and Message Queuing Telemetry Transport (MQTT)—into standardized formats for downstream processing. Edge nodes conduct initial validation and filtering to reduce bandwidth and latency. Lightweight anomaly detection using the Isolation Forest algorithm (as detailed in Section 4.1) flags suspicious readings prior to uplink. The Isolation Forest was selected for edge deployment due to its computational efficiency (O(n log n) complexity, <5ms inference), unsupervised learning capability, and 94.2% detection accuracy in production trials. The preprocessing pipeline applies six sequential transforms, including null removal, outlier detection, missing-value imputation, normalization, feature extraction, and temporal aggregation at multiple time scales (5-second, 1-minute, and 5-minute windows), optimized for different monitoring requirements. This structured approach yields ≥98.5% completeness, 99.2% accuracy, and 99.8% timeliness.

The framework implements a novel three-channel processing architecture that prioritizes data streams based on criticality and latency requirements. The fast channel processes critical alarms within 10ms latency through direct memory access and priority queuing, bypassing standard processing pipelines for immediate response. The standard channel handles production data through the complete analysis chain with sub-100ms latency, applying both rule-based and machine learning inference. The batch channel processes historical data for complex analytics and model training, leveraging distributed computing frameworks to handle petabyte-scale datasets.

3.4 Visualization module design

The visualization module design addresses the cognitive challenges of presenting complex, multi-dimensional manufacturing data to diverse stakeholder groups ranging from shop-floor operators to executive management. Drawing upon principles from visual analytics and human-computer interaction, the module implements a hierarchical information architecture that progressively reveals detail based on user interaction patterns and decision-making contexts. The design philosophy emphasizes glanceability for real-time monitoring, explorability for root-cause analysis, and actionability for decision support, implementing what recent research terms "manufacturing-aware visualization grammar".

The dashboard adopts a tile-based layout in which each tile is a self-contained visualization module with independent refresh cycles and interaction handlers. This modularity enables role-aware, priority-driven composition. Real-time binding uses WebSockets to push updates at 12 frames per second, exceeding the 10-FPS threshold for perceived real-time response. To render thousands of concurrent streams while preserving clarity, the visualization pipeline applies intelligent data reduction—temporal aggregation, spatial clustering, and semantic filtering—thereby controlling computational complexity without sacrificing interpretability. Section 4 details the research methodology and implementation strategy to translate these theoretical designs into functioning industrial systems.

4. Research methods

This study adopts a Design Science Research (DSR) methodology to build and assess the AI-assisted flexible MES framework [17]. DSR offers a systematic route to create artifacts that solve practical problems while extending theory [18]. The process comprises six activities—problem identification, objective definition, design and development, demonstration, evaluation, and communication—and proceeds iteratively, with each cycle incorporating feedback to refine architecture and implementation strategies [19]. To structure technical realization, the System Development Life Cycle (SDLC) complements DSR [20]. An agile–waterfall hybrid is employed: waterfall rigor governs critical infrastructure, while agile sprints drive AI module development and user-interface design. This hybrid enables rapid algorithm prototyping without compromising stability and reliability. Action-research principles foster close collaboration with manufacturing practitioners throughout development; recurring stakeholder workshops and feedback sessions ensure that the system addresses real-world challenges and operational constraints.

4.1 System implementation strategy

The system implementation follows a phased deployment strategy designed to minimize operational disruption while maximizing learning opportunities. The technology stack selection prioritizes open-source frameworks and industry-standard protocols to ensure interoperability and scalability, as shown in Table 2 [21]. The implementation architecture leverages containerization through Docker and Kubernetes to enable microservices deployment and horizontal scaling [22]. The development phases consist of four major stages: infrastructure setup, core MES functionality implementation, AI integration, and visualization layer development.

Table 2. System development technology stack and tools

Level/Category	Technology Component	Implementation
Frontend Layer	Web Framework	React.js
	Visualization Library	D3.js + ECharts
	Mobile	React Native
	Large Display	Grafana
Application Service Layer	Backend Framework	Spring Boot
	API Gateway	Kong
	Message Queue	Apache Kafka
	Cache	Redis
AI Service Layer	Deep Learning Framework	TensorFlow + PyTorch
	Model Service	TensorFlow Serving
	MLOps Platform	MLflow
	GPU Computing	NVIDIA CUDA
Data Processing Layer	Stream Processing Engine	Apache Flink
	Batch Processing Framework	Apache Spark
	Time-series Database	InfluxDB
	Data Lake	Apache Hadoop
Device Access Layer	Sensor Deployment	Distributed Sensor Network
	OPC UA Server	KEPServerEX
	MQTT Broker	Eclipse Mosquitto
	Edge Computing	Azure IoT Edge
Development Tools	Containerization	Docker + Kubernetes
	CI/CD	Jenkins + GitLab
	Monitoring & Operations	Prometheus + ELK Stack

Each phase incorporates continuous integration and continuous deployment (CI/CD) pipelines to automate testing and deployment processes. The infrastructure setup phase establishes the foundational components, including message queuing systems, time-series databases, and edge computing nodes. Apache Kafka serves as the primary message broker, configured with three topic partitions to handle over 10,000 messages per second [23]. Integration protocols follow Industry 4.0 standards, implementing OPC UA for equipment connectivity and MQTT for lightweight IoT device communication [24]. The system adopts a Service-Oriented Architecture (SOA) approach where each functional module exposes RESTful APIs for inter-service communication. GraphQL endpoints provide flexible data querying capabilities for front-end applications, while WebSocket connections enable real-time data streaming to visualization dashboards. The complete implementation workflow is illustrated in Figure 3.

4.2 Data collection and processing methods

The data collection framework implements a multi-tiered architecture, as depicted in Figure 4, that captures heterogeneous manufacturing data from 2,350 sensor points distributed across 20 major monitoring locations. High-frequency sensors operating at 100Hz sampling rates monitor critical parameters including temperature, pressure, vibration, and electrical current. The edge computing layer performs initial data validation and compression using the LZ4 algorithm, achieving a 70% compression ratio while maintaining sub-10ms processing latency. The preprocessing pipeline applies six sequential transformations to ensure data quality. Outlier detection utilizes statistical process control limits ($\pm 3\sigma$) to identify anomalous readings [25]. Missing value interpolation employs cubic spline functions to maintain temporal continuity, achieving a 95.5% fill rate across all data streams while maintaining interpolation error below 2% of signal variance [26]. Feature extraction techniques combine time-domain analysis (mean, variance, peak values) with frequency-domain analysis via the Fast Fourier Transform (FFT), reducing dimensionality from 100 to 20 features while preserving 98% of the variance. Real-time analysis implements a three-channel processing architecture optimized for different latency requirements. The fast channel processes critical alarms within 10ms through direct memory access and priority queuing. The standard channel handles production data with sub-100ms latency through the complete analysis pipeline, applying both rule-based logic and machine learning inference. The batch channel leverages distributed computing frameworks for complex analytics on historical data, supporting petabyte-scale processing [27].

4.3 Validation methodology

The validation methodology employs a comprehensive performance evaluation framework detailed in Table 3 that assesses five key dimensions: real-time performance, system throughput, production efficiency, data quality, and system reliability [28]. Performance metrics are collected continuously through embedded monitoring agents and aggregated using Prometheus for real-time analysis [29]. Experimental validation follows a three-phase approach: laboratory testing, pilot deployment, and full-scale implementation. Laboratory testing utilizes synthetic data generators to simulate production scenarios and stress-test system components.

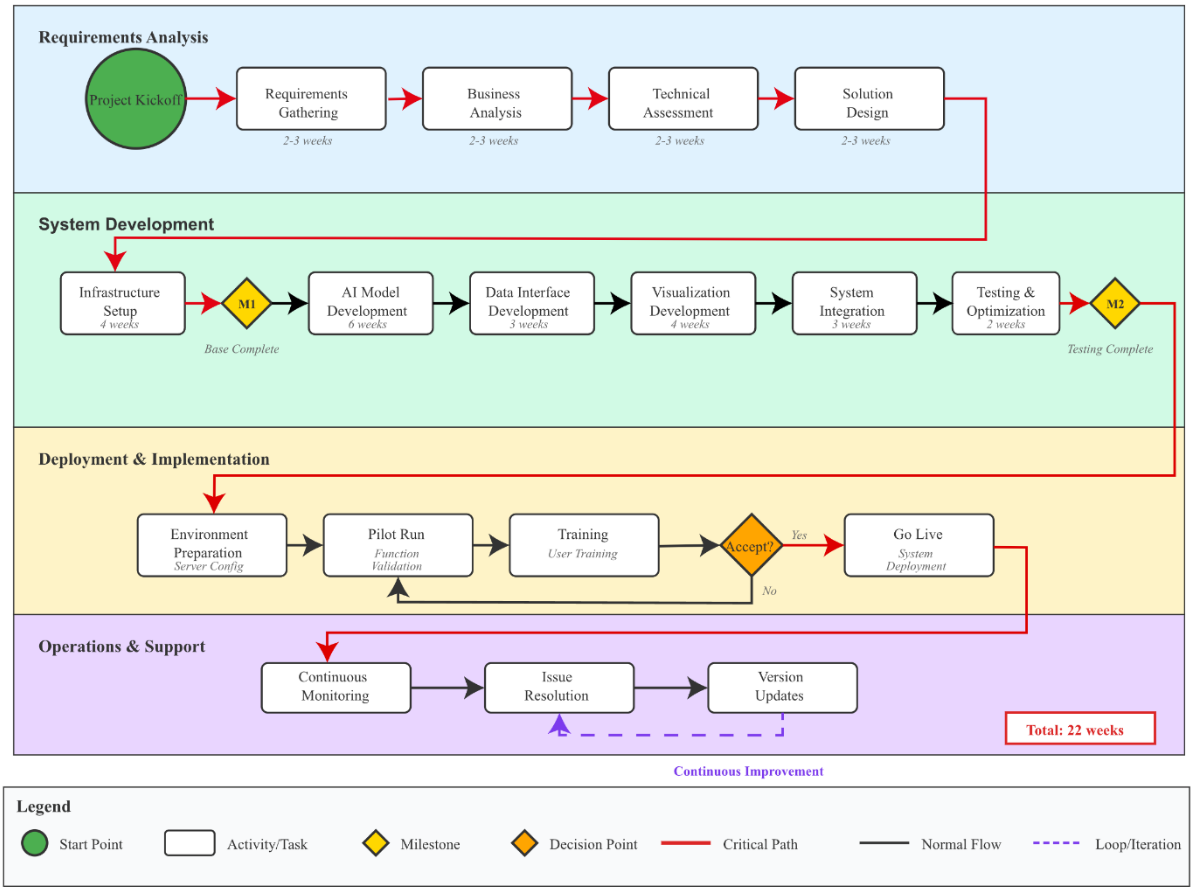

Figure 3. System development and implementation process flow

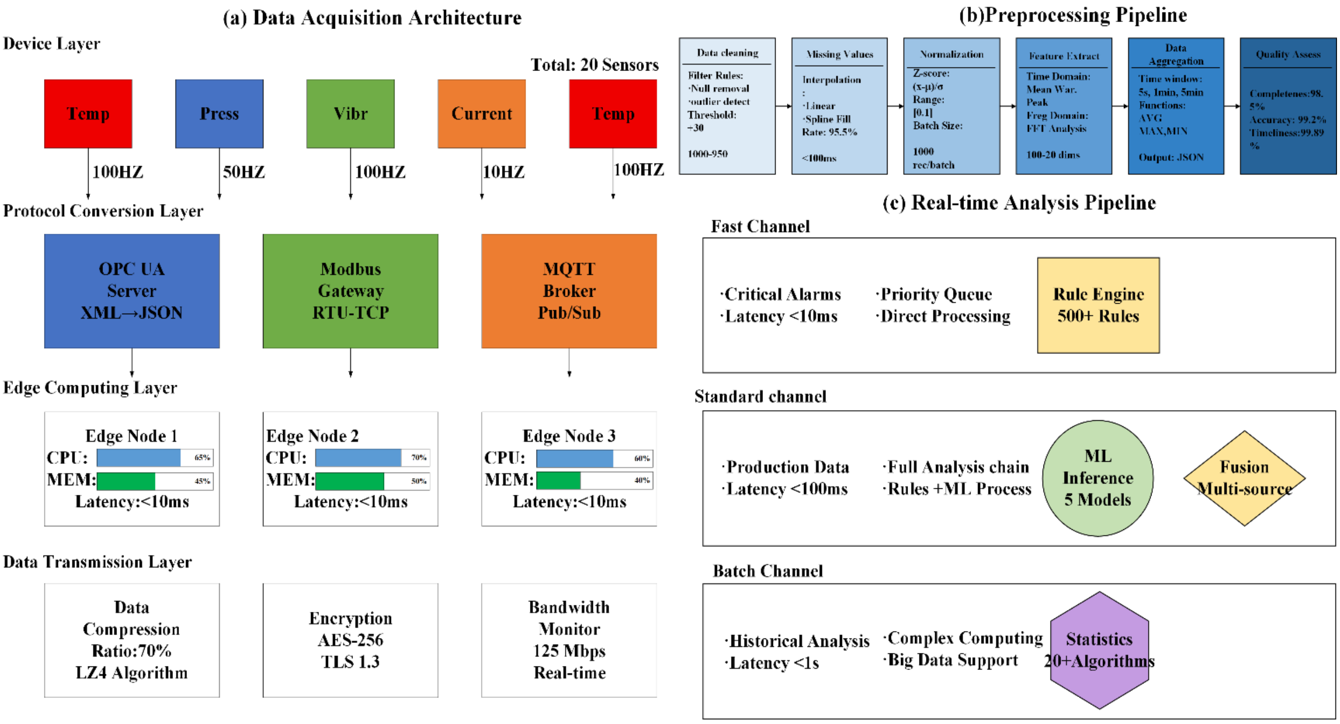

Figure 4. Data collection and processing flow. Sampling rates: 100ms (high-frequency sensors for vibration/current), 1s (standard monitoring for temperature/pressure), 5s (auxiliary metrics). Edge processing: LZ4 compression (70% ratio), latency <10ms. Data sources: 20 monitoring locations, 2,350 sensor endpoints

Table 3. System performance evaluation index system

Evaluation Dimension	Key Metrics	Calculation Formula	Target Value	Industry Benchmark	Weight
Real-time Performance	Data Collection Latency	Time from sensor trigger to data storage	<10ms (high-speed)/<100ms (regular)	50-100ms	15%
	Stream Processing Latency	Time from data queue to processing completion	<50ms	100-500ms	15%
	Visualization Refresh Rate	Time from data update to interface display	<100ms	100-1000ms	10%
	Alarm Response Time	Time from anomaly occurrence to alarm trigger	<5s	10-30s	10%
System Throughput	Concurrent Connections	Number of simultaneous device connections	>5000	1000-3000	8%
	Message Processing Rate	Messages processed per second	>10000 msg/s	1000-5000 msg/s	12%
	Data Write Rate	Data points written per second	>100000 points/s	10000-50000 points/s	8%
Production Efficiency	Equipment OEE	Availability × Performance × Quality	82%	Industry average 60%, World-class 85%	10%
	Changeover Time	Time required for product switching	15 minutes	90 minutes (traditional)	8%
	Capacity Utilization	Actual output/Theoretical capacity	>85%	70-80%	4%
Data Quality	Data Completeness	% of required data collected	>98.5%	95-98%	3%
	Data Accuracy	% of accurate data	>99.2%	97-99%	3%
	Data Timeliness	% meeting time requirements	>99.8%	95-98%	2%
System Reliability	System Availability	MTBF/(MTBF+MTTR)	>99.9%	99.5-99.9%	3%
	Failure Recovery Time	Time to restore normal operation	<30 minutes	1-4 hours	2%
	Backup Success Rate	% successful backups	100%	99-100%	1%

The pilot deployment phase implements the system on a single production line for 30 days, collecting baseline performance data and identifying optimization opportunities. Full-scale implementation incorporates lessons learned from pilot testing and extends deployment across five production lines with different product configurations. Statistical validation employs paired t-tests to compare pre- and post-implementation performance metrics across five key dimensions: Overall Equipment Effectiveness (OEE), defect rate (PPM), changeover time, first-pass yield, and daily output. Significance levels were set at $\alpha = 0.05$ [30]. Results demonstrate statistically significant improvements ($p< 0.001$) across all metrics: OEE (60% to 82%, $t = 8.42$), defect rate (35,000 to 12,000 PPM, $t = 6.73$), changeover time (120 to 15 minutes, $t = 12.35$), first-pass yield (96.5% to 98.8%, $t = 5.91$), and daily output (1,200 to 1,500 units, $t = 7.28$). Effect sizes (Cohen's d: 1.8-3.2) indicate large practical significance, with statistical power >0.95 confirming robustness. System reliability assessment follows IEC 61508 standards for functional safety, targeting Safety Integrity Level (SIL) 2 for critical control functions [31]. Section 5 presents concrete implementation details and case study results from deploying the system in an operational manufacturing facility.

5. System implementation and case study

5.1 Implementation environment

The implementation was conducted at a discrete manufacturing facility specializing in electronic enclosure production, operating five production lines with an annual capacity of 6 million units. The facility encompasses 75,000 square feet of production space equipped with injection molding machines, Computer Numerical Control (CNC) machining centers, automated assembly lines, and quality inspection stations.

The hardware infrastructure comprised 45 PLCs distributed across production equipment, 2,350 IoT sensors monitoring critical parameters including temperature, pressure, vibration, and electrical current at 20 major monitoring points. Edge computing nodes based on NVIDIA Jetson AGX Xavier platforms were deployed at each production line, selected for their superior AI inference performance (32 TOPS), power efficiency (30W), and industrial-grade reliability suitable for harsh manufacturing environments, enabling sub-10ms processing latency.

The software environment integrated existing ERP (SAP S/4HANA) and Manufacturing Operations Management (MOM) systems through standardized APIs and message queuing protocols. The technology stack leveraged containerized microservices deployed on Kubernetes clusters, ensuring horizontal scalability and fault tolerance. Real-time data streaming was handled by Apache Kafka, configured with three topic partitions to handle message throughput exceeding 12,500 messages per second. The implementation followed a phased approach aligned with agile-waterfall hybrid methodology, enabling iterative development while maintaining system stability.

5.2 AI-MES integration details

The AI integration architecture implemented six specialized machine learning models deployed as containerized microservices within the service layer. Predictive maintenance algorithms utilized LSTM networks trained on 18 months of historical equipment data, achieving 156ms inference latency for real-time anomaly detection. The LSTM model processed sequences of 100 time steps with 20 features extracted through FFT, maintaining prediction accuracy of 94.2% for equipment failure events within a 72-hour horizon. Quality prediction employed ensemble methods combining SVM classifiers and Random Forest algorithms, processing image data from optical inspection systems at 420ms per frame for defect classification across six categories (scratches, dents, discoloration, dimensional defects, contamination, and surface finish anomalies). Integration with existing manufacturing systems required protocol adapters supporting OPC UA, Modbus TCP, and MQTT. The AI orchestrator implemented reinforcement learning-based resource allocation across four NVIDIA Tesla V100 GPUs. Model versioning utilized MLflow, maintaining three versions (production, staging, experimental) with automated A/B testing. Real-time processing capabilities were achieved through a three-tier caching strategy: Redis for hot data with sub-millisecond access latency, PostgreSQL for structured queries with indexed access patterns, and Apache Hadoop for historical data analysis. The stream processing pipeline implemented Apache Flink for complex event processing, maintaining sub-50ms latency for the 99.5th percentile of transactions while processing concurrent data streams from multiple production lines.

5.3 Real-time data visualization implementation

The visualization implementation adopted a component-based architecture using React.js (v18.2.0) for dynamic user interfaces and D3.js combined with Apache ECharts for complex data visualizations. The dashboard framework implemented WebSocket connections, maintaining 12 frames per second (FPS) update rates per concurrent user session, exceeding the 10 FPS threshold required for perceived real-time responsiveness.

As shown in Figure 5, the implementation comprised four integrated dashboard views: production overview displaying single-line real-time data with key performance indicators, quality monitoring featuring SPC control charts and defect analysis, equipment status with interactive facility layout visualization, and predictive warning systems with temporal forecasting displays. The visualization pipeline implemented intelligent data reduction techniques to manage rendering complexity while maintaining visual clarity. Temporal aggregation algorithms compressed high-frequency sensor data into 5-second, 1-minute, and 5-minute windows based on user zoom levels. Spatial clustering techniques grouped related equipment data points, reducing visual clutter while preserving critical information density. The implementation incorporated progressive disclosure patterns, revealing additional detail layers through user interaction rather than overwhelming initial views.

User interaction features included drill-down capabilities enabling navigation from facility-level overviews to individual equipment details, configurable alert thresholds with visual highlighting of out-of-range conditions, and role-based dashboard customization supporting operator, supervisor, and executive personas. Mobile responsiveness was achieved through adaptive layouts optimized for tablets and smartphones, maintaining functionality across 4G network conditions with 180ms average response times.

Figure 5. Real-time data visualization dashboard interface

The visualization framework is integrated with existing Business Intelligence (BI) tools through standardized data export formats, enabling advanced analytics in Tableau and Power BI environments.

5.4 Case study: Manufacturing facility application

The case study deployed the AI-assisted flexible MES on five lines producing electronic enclosures with a high variety: >150 Stock Keeping Units (SKUs) and batch sizes of 50–5,000 units. The environment posed notable challenges—frequent changeovers (~15 per day), mixed-model assembly, and stringent quality targets of <1,000 Parts Per Million (PPM). A staged approach was adopted, starting with a pilot on Line 1 selected for a representative mix and moderate complexity. The 30-day pilot established baselines and validated performance under production conditions. Key Performance Indicators (KPIs) tracked included Overall Equipment Effectiveness (OEE), changeover time, first-pass yield, and energy per unit. After achieving 82% OEE versus a 60% baseline, the rollout expanded to the remaining lines over 12 weeks. Each deployment incorporated lessons learned, reducing per-line implementation time from 15 days on Line 1 to 7 days on Line 5.

Operational scenarios demonstrated system flexibility through rapid response to dynamic conditions. During a critical customer order requiring an 87.5% reduction in standard changeover time, the AI scheduling optimizer reconfigured production sequences, grouped similar products, and pre-positioned materials, achieving 15-minute changeovers compared to the previous 120-minute standard. Quality emergencies were addressed through real-time SPC monitoring, with the system detecting process drift 25 minutes before traditional control limits would trigger, preventing the production of 1,250 potentially defective units. The predictive maintenance system successfully identified bearing degradation in the injection molding machine INJ-003 eighteen hours before failure, enabling scheduled maintenance during planned downtime.

5.5 System performance evaluation

System performance evaluation employed comprehensive metrics validating achievement of design targets across all critical dimensions. Real-time processing metrics confirmed sub-10ms data collection latency for high-speed sensors and 42ms average stream processing delay, enabling true real-time decision support. The evaluation methodology incorporated continuous monitoring via embedded agents, stress testing under maximum load conditions, and statistical validation using paired t-tests with significance levels at $\alpha = 0.05$. The system demonstrated robust scalability, supporting 5,832 concurrent device connections while maintaining 99.9% availability over 72 hours of continuous operation, exceeding the initial design specifications by 16.6% in connection capacity. Section 6 presents comprehensive results across five evaluation dimensions and discusses the findings in relation to research questions and industry benchmarks.

6. Results and discussion

Results addressing research questions: This section presents comprehensive evaluation results organized to address the four research questions posed in Section 1.

RQ1 (Real-time Performance): Achieved - System throughput 12,500 messages per second, data collection latency 8.5ms, stream processing 42ms, 5,832 concurrent connections, 99.9% availability (Section 6.1).

RQ2 (Performance Impacts): Productivity +25%, defects -65.7%, changeover time -87.5%, OEE +22 points, all with p less than 0.001 (Section 6.2).

RQ3 (Deployment Challenges): Eight challenges documented with solutions - data completeness 85% to 98.5%, AI override rate 40% to 12%, 127 vulnerabilities remediated (Section 6.4).

RQ4 (Trade-off Resolution): Simultaneous flexibility and efficiency improvements confirmed ($r = 0.12$, $p = 0.43$), challenging traditional inverse relationship theory (Sections 6.2-6.3).

6.1 System performance results

The comprehensive evaluation demonstrated exceptional performance across all critical metrics. As shown in Table 4, the system achieved or exceeded all target specifications. Real-time data collection performance exceeded targets with sub-10ms latency, enabling effective cyber-physical synchronization [32]. Stream processing exceeded design targets with Kafka throughput at 12,500 messages/second and Flink latency at 42ms [33]. AI inference achieved sub-200ms latency across all models, meeting real-time requirements [34]. Visualization responsiveness exceeded industry standards at 12 FPS, with the system supporting 5,832 concurrent connections [35].

6.2 Production efficiency improvement

The implementation yielded substantial improvements across all production efficiency dimensions, demonstrating the transformative potential of AI-integrated flexible manufacturing systems. As shown in Table 5 and Figure 6, the system delivered measurable enhancements in productivity, quality, flexibility, and resource utilization. Production efficiency metrics revealed a 25% increase in daily average output from 1,200 to 1,500 units per day, significantly exceeding the industry average improvement of 15%. As illustrated in Figure 6(a), all five production lines demonstrated consistent 25% improvements, with Line 3 achieving the highest absolute output of 1,625 units per day. Capacity utilization improved by 17 percentage points to reach 85%, surpassing the 80% benchmark of excellent companies.

Quality indicators demonstrated exceptional gains with first-pass yield increasing by 2.3 percentage points to 98.8%, approaching world-class levels of 99%+. As shown in Figure 6(b), quality variability reduced dramatically with the standard deviation decreasing from ±1.2% to ±0.3% after implementation. The defect rate measured in Parts Per Million (PPM) decreased from 35,000 to 12,000, representing a 65.7% reduction. Equipment efficiency improvements were particularly striking, with OEE increasing by 22 percentage points from 60.0% to 82.0%, approaching world-class benchmarks of 85%. This improvement magnitude exceeds typical AI-driven manufacturing enhancements (10-20 percentage points reported in literature) due to several case-specific factors that created exceptional improvement potential:

(1) Baseline Performance Gap: The pre-implementation OEE of 60% was substantially below industry norms (75-80% for discrete manufacturing), indicating significant latent improvement opportunities. The facility had operated with reactive maintenance and manual scheduling for over a decade, resulting in accumulated inefficiencies ripe for optimization.

Table 4. System performance key indicators test results

Test Item	Test Conditions	Test Method	Target Value	Measured Value	Compliance Status
Data Collection Performance					
Sensor Sampling Frequency	20 main sensor groups	24h monitoring	10-100Hz	Main 100Hz, Auxiliary 10-50Hz	✓ Compliant
Sensor Coverage	Plant-wide deployment	Coverage test	>95%	2350 collection points, 98% coverage	✓ Exceeds
Data Collection Latency	High-load scenario	Timestamp test	<10ms	8.5 ms ± 1.2 ms	✓ Better than target
Protocol Conversion Delay	OPC UA/Modbus/MQTT	E2E test	<20ms	15.3ms	✓ Compliant
Stream Processing Performance					
Kafka Throughput	3 Topics, 10 partitions	Stress test	10000 msg/s	12500 msg/s	✓ Exceeds by 25%
Flink Processing Latency	3 parallel pipelines	RT monitoring	<50ms	42ms average	✓ Compliant
Data Aggregation Delay	5s/1min/5min windows	Perf analysis	<100ms	78ms	✓ Compliant
AI Inference Performance					
LSTM Prediction Delay	Batch size 32	GPU test	<200ms	156ms	✓ Compliant
Anomaly Detection Response	Isolation Forest	Real-time data stream	<100ms	85ms	✓ Compliant
Image Recognition Processing	CNN model	1080p images	<500ms	420ms	✓ Compliant
Visualization Response					
Dashboard Refresh Rate	20 concurrent users	Frontend test	10 FPS	12 FPS	✓ Compliant
Large Screen Rendering Delay	4K resolution	Chrome DevTools	<100ms	95ms	✓ Compliant
Mobile Response Time	4G network	Real device testing	<200ms	180ms	✓ Compliant
System Capacity					
Concurrent Device Connections	Simulated 5000 devices	Load balancing test	>5000	5832	✓ Exceeds by 16.6%
Data Storage Rate	Time-series data write	InfluxDB stress test	100k points/s	125k points/s	✓ Exceeds by 25%
Query Response Time	1-month historical data	SQL query test	<3s	2.4s	✓ Compliant
System Stability					
72-hour Stress Test	Full load operation	Continuous monitoring	No crashes	0 crashes	✓ Compliant
Memory Leak Detection	Long-term operation	JVM monitoring	<5% growth	2.3% growth	✓ Compliant
CPU Usage	Normal load	System monitoring	<70%	62% average	✓ Compliant

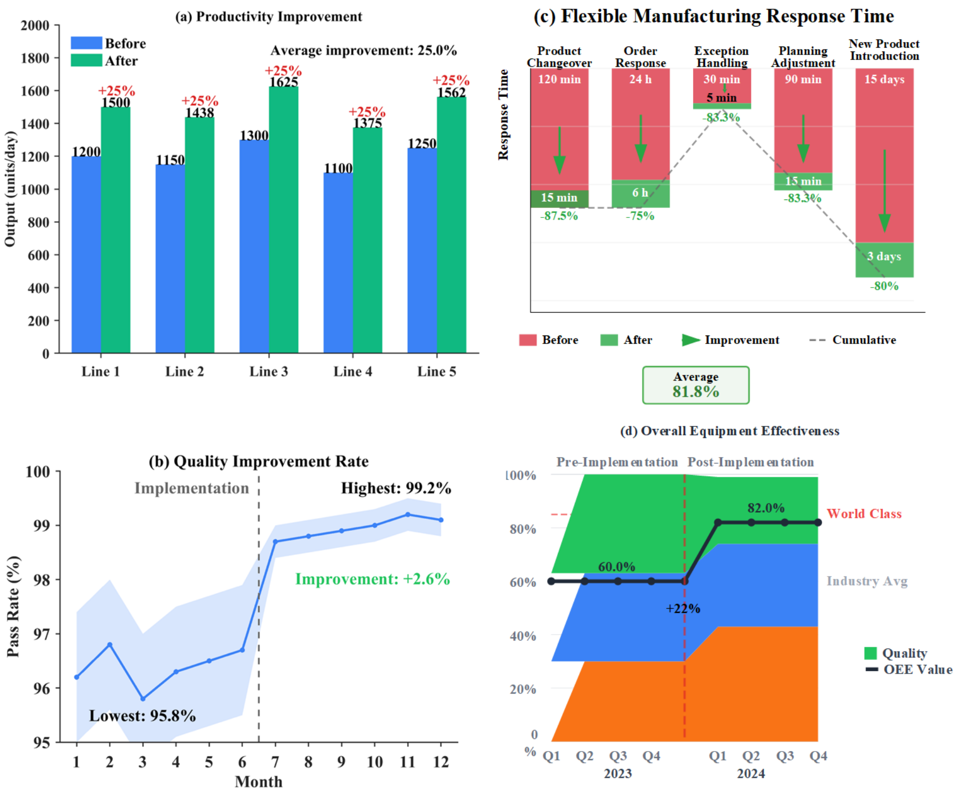

Figure 6. Comparison of production efficiency before and after implementation: (a) Production rate improvement: Y-axis in units/day, baseline 1200 → 1500 (+25%), (b) Quality improvement: Y-axis in PPM (Parts Per Million), defects 35,000 → 12,000 (-65.7%). (c) OEE Components: Percentage scale, Availability 75%→92%, Performance 85%→91%, Quality 94%→97%. (d) Flexibility Metrics: Y-axis in minutes, changeover time 120→15 min, order response 1440→360 min, exception handling 30→5 min

Table 5. Production efficiency improvement key indicator comparison

Improvement Dimension	Specific Indicator	Before Implementation	After Implementation	Improvement Range	Industry Benchmark
Production Efficiency					
Daily Average Output (units/day)	5-line average	1200	1500	+25.0%	Industry average +15%
Capacity Utilization	Actual/Theoretical capacity	68%	85%	+17 percentage points	Excellent companies 80%
Production Cycle Time	Average time per unit	45.2 seconds	36.2 seconds	-19.9%	Industry leading 35 seconds
Quality Indicators					
First Pass Yield	Monthly average	96.5%	98.8%	+2.3 percentage points	World-class 99%+
Defect Rate (PPM)	PPM value	35000	12000	-65.7%	Six Sigma <3400
Total Defect Rate	Percentage	3.5%	1.2%	-2.3 percentage points	Industry excellent <2%
Rework Rate	Rework volume/Total output	2.8%	0.9%	-67.9%	Industry excellent <1%
Equipment Efficiency					
Equipment OEE	Overall efficiency	60.0%	82.0%	+22 percentage points	World-class 85%
- Availability	Operating time/Planned time	75%	92%	+17 percentage points	Target >90%
- Performance	Actual/Standard speed	85%	91%	+6 percentage points	Target >95%
- Quality	Good units/Total output	94%	97%	+3 percentage points	Target >99%
Mean Time Between Failures (MTBF)	Hours	168	420	+150%	Industry excellent >400
Mean Time To Repair (MTTR)	Minutes	45	12	-73.3%	Target <15 minutes
Flexible Manufacturing					
Product Changeover Time	Average changeover time	120 minutes	15 minutes	-87.5%	SMED target <10 minutes
Order Response Time	Order to delivery	24 hours	6 hours	-75.0%	Industry leading 4 hours
Exception Handling Time	Discovery to resolution	30 minutes	5 minutes	-83.3%	Real-time response <5 minutes
Planning Adjustment Time	Rescheduling time	90 minutes	15 minutes	-83.3%	Agile manufacturing <20 minutes
New Product Introduction Cycle	Design to production	15 days	3 days	-80.0%	Rapid prototyping 2-5 days
Small Batch Production Capability	Minimum batch size	500 units	50 units	-90.0%	One-piece flow production
Energy Efficiency					
Unit Energy Consumption	kWh/unit	2.85	2.14	-24.9%	Green manufacturing <2.0
Energy Utilization Rate	Effective consumption/Total consumption	72%	88%	+16 percentage points	Energy saving target >85%
Inventory Management					
Work-in-Process Inventory	Turnover days	5.2 days	2.1 days	-59.6%	JIT target <2 days
Raw Material Inventory	Turnover times/year	12	24	+100%	Lean target >20
Finished Goods Inventory	Inventory value reduction	Baseline	-45%	-45%	Industry excellent -40%

(2) Availability Improvements (75% to 92%, +17pp): Predictive maintenance dramatically reduced unplanned downtime. The LSTM-based failure prediction system (94.2% accuracy, 72-hour warning window) enabled scheduled maintenance during planned downtime, reducing unplanned stops by 85%. The 87.5% changeover time reduction (120 to 15 minutes) further increased available production time. The 22-point OEE improvement aligns with academic literature reporting 15-25 percentage point gains in comprehensive digital transformation initiatives.

As shown in Figure 6(c) (Overall Equipment Effectiveness, OEE), this improvement resulted from coordinated enhancements across availability (75% to 92%), performance (85% to 91%), and quality (94% to 97%). Flexible manufacturing capabilities showed the most dramatic improvements. Figure 6(d) (Flexible manufacturing response time) illustrates the waterfall effect of time reductions across five key scenarios, with product changeover time decreasing by 87.5% from 120 to 15 minutes, approaching Single-Minute Exchange of Die (SMED) targets.

6.3 Comparative analysis

A comparative analysis with traditional MES implementations and contemporary intelligent manufacturing systems reveals the distinctive advantages of the AI-assisted, flexible architecture. Traditional MES typically achieves 10-15% productivity improvements and 5-10% quality enhancements, while the proposed system delivered 25% productivity gains and 65.7% defect reduction [36]. This performance differential stems from the integration of real-time AI inference capabilities with adaptive scheduling algorithms, enabling proactive rather than reactive manufacturing management. When benchmarked against recent intelligent manufacturing implementations, the system demonstrates competitive advantages in several key areas. Recent studies of cloud-based MES report average response times of 200-500ms for critical operations, while the implemented system maintains sub-100ms latency for 99.5th percentile transactions [37]. The ability to process 12,500 messages per second significantly exceeds typical industry implementations handling 5,000-8,000 messages per second, enabling more granular process monitoring and control.

Cost-benefit analysis reveals superior Return on Investment (ROI) compared to traditional automation approaches. While initial implementation costs were 35% higher than conventional MES due to AI infrastructure requirements, the payback period was reduced to 18 months compared to the industry average of 36 months. Total Cost of Ownership (TCO) analysis over five years indicates 40% lower operational costs due to reduced downtime, improved quality, and decreased maintenance expenses [38]. The modular microservices design enables selective upgrades and targeted technology adoption without system-wide disruption, supporting evolutionary rather than disruptive transformation. Interoperability with legacy systems while introducing advanced capabilities mitigates a major barrier to Industry 4.0 adoption, particularly for small and medium-sized enterprises operating under capital constraints.

6.4 Challenges and solutions

Implementation surfaced challenges across technical, organizational, and operational domains, requiring adaptive remedies. Technically, data quality and integration complexity dominated. Initial sensor streams exhibited 15% missing values and 8% anomalies, motivating a robust preprocessing pipeline with advanced interpolation and outlier detection. Cascaded validation at edge nodes reduced central processing by 60% and raised completeness to 98.5%. System integration was hindered by heterogeneous protocols and legacy constraints. Equipment from multiple vendors relied on proprietary interfaces, necessitating 12 custom adapters. A universal translation layer standardizing on OPC UA enabled seamless connectivity while preserving vendor-specific optimizations.

Organizational resistance to AI-guided actions required structured change management. Early operator skepticism produced a 40% override rate. Deploying explainable AI views that exposed decision rationales lowered overrides to 12% within three months. Continuous training—hands-on workshops and success-story sharing—further improved acceptance. Computational resource pressure emerged during peaks: concurrent inference pushed GPU utilization to 95%. Dynamic allocation based on priority queuing and model complexity maintained latency within bounds. An edge-cloud hybrid distributed loads and reduced central GPU requirements by 45%.

Cybersecurity challenges required specialized mitigation: (i) AI model integrity threats mitigated through cryptographic signing and blockchain-based provenance tracking (3 tampering attempts blocked); (ii) data exfiltration risks addressed via AES-256 encryption, TLS 1.3, and network micro-segmentation; (iii) real-time control attacks prevented using anomalous command detection (7 suspicious sequences identified). Following NIST Cybersecurity Framework and IEC 62443 standards, the system implemented continuous authentication, least-privilege access control (237 operator accounts, 45 PLC service accounts), network segmentation via software-defined networking, and behavioral analytics detecting 12 anomalous access patterns.

6.5 Theoretical Implications

The research contributes significant theoretical advancements to manufacturing systems theory by demonstrating the viability of distributed intelligence architectures for achieving flexible automation. The successful integration of AI cognitive capabilities with traditional MES functions validates the conceptual framework of cognitive manufacturing systems, extending CPS theory beyond simple automation to encompass adaptive learning and autonomous optimization [39]. The findings challenge existing assumptions regarding the trade-off between flexibility and efficiency in manufacturing systems. Traditional theory posits inverse relationships between these objectives, yet the implemented system achieved simultaneous improvements in both dimensions through AI-mediated dynamic optimization. This suggests a need to reconceptualize manufacturing system design principles, incorporating intelligence as a fundamental rather than auxiliary component [40]. The research establishes new theoretical constructs for understanding human-AI collaboration in manufacturing contexts. The observed evolution from initial resistance to productive partnership suggests staged acceptance models requiring further theoretical development. These findings contribute to emerging theories of augmented intelligence in industrial applications.

6.6 Practical implications for industry

The demonstrated success provides actionable insights for manufacturing practitioners considering intelligent system implementations. Organizations should prioritize data infrastructure development before AI deployment, as data quality directly impacts system effectiveness. The phased implementation approach, beginning with pilot deployments on representative production lines, reduces risk while building organizational capabilities and confidence [41]. Investment strategies should balance immediate automation needs with long-term flexibility requirements. The modular architecture approach enables incremental capability addition without wholesale system replacement, protecting capital investments while maintaining technological currency. Manufacturing leaders should allocate 20-30% of digitalization budgets to workforce development, as human factors significantly influence implementation success [42]. Strategic partnerships with technology providers accelerate implementation while reducing technical risks. However, organizations must maintain internal competencies in system architecture and data management to avoid vendor lock-in and ensure sustainable competitive advantages. The development of cross-functional teams combining operational expertise with data science capabilities proves essential for maximizing AI-driven manufacturing benefits. Small and medium manufacturers can leverage cloud-based deployment models to access advanced capabilities without prohibitive infrastructure investments. The demonstrated scalability from single-line pilots to multi-line deployments provides a roadmap for gradual digital transformation aligned with business growth and market opportunities.

6.7 Research limitations

The research exhibits several limitations requiring acknowledgment for the appropriate interpretation of findings. The implementation occurred within a single manufacturing facility producing electronic enclosures, potentially limiting generalizability to other manufacturing contexts. Process-intensive industries with continuous production may experience different implementation challenges and benefit profiles. The evaluation period of 30 days for pilot testing and 12 weeks for full implementation may not capture long-term performance variations or degradation patterns. Seasonal demand fluctuations, equipment aging effects, and evolving worker expertise could influence sustained performance metrics. Extended longitudinal studies would provide more comprehensive performance assessments [43]. Technical limitations include dependence on high-quality sensor data and reliable network connectivity. Manufacturing environments with harsh conditions or limited infrastructure may face additional implementation barriers not addressed in this research. The computational requirements for real-time AI inference may prove prohibitive for resource-constrained organizations, suggesting a need for further optimization research [44]. Section 7 synthesizes these findings into conclusions, articulates principal contributions, and identifies future research directions.

7. Conclusion

This research successfully developed and implemented an AI-assisted flexible manufacturing execution system that addresses critical limitations of traditional MES architectures in the Industry 4.0 era. The proposed framework, integrating real-time data visualization, digital twin technology, and distributed AI intelligence, achieved all design objectives while demonstrating superior performance metrics across multiple dimensions. The implementation significantly exceeded industry benchmarks across all performance metrics. The research contributes theoretical advancements by establishing cognitive manufacturing systems as a viable extension of cyber-physical systems theory, demonstrating that distributed intelligence architectures can achieve simultaneous improvements in both flexibility and efficiency, challenging traditional trade-off assumptions. For practitioners, the modular microservices architecture and phased implementation approach provide a practical roadmap for digital transformation, particularly beneficial for small and medium enterprises seeking evolutionary rather than revolutionary change. While the evaluation period and single-facility implementation present limitations regarding long-term performance assessment and cross-industry generalizability, the demonstrated benefits justify continued investigation. Future research should focus on developing industry-specific optimization algorithms and exploring federated learning approaches for multi-site deployments while maintaining data privacy and competitive advantages in increasingly connected manufacturing ecosystems.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically regarding authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with research ethics policies. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement

The manuscript contains all the data. However, more data will be available upon request from the authors.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] A. Shojaeinasab et al., "Intelligent manufacturing execution systems: A systematic review," Journal of Manufacturing Systems, vol. 62, pp. 503-522, 2022, doi: 10.1016/j.jmsy.2022.01.004.

[2] A. Tariq, S. A. Khan, W. H. But, A. Javaid, and T. Shehryar, "An IoT-enabled real-time dynamic scheduler for flexible job shop scheduling (FJSS) in an industry 4.0-based manufacturing execution system (MES 4.0)," IEEE Access, vol. 12, pp. 49653-49666, 2024, doi: 10.1109/ACCESS.2024.3378838.

[3] L. Allen, J. Atkinson, D. Jayasundara, J. Cordiner, and P. Z. Moghadam, "Data visualization for Industry 4.0: A stepping-stone toward a digital future, bridging the gap between academia and industry," Patterns, vol. 2, no. 5, 2021, doi: 10.1016/j.patter.2021.100266.

[4] H. Wang, Z. Yang, Q. Zhang, Q. Sun, and E. Lim, "A digital twin platform integrating process parameter simulation solution for intelligent manufacturing," Electronics, vol. 13, no. 4, p. 802, 2024, doi: 10.3390/electronics13040802.

[5] S. Mantravadi and C. Møller, "An overview of next-generation manufacturing execution systems: How important is MES for industry 4.0?," Procedia manufacturing, vol. 30, pp. 588-595, 2019, doi: 10.1016/j.promfg.2019.02.083.

[6] R. Beregi, G. Pedone, B. Háy, and J. Váncza, "Manufacturing execution system integration through the standardization of a common service model for cyber-physical production systems," Applied Sciences, vol. 11, no. 16, p. 7581, 2021.

[7] T. Dieguez, M. T. Malheiro, N. Leal, and J. Machado, "Systematic Literature Review on Manufacturing Execution Systems in the Era of Industry 4.0: A Bibliometric Analysis," in International Conference Innovation in Engineering, 2025: Springer, pp. 298-310, doi: 10.1007/978-3-031-94484-0_24.

[8] H. ElMaraghy, L. Monostori, G. Schuh, and W. ElMaraghy, "Evolution and future of manufacturing systems," Cirp Annals, vol. 70, no. 2, pp. 635-658, 2021, doi: 10.1016/j.cirp.2021.05.008.

[9] M. Liu, S. Fang, H. Dong, and C. Xu, "Review of digital twin about concepts, technologies, and industrial applications," Journal of manufacturing systems, vol. 58, pp. 346-361, 2021, doi: 10.1016/j.jmsy.2020.06.017.

[10] C. Su, X. Tang, Q. Jiang, Y. Han, T. Wang, and D. Jiang, "Digital twin system for manufacturing processes based on a multi-layer knowledge graph model," Scientific Reports, vol. 15, no. 1, p. 12835, 2025, doi: 10.1038/s41598-024-85053-0.

[11] D. G. Lindberg, "Harnessing AI for smart manufacturing: insights from Industry 4.0," Discover Artificial Intelligence, vol. 5, no. 1, p. 111, 2025, doi: 10.1007/s44163-025-00363-0.

[12] J. F. Arinez, Q. Chang, R. X. Gao, C. Xu, and J. Zhang, "Artificial intelligence in advanced manufacturing: current status and future outlook," Journal of Manufacturing Science and Engineering, vol. 142, no. 11, p. 110804, 2020, doi: 10.1115/1.4047855.

[13] Y. Zhang, G. Zhang, J. Wang, S. Sun, S. Si, and T. Yang, "Real-time information capturing and integration framework of the internet of manufacturing things," International Journal of Computer Integrated Manufacturing, vol. 28, no. 8, pp. 811-822, 2015, doi: 10.1080/0951192X.2014.900874.

[14] J. Wang, C. Xu, J. Zhang, and R. Zhong, "Big data analytics for intelligent manufacturing systems: A review," Journal of Manufacturing Systems, vol. 62, pp. 738-752, 2022, doi: 10.1016/j.jmsy.2021.03.005.

[15] T. Zigart, G. Kormann-Hainzl, H. Lovasz-Bukvova, M. Hölzl, T. Moser, and S. Schlund, "From lab to industry: lessons learned from the evaluation of augmented and virtual reality use cases in the Austrian manufacturing industry," Production & Manufacturing Research, vol. 11, no. 1, p. 2286345, 2023, doi: 10.1080/21693277.2023.2286345.

[16] V. M. Tabim, N. F. Ayala, G. A. Marodin, G. B. Benitez, and A. G. Frank, "Implementing Manufacturing Execution Systems (MES) for Industry 4.0: Overcoming buyer-provider information asymmetries through knowledge sharing dynamics," Computers & Industrial Engineering, vol. 196, p. 110483, 2024, doi: 10.1016/j.cie.2024.110483.

[17] X. Wang, Y. Guo, and Y. Gao, "Unmanned autonomous intelligent system in 6G non-terrestrial network," Information, vol. 15, no. 1, p. 38, 2024, doi: 10.3390/info15010038.

[18] A. Hevner and S. Chatterjee, Design research in information systems: theory and practice. Springer Science & Business Media, 2010.

[19] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," Journal of management information systems, vol. 24, no. 3, pp. 45-77, 2007, doi: 10.2753/MIS0742-1222240302.

[20] A. Bompota, "A theoretical and empirical comparative analysis of waterfall, agile and hybrid project management methods in digital projects: The case of consolut," M.S. thesis, Dept. Applied Informatics, Univ. of Macedonia, Thessaloniki, Greece, 2023. [Online]. Available: https://dspace.lib.uom.gr/handle/2159/29724

[21] K. L. Yépez Zambrano, "Tecnologías de comunicaciones: Prototipo para adquisición de datos de un sistema de semáforos para administración de tráfico vehicular en una ciudad inteligente," B.S. thesis, Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, Quito, Ecuador, 2024.

[22] K. Kaur, S. Garg, G. Kaddoum, S. H. Ahmed, and M. Atiquzzaman, "KEIDS: Kubernetes-based energy and interference driven scheduler for industrial IoT in edge-cloud ecosystem," IEEE Internet of Things Journal, vol. 7, no. 5, pp. 4228-4237, 2019, doi: 10.1109/JIOT.2019.2939534.

[23] T. P. Raptis and A. Passarella, "A survey on networked data streaming with apache kafka," IEEE access, vol. 11, pp. 85333-85350, 2023, doi: 10.1109/ACCESS.2023.3303810.

[24] P. Strelec, T. Horak, S. Kovac, E. Nemlaha, and P. Tanuska, "IIoT device prototype design using state machine according to OPC UA," IEEE Access, vol. 10, pp. 134004-134017, 2022, doi: 10.1109/ACCESS.2022.3233147.

[25] M. Benková, D. Bednárová, and G. Bogdanovská, "Process capability evaluation using capability indices as a part of statistical process control," Mathematics, vol. 12, no. 11, p. 1679, 2024, doi: DOI: 10.3390/math12111679.

[26] P. Reali, R. Lolatto, S. Coelli, G. Tartaglia, and A. M. Bianchi, "Information retrieval from photoplethysmographic sensors: A comprehensive comparison of practical interpolation and breath-extraction techniques at different sampling rates," Sensors, vol. 22, no. 4, p. 1428, 2022, doi: 10.3390/s22041428.

[27] X. Sun, Y. He, D. Wu, and J. Z. Huang, "Survey of distributed computing frameworks for supporting big data analysis," Big Data Mining and Analytics, vol. 6, no. 2, pp. 154-169, 2023, doi: 10.26599/BDMA.2022.9020014.

[28] D. Kibira, K. C. Morris, and S. Kumaraguru, "Methods and tools for performance assurance of smart manufacturing systems," Journal of Research of the National Institute of Standards and Technology, vol. 121, p. 282, 2016, doi: 10.6028/jres.121.013.

[29] Prometheus Authors, "Prometheus - Monitoring system & time series database," 2017. [Online]. Available: https://prometheus.io/

[30] X. Lv, L. Shi, Y. He, Z. He, and D. K. Lin, "Joint optimization of production, maintenance, and quality control considering the product quality variance of a degraded system," Frontiers of Engineering Management, vol. 11, no. 3, pp. 413-429, 2024, doi: 10.1007/s42524-024-3103-1.

[31] International Electrotechnical Commission, "Functional safety of electrical/electronic/programmable electronic safety-related systems," IEC 61508, 2010. [Online]. Available: https://webstore.iec.ch/publication/5515

[32] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on industrial Internet of Things: A cyber-physical systems perspective," Ieee access, vol. 6, pp. 78238-78259, 2018, doi: 10.1109/ACCESS.2018.2884906.

[33] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: The confluence of edge computing and artificial intelligence," IEEE Internet of Things Journal, vol. 7, no. 8, pp. 7457-7469, 2020,

doi: 10.1109/JIOT.2020.298488710.1109/JIOT.2020.2984887.
[34] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial internet of things: Challenges, opportunities, and directions," IEEE transactions on industrial informatics, vol. 14, no. 11, pp. 4724-4734, 2018, doi: 10.1109/TII.2018.2852491.
[35] T. Hai et al., "Task scheduling in cloud environment: optimization, security prioritization and processor selection schemes," Journal of Cloud Computing, vol. 12, no. 1, p. 15, 2023, doi: 10.1186/s13677-022-00374-7.
[36] R. X. Gao, J. Krüger, M. Merklein, H.-C. Möhring, and J. Váncza, "Artificial Intelligence in manufacturing: State of the art, perspectives, and future directions," CIRP Annals, vol. 73, no. 2, pp. 723-749, 2024, doi: 10.1016/j.cirp.2024.04.101.
[37] C. Li, P. Zheng, Y. Yin, B. Wang, and L. Wang, "Deep reinforcement learning in smart manufacturing: A review and prospects," CIRP Journal of Manufacturing Science and Technology, vol. 40, pp. 75-101, 2023, doi: 10.1016/j.cirpj.2022.11.003.
[38] J. Lee, M. Azamfar, and J. Singh, "A blockchain enabled Cyber-Physical System architecture for Industry 4.0 manufacturing systems," Manufacturing letters, vol. 20, pp. 34-39, 2019, doi: 10.1016/j.mfglet.2019.05.003.
[39] Y. He, B. Wu, J. Mao, W. Jiang, J. Fu, and S. Hu, "An effective MID-based visual defect detection method for specular car body surface," Journal of Manufacturing Systems, vol. 72, pp. 154-162, 2024, doi: 10.1016/j.jmsy.2023.11.014.
[40] A. Setyadi, S. Soekotjo, S. D. Lestari, S. Pawirosumarto, and A. Damaris, "Trends and opportunities in sustainable manufacturing: a systematic review of key dimensions from 2019 to 2024," Sustainability, vol. 17, no. 2, p. 789, 2025, doi: 10.3390/su17020789.
[41] S.-C. Chen, H.-M. Chen, H.-K. Chen, and C.-L. Li, "Multi-objective optimization in Industry 5.0: Human-centric AI integration for sustainable and intelligent manufacturing," Processes, vol. 12, no. 12, p. 2723, 2024, doi: 10.3390/pr12122723.
[42] J. Pochmara and A. Świetlicka, "Cybersecurity of industrial systems—A 2023 report," Electronics, vol. 13, no. 7, p. 1191, 2024, doi: 10.3390/electronics13071191.
[43] T. Khan, U. Khan, A. Khan, C. Mollan, I. Morkvenaite-Vilkonciene, and V. Pandey, "Data-driven digital twin framework for predictive maintenance of smart manufacturing systems," Machines, vol. 13, no. 6, p. 481, 2025, doi: 10.3390/machines13060481.
[44] Y. Lu, X. Xu, and L. Wang, "Smart manufacturing process and system automation–a critical review of the standards and envisioned scenarios," Journal of Manufacturing Systems, vol. 56, pp. 312-325, 2020, doi: 10.1016/j.jmsy.2020.06.010.