

# AI-Powered Scheduling in High-Mix, Low-Volume (HMLV) Manufacturing: An Appraisal of the Challenges, Approaches, and Future Directions


**Chukwuma Godfrey Ono**

Industrial/Production Engineering Department, Nnamdi Azikiwe University, P.M.B. 5025 Awka, Anambra State – Nigeria, [cg.ono@unizik.edu.ng](mailto:cg.ono@unizik.edu.ng)

**Fredrick Nnaemeka Okeagu**

Industrial/Production Engineering Department, Nnamdi Azikiwe University, P.M.B. 5025 Awka, Anambra State – Nigeria, [fn.okeagu@unizik.edu.ng](mailto:fn.okeagu@unizik.edu.ng)

\*Corresponding Author's Email: [cg.ono@unizik.edu.ng](mailto:cg.ono@unizik.edu.ng)

ARTICLE INFO	ABSTRACT
<p><b>Keywords:</b> <i>High-Mix Low-Volume manufacturing, AI-powered scheduling, reinforcement learning, digital twins, edge computing</i></p> <p><i>Received: 24, Jan. 2026</i> <i>Revised: 16, Feb. 2026</i> <i>Accepted: 17, Feb. 2026</i></p> <p>©2026 Author(s): This is an open-access article distributed under the terms of the <a href="https://creativecommons.org/licenses/by/4.0/">Creative Commons Attribution 4.0 International</a></p> 	<p><i>High-Mix Low-Volume manufacturing faces persistent scheduling pressure from high product variety, frequent changeovers, alternative routings, volatile demand, and disruption events such as rush orders and equipment failures. These conditions expose limits in fixed dispatching rules, static expert policies, and large-scale mathematical programming, especially under uncertainty and tight decision windows. This review synthesizes AI-powered scheduling approaches developed for HMLV settings, with emphasis on predictive, adaptive, and decentralized control. Machine learning models support processing-time estimation, bottleneck prediction, and schedule recommendations. Reinforcement learning and deep reinforcement learning frame scheduling as sequential decision-making and learn policies that adapt under nonstationary shop-floor states, often trained and validated through simulation or digital twin environments. Metaheuristics and learning-guided hybrid methods address combinatorial scale by steering search toward feasible, high-quality schedules within practical runtimes. Multi-agent systems distribute scheduling decisions across job and resource agents through negotiation and coordination, supporting resilience and scalability in modular production lines. The review also consolidates integration requirements across IoT sensing, MES and ERP connectivity, edge deployment for low-latency execution, and digital-twin-driven verification. Comparative evaluation metrics include makespan, tardiness, utilization, and disruption responsiveness. Open challenges include sparse and inconsistent data, limited interpretability of high-capacity models, compute burden, trust and governance in human-in-the-loop operation, and sustainability and fairness constraints. Future directions prioritize explainable scheduling, continuous learning, federated training across sites, LLM-assisted scheduling knowledge, and alignment with autonomous manufacturing architectures.</i></p>

## INTRODUCTION

### 1.1 Context of High-Mix, Low-Volume (HMLV) Manufacturing

High-Mix, Low-Volume (HMLV) manufacturing refers to production environments characterized by a wide variety of products, customized designs, and small batch sizes (Gan et al., 2023; Normand and Bradley, 2024; Tahmina et al., 2022). Unlike mass production systems that emphasize scale and standardization, HMLV prioritizes flexibility and responsiveness to changing customer requirements, regulatory landscapes, and dynamic demand patterns (Normand and Bradley, 2024; Tahmina et al., 2022). Industries exemplifying HMLV include aerospace manufacturing with custom-engineered components, medical device manufacturing where regulatory compliance and patient-specific customization are critical, and custom machinery sectors where unique client requirements drive substantial product variability (Normand and Bradley, 2024; Tahmina et al., 2022). The inherent diversity of product types and the frequent switching between production tasks

make HMLV environments fundamentally complex to manage, challenging planning, sequencing, and shop-floor control (Nwamekwe et al., 2025; Ojsteršek et al., 2020).

To cope with this complexity, AI-powered scheduling, digital twins, and Industry 4.0-enabled decision support have emerged as core approaches to HMLV orchestration. Real-time decision-support systems for high-mix, low-volume production scheduling have been proposed to handle variability and enable rapid reallocation of resources in Industry 4.0 settings (Kocsi et al., 2020). Digital twin-based designs for HMLV contexts enable simulation-based optimization of job allocation and scheduling to balance productivity with mass personalization (Sit and Lee, 2023). The integration of robotics and human-in-the-loop control is supported by capability-based frameworks for industrial robot skills, highlighting the role of automation in enabling flexible, on-demand production (Pantano, 2022). Moreover, 5G-connected smart manufacturing architectures provide the communication backbone for data-driven scheduling and remote monitoring in HMLV environments, facilitating edge computing and rapid data exchange (Ludwig et al., 2022). Future directions also emphasize sustainability in flexible manufacturing and demand-driven collaborations among SMEs to enable ramp-up and resilience in HMLV settings (Ojsteršek et al., 2020). Additionally, reinforcement learning and other AI-based optimization methods have been explored for scheduling under high variability, underscoring the AI-driven direction of scheduling in HMLV contexts (Julaiti et al., 2022).

## 1.2 Importance of Scheduling in HMLV Environments

Scheduling in High-Mix Low-Volume (HMLV) manufacturing is central to operational performance because it determines how tasks are allocated to machines, the sequencing of jobs, and the efficient use of scarce shop-floor resources (Gan et al., 2023; Kocsi et al., 2020; Zanchettin, 2021). The HMLV milieu is defined by a broad product variety, customized designs, and small batch sizes, which substantially elevate planning and control complexity relative to repetitive production systems (Gan et al., 2023; (Kocsi et al., 2020). The resulting variability in demand, diverse routing, and nontrivial setup/changeover times render schedules highly dynamic and prone to disruption, underscoring the need for adaptive, real-time decision support rather than static planning (Kocsi et al., 2020; (Julaiti et al., 2022). In response, researchers have advanced real-time scheduling frameworks and digital-twin driven optimization to manage high-mix variability and support rapid resource reallocation within Industry 4.0 environments (Kocsi et al., 2020; Zanchettin, 2021).

Beyond the immediate production impacts, suboptimal scheduling in HMLV yields tangible penalties in lead times, operating costs, and resource underutilization, particularly when rush orders or unplanned machine failures occur (Zanchettin, 2021; Julaiti et al., 2022). This reality motivates the exploration of robust, adaptive strategies that integrate AI and constraint-based techniques to balance utilization with responsiveness, including reinforcement learning for stochastic, nonstationary settings and POLCA-inspired constraints embedded in scheduling models for high-mix contexts (Julaiti et al., 2022; Zhou et al., 2024). Moreover, sustainability considerations are increasingly embedded in scheduling objectives, driving energy-aware and low-carbon scheduling research in flexible job shops and HMLV settings (Liu et al., 2024). Collectively, these threads, AI methods, digital twins, and Industry 4.0 data infrastructures, position scheduling in HMLV not merely as a control problem but as a strategic capability for mass personalization and operational resilience (Kocsi et al., 2020; Zanchettin, 2021).

## 1.3 Motivation for AI-Powered Scheduling

Traditional scheduling methods, including heuristics, rule-based strategies, and mathematical programming, have long been applied in manufacturing, but they struggle to cope with the uncertainty, complexity, and dynamic nature of high-mix, low-volume (HMLV) systems. In particular, optimization-based approaches often become computationally intractable as problem size

and variability grow, while heuristic rules can fail to adapt to unforeseen disruptions and changing priorities (Gan et al., 2023; Castañé et al., 2022). The HMLV context characterized by a broad product variety, frequent routing changes, and nontrivial setup times, exposes these limitations and motivates the search for more flexible, data-driven scheduling solutions (Gan et al., 2023; Nwamekwe et al., 2025; Castañé et al., 2022). AI-enabled scheduling, which leverages machine learning, reinforcement learning, and multi-agent systems, offers a promising alternative by enabling predictive, adaptive, and data-driven decision making that can learn from history, respond to real-time changes, and optimize across multiple, often conflicting objectives (Julaiti et al., 2022; Gallo et al., 2023).

The motivation to adopt AI-powered scheduling in HMLV is reinforced by evidence that AI can improve responsiveness, resilience, and overall performance under variability and disturbances. AI-based approaches can anticipate delays and disruptions, adapt to nonstationary conditions, and balance competing goals such as throughput, lead times, energy use, and cost (Julaiti et al., 2022; Gallo et al., 2023; Chen et al., 2022). Moreover, contemporary Industry 4.0 concepts, digital twins, asset-and-agent modelling, and real-time data integration, provide the enabling infrastructure for AI scheduling to operate at the ecosystem level, enabling simulation-based optimization, real-time reallocation, and coordinated human-machine collaboration (Nwamekwe et al., 2024). As organizations seek mass personalization and sustainable operations, AI-driven scheduling becomes not just a control mechanism but a strategic capability to enhance agility, resilience, and efficiency in HMLV manufacturing (Castañé et al., 2022; Nwamekwe et al., 2024; Thakre et al., 2023).

#### **1.4 Scope and Contributions of This Review**

This review paper synthesizes the current state of AI-powered scheduling in HMLV manufacturing. It examines the scheduling challenges inherent in such environments, evaluates different AI approaches including supervised learning, reinforcement learning, metaheuristics, and multi-agent systems, and discusses integration with digital technologies such as edge computing and digital twins. Furthermore, it highlights comparative studies, open research questions, and future directions, with a focus on bridging theoretical advances and industrial adoption. The novelty of this review lies in its exclusive focus on scheduling algorithms tailored for HMLV systems, a niche but critical domain within smart manufacturing research.

## **2. The Scheduling Challenge in HMLV Manufacturing**

### **2.1 Characteristics of HMLV Systems**

High-Mix Low-Volume (HMLV) manufacturing is defined by substantial product diversity, fluctuating demand, and dynamic routing requirements, with jobs that vary significantly in size, complexity, and processing time (Gan et al., 2023; Normand and Bradley, 2024; Tahmina et al., 2022). Frequent product changes drive nontrivial machine setup times, and routing often involves multiple alternative machines and path choices, amplifying the complexity of feasible schedules (Normand and Bradley, 2024; Nwamekwe et al., 2025; Tahmina et al., 2022). These intertwined characteristics create highly intricate decision-making environments in which traditional scheduling models struggle to scale or adapt, limiting their effectiveness in capturing variability and mass personalization needs (Gan et al., 2023; Kocsi et al., 2020; Pantano, 2022).

The resultant scheduling challenge is further intensified by disruptions such as rush orders and machine failures, which amplify variability and stress existing planning approaches (Kocsi et al., 2020; Sit and Lee, 2023; Pantano, 2022). Consequently, researchers have pursued AI-enabled and data-driven scheduling frameworks encompassing machine learning, reinforcement learning, and multi-agent systems to enable predictive, adaptive decision-making across competing objectives under nonstationary conditions (Ludwig et al., 2022; Julaiti et al., 2022; Nwamekwe et al., 2025; Ojsteršek et al., 2020). The enabling ecosystem of Industry 4.0 infrastructures, digital twins, real-time data integration, and distributed scheduling agents facilitates simulation-based optimization, rapid

resource reallocation, and tighter human-machine collaboration, thereby supporting robust operation in HMLV contexts (Nwamekwe et al., 2025; Julaiti et al., 2022; Ojsteršek et al., 2020).

Table 1. HMLV Scheduling Stressors and Operational Consequences

HMLV characteristic	Scheduling implication	Shop-floor symptom	Planning consequence
High product variety	Frequent priority shifts across diverse job families	Queue reshuffling, sequencing conflicts	Higher dispatching volatility, replan frequency increases
Small batch sizes	Limited repetition for stable cycle-time learning	Sparse history for many part numbers	Higher uncertainty in duration and due-date promises
Frequent changeovers and setups	Sequence-dependent setup penalties dominate objectives	Idle time during setups, changeover bottlenecks	Need setup-aware sequencing and family grouping
Alternative routings and re-entrance	Expanded feasible schedule space and constraint interactions	Routing swaps under congestion	Higher combinatorial complexity for exact solvers
Volatile demand and rush orders	Dynamic insertions and reprioritization requirements	Expedites interrupt planned flow	Schedule instability, higher tardiness risk
Machine failures and downtime	Unplanned capacity loss and cascading delays	WIP accumulation, missed due dates	Need disruption-aware reallocation and recovery policies
Processing-time variability	State uncertainty for dispatch and completion predictions	Late finishes, missed handoffs	Need predictive duration models and robust buffers
Due-date variability and tight windows	Multi-objective tension between tardiness and utilization	Frequent rescheduling meetings	Need penalty-aware policies and service-level control

Table 1 traces how structural HMLV variability propagates into shop-floor disturbances and planning instability. It demonstrates scheduling challenges arise from system design features, reinforcing the need for predictive models, adaptive sequencing, and disruption-aware control.

## 2.2 Traditional Scheduling Approaches

Historically, scheduling in HMLV systems has relied on:

- i. Heuristics, such as dispatching rules (e.g., shortest processing time, earliest due date). Heuristics have historically served as fast and implementable tools for HMLV scheduling due to their low computational burden and ease of integration with shop-floor control, which suits environments with high product variety and frequent changeovers (Tahmina et al., 2022; Zanchettin, 2021). In HMLV contexts, dispatching rules provide immediate guidance for task selection and sequencing without solving large optimization problems, making them attractive for real-time or near-real-time decision support (Tahmina et al., 2022; Zanchettin, 2021). However, these rules optimize local objectives and often neglect global considerations such as setup sequencing, routing alternatives, and multi-criteria trade-offs, potentially leading to suboptimal lead times and underutilized capacity when variability and non stationarity are high (Tahmina et al., 2022; Zanchettin, 2021). The brittleness of purely heuristic policies under disruptions such as rush orders or unforeseen machine failures, further motivates

embedding heuristics within adaptive, data-driven decision-support frameworks rather than relying on fixed rules alone (Kocsi et al., 2020; Zanchettin, 2021).

- ii. Mathematical programming, including mixed-integer linear programming (MILP), which provides exact solutions but at high computational cost. MILP offers rigorous representations of constraints (capacities, setup times, and routing) and the potential for globally optimal schedules, but the combinatorial complexity induced by large product assortments, numerous setups, and routing choices in HMLV often renders exact solutions computationally intractable for timely decision making (Emeka et al., 2025). Techniques to improve tractability such as using record-keeping variables (RKVs) to reduce branching and accelerate MILP search in discrete-time production scheduling, demonstrate how exact formulations can be made more scalable in practice, albeit with continued sensitivity to problem size. Additional work shows that strengthening bounds through preprocessing and cutting planes can yield substantial solve-time reductions for RCPSP-like scheduling problems common in flexible environments (Emeka et al., 2025). More recently, efforts to augment MILP with learning-based cut selection or objective-value prediction indicate a path toward hybrid, data-driven optimization that can cope with nonstationary, large-scale HMLV instances.
- iii. Rule-based methods, where human experts encode domain knowledge into scheduling policies. Rule-based scheduling captures practitioner expertise, enabling interpretable and deployable policies that align with established manufacturing practices and capacity awareness in HMLV contexts (Kocsi et al., 2020; Zanchettin, 2021). These policies can provide robust performance in well-characterized processes and can be effectively deployed in real-time settings when paired with continuous monitoring and feedback (Kocsi et al., 2020). Yet, reliance on static, expert-crafted rules can limit adaptability to novel product mixes, routing permutations, and unexpected disturbances, reducing resilience in highly dynamic HMLV environments without ongoing expert maintenance and periodic rule revision (Zanchettin, 2021). Consequently, contemporary practice increasingly treats rule-based approaches as components of hybrid AI–OR frameworks that balance explainability with optimization-driven performance (Kocsi et al., 2020; Zanchettin, 2021). But while these methods offer structure, they often fail to scale or adapt in dynamic environments.

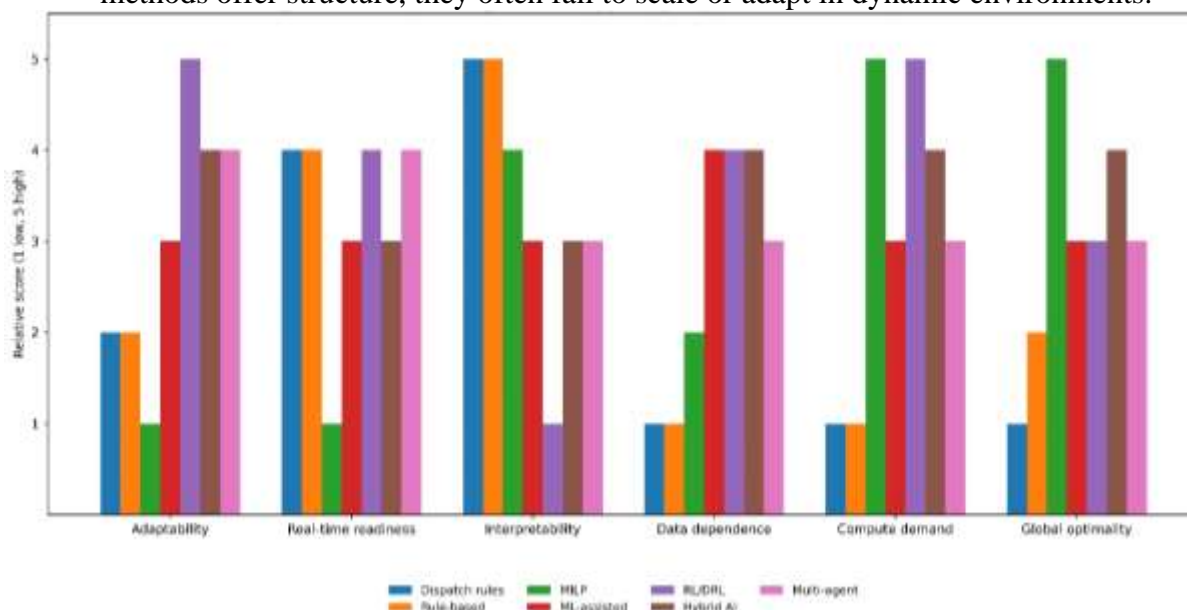


Figure 1. Relative Positioning of Scheduling Paradigms Against HMLV Decision Requirements

Figure 1 compares scheduling paradigms against responsiveness, transparency, computational burden, and optimization capacity. It explains why adaptive learning approaches outperform static

rules in HMLV environments where disruptions, routing diversity, and uncertainty dominate scheduling decisions.

### 2.3 Limitations of Conventional Approaches in HMLV

The limitations of traditional methods include:

1. Computational complexity remains a fundamental limitation of conventional exact optimization approaches in high-mix low-volume (HMLV) scheduling, as mixed-integer linear programming (MILP) formulations, while capable of representing capacities, setups, and routing, rapidly become intractable as product variety and disruption frequencies grow in high-mix environments (Zhou et al., 2024). Empirical work demonstrates that even with improvements like data-driven cut selection and preprocessing, solving large-scale MILP-based scheduling problems can be impractical for near-real-time decision-making in dynamic shops (Tang et al., 2023; Zhou et al., 2024), prompting exploration of alternative paradigms such as learning-enhanced optimization and predictive analytics to restore tractability under nonstationary conditions (Tang et al., 2023). The persistent tension between solution quality and computational burden motivates hybrid, approximate, and data-driven approaches that can deliver timely, good-enough schedules without sacrificing responsiveness to mass customization demands (Zhao et al., 2022).
2. Rigidity is another central drawback of traditional rule-based and heuristic scheduling in HMLV contexts, where predefined policies and fixed priority rules often fail to adapt to new products, routing permutations, or unexpected disturbances characteristic of HMLV systems (Zanchettin, 2021; Zhou et al., 2024). While heuristics such as dispatching rules offer fast, implementable guidance, their local focus and lack of global coordination can degrade performance when setups, routes, and urgent orders shift, unless they are embedded within adaptive, data-informed decision-support frameworks (Zanchettin, 2021). Empirical evidence from studies on reinforcement learning indicates that adaptable, learning-based controllers outperform static rule sets in the face of disruption and non stationarity, underscoring the need to move beyond rigid policy catalogues toward flexible, capable scheduling architectures (Zhou et al., 2024).
3. Uncertainty handling is another critical limitation of conventional approaches, as many deterministic models assume fixed processing times, stable demand, and predictable disruptions, which rarely hold in HMLV manufacturing. Uncertainties in processing times, due-date variability, and sudden events necessitate robust or stochastic strategies; however, many classical methods lack explicit mechanisms to hedge against such variability. Recent advances incorporating fuzzy, stochastic, or neutrosophic representations, as well as robust optimization and reinforcement-learning-based online adaptation, have demonstrated improved resilience to uncertainty and better maintenance of service levels in heterogeneous, variable environments. These approaches illustrate why conventional scheduling methods often perform poorly under real-world fluctuations and motivate the development of AI-powered scheduling solutions that can learn, adapt, and replan in the presence of uncertainty (Zhou et al., 2024).

These gaps provide a strong rationale for AI-powered approaches.

## 3. AI-Powered Scheduling Approaches

### 3.1 Machine Learning-Based Scheduling

Machine learning has been applied to predict job durations, identify potential bottlenecks, and generate scheduling recommendations in manufacturing contexts, including high-mix low-volume (HMLV) environments, where data-driven insights are essential for timely reallocation of resources

(Lu et al., 2020; Zhou et al., 2024). Supervised learning models are widely used to estimate processing times and resource requirements, providing valuable inputs for dispatching and sequencing decisions; empirical work demonstrates effective use of machine learning for predicting standard times and job durations, informing realistic due dates and capacity planning (Igbokwe and Nwamekwe, 2025; Liu et al., 2024). Reinforcement learning models, in contrast, continuously improve scheduling decisions through trial-and-error interaction with a simulated environment or digital twin, thereby enhancing adaptability to disturbances and nonstationary demand (Zhou et al., 2024; Vitalis et al., 2024; Zhang et al., 2022), with recent work illustrating deep reinforcement learning-based policy selection and real-time adaptation in job shop and flow shop contexts (Zhou et al., 2024; Zhang et al., 2022). The enabling role of digital twins and Industry 4.0 data infrastructures is also evident, as machine learning-enabled schedulers leverage simulated environments and real-time data streams to explore alternatives and reallocate capacity rapidly (Lu et al., 2020).

Beyond purely supervised or reinforcement learning approaches, hybrid AI strategies are emerging to combine predictive modelling with optimization, reducing search spaces and accelerating convergence in complex HMLV settings (Chen et al., 2022). For example, machine learning components can predict promising search regions for metaheuristics or guide the coupling of learning with exact optimization in matheuristics, improving scalability without compromising solution quality (Chen et al., 2022). Two-stage learning frameworks illustrate a concrete path where supervised predictions feed subsequent optimization, delivering high-quality schedules under variability (Ahn et al., 2024), while federated digital-twin scheduling and multi-agent reinforcement learning highlight distributed, adaptive decision-making in dynamic shop floors (Castañé et al., 2022). Collectively, these machine-learning-driven approaches illuminate a practical trajectory for AI-powered scheduling in HMLV manufacturing, where predictive accuracy, adaptive control, and scalable computation are essential to achieving mass personalization with operational resilience (Igbokwe and Nwamekwe, 2025; Castañé et al., 2022).

### 3.2 Metaheuristic and Hybrid AI Methods

Metaheuristics such as genetic algorithms (GA), particle swarm optimization (PSO), and related approaches have been extensively employed to tackle scheduling in manufacturing, including HMLV settings, owing to their ability to explore large, nonconvex search spaces and to find high-quality solutions within reasonable time frames. AI-enhanced iterations of these methods contribute predictive insights, enable informed initialization, and guide search trajectories to accelerate convergence, as demonstrated by frameworks that couple neural networks with GA and other metaheuristics to produce data-driven scheduling recommendations (Igbokwe et al., 2025), and by ML-assisted approaches that predict promising regions for search in specialized scheduling problems such as ground-station scheduling (Xu et al., 2024). In parallel, hybrid “matheuristic” models where ML or AI components are integrated with metaheuristics or exact optimization illustrate practical pathways to combine the exploration power of metaheuristics with the precision of optimization, such as in two-hybrid flow-shop scheduling lines that merge compatibility-aware constraints with learning-informed search strategies (Muñoz-Sánchez et al., 2024), and in broader AI-assisted scheduling studies for distributed manufacturing systems (Igbokwe et al., 2025).

The appeal of hybrid AI methods lies in their capacity to reduce search spaces and to improve convergence by learning where to search and how to adjust operators, rather than blindly exploring the entire space. Empirical and methodological work shows that ML can steer genetic search toward regions with higher promise, enabling faster convergence to near-optimal solutions in complex production environments, including high-mix contexts (Xu et al., 2024; Igbokwe et al., 2025). Matheuristics further demonstrate how predictive models can be embedded within exact or near-exact search procedures to boost scalability without sacrificing solution quality, as seen in hybrid flow-shop problems that blend optimization with data-driven guidance (Muñoz-Sánchez et al., 2024). Real-

time, Industry 4.0-enabled scheduling studies also reveal the value of such hybrids in dynamic settings, where a real-time decision-support orientation can benefit from AI-guided metaheuristics to cope with disturbances and nonstationary demand (Kocsi et al., 2020). Collectively, these approaches position metaheuristics and their AI-augmented hybrids as a practical and scalable path for AI-powered scheduling in HMLV environments.

### 3.3 Reinforcement Learning (RL) and Deep RL

Reinforcement learning (RL), including its deep variants (deep reinforcement learning, DRL), has been increasingly applied to scheduling by framing it as a sequential decision-making problem where agents learn policies that aim to minimize make span, tardiness, or other joint objectives under uncertainty in demand, processing times, and disruptions (Liu et al., 2024; Yang et al., 2024). In particular, Deep Q-Networks (DQN) and policy-gradient methods have been used to dynamically allocate tasks and adapt dispatching under changing shop-floor states, supporting more responsive and autonomous scheduling in HMLV environments (Zhao et al., 2022; Zhang et al., 2022). The effectiveness of DRL approaches is often demonstrated in conjunction with digital twin or simulation environments, where agents can safely explore alternatives and continuously improve policies before deployment in real production (Chang et al., 2022). These trends position RL/DRL as central components of AI-powered scheduling capable of handling nonstationary demand and complex routing typical of HMLV settings (Liu et al., 2024; Yang et al., 2024; Zhao et al., 2022).

Beyond single-agent paradigms, recent work shows that RL and DRL can be extended to robust and distributed scheduling in dynamic manufacturing through re-entrant lines, random job arrivals, and multi-agent frameworks. Studies on DRL for dynamic or flexible job-shop problems with randomness or disturbances demonstrate improved resilience and responsiveness compared with static heuristics (Chang et al., 2022; Vitalis et al., 2024), while PPO-based and actor-critic DRL variants offer stable learning and real-time adaptability in complex shop floors (Zhang et al., 2022; Zhang et al., 2022). Moreover, the literature increasingly reports on hybrid configurations combining DRL with learning-informed search or multi-agent coordination to achieve scalable, reliable performance in cloud and distributed manufacturing environments (Nwamekwe and Nwabunwanne, 2025). Collectively, these RL/DRL developments, spanning single-agent DRL, deep Q-learning, and multi-agent frameworks offer a practical, scalable pathway for AI-powered scheduling that can cope with the variability and disruption inherent to HMLV manufacturing (Yang et al., 2024; Vitalis et al., 2024; Nwamekwe and Nwabunwanne, 2025).

### 3.4 Multi-Agent Systems for Decentralized Scheduling

In multi-agent systems (MAS) for scheduling, machines and jobs are modelled as autonomous agents that negotiate, coordinate, and reallocate tasks to form feasible schedules. This decentralized paradigm enhances scalability and resilience, as local decisions can be made in real-time in response to disturbances or partial information, without relying on a single centralized planner (Pal et al., 2022; Kölle et al., 2022). Coordination mechanisms such as contract nets and auction-based bidding have been employed to resolve task assignments and routing in HMLV contexts, enabling adaptive decision-making as product mixes and routing options evolve (Pal et al., 2022; Kölle et al., 2022). Ontological representations and knowledge-based interactions further support interoperability and shared understanding across agents, strengthening distributed control in manufacturing systems. Delegated MAS approaches also illustrate how route planning can be distributed among specialized agents in modular lines, reducing bottlenecks and enabling more flexible. Together, these works demonstrate MAS as a practical framework for decentralized scheduling in diverse HMLV settings, where agents negotiate to balance throughput, lead times, and setup considerations under uncertainty (Kölle et al., 2022).

Beyond deployment aspects, the literature on MAS for scheduling highlights key design and performance considerations. Decentralized planning under uncertainty is often framed via decentralized partially observable Markov decision processes (Dec-POMDPs) or macro-actions, providing a theoretical basis for robust coordination and decision-making when agents have partial observability and limited communication (Chidiebube et al., 2025). Effective MAS also relies on communication-efficient mechanisms, with importance-aware message exchange improving learning and coordination in multi-agent reinforcement learning contexts, thereby mitigating communication bottlenecks in dynamic shops (Igbokwe et al., 2024). Human-in-the-loop considerations emerge in MAS planning, where high-level guidance can be integrated to preserve governance and domain knowledge while maintaining autonomous agent operation (Wu et al., 2021). Collectively, these studies illustrate a mature trajectory for MAS-based scheduling in HMLV, where decentralized coordination, adaptive bargaining, and knowledge-rich representations enable scalable, resilient, and context-aware production orchestration (Pal et al., 2022; Chidiebube et al., 2025; Igbokwe et al., 2024; Wu et al., 2021).

Table 2. AI-Powered Scheduling Approach Matrix with Mechanisms, Inputs, and Outputs

Approach family	Decision mechanism	Typical inputs	Typical outputs	Primary strength	Primary constraint
Supervised ML scheduling	Predict durations, bottlenecks, priority scores	MES and ERP logs, routing, historical sequences	Duration estimates, dispatch priorities, risk flags	Improves prediction and planning realism	Needs consistent labels and coverage across variants
RL and DRL scheduling	Learn sequencing and dispatch policies through interaction	State features, WIP, machine status, due dates, digital twin simulator	Policy actions for sequencing, allocation, replan triggers	High adaptability under disruptions	Low transparency, high training burden
Metaheuristics and hybrid AI	Search over schedules with learned guidance or matheuristics	Constraints, processing times, setups, learned promising regions	Near-optimal schedules under multi-objectives	Scales in large combinatorial spaces	Tuning effort and runtime sensitivity
Multi-agent scheduling	Distributed negotiation and coordination among agents	Local machine states, job attributes, messaging protocols	Decentralized assignments and routing decisions	Scalability and resilience through locality	Coordination overhead and stability under partial observability

Table 2 is a matrix that explains operational roles of AI scheduling families by linking mechanisms, data requirements, outputs, benefits, and limitations. It clarifies complementary contributions of predictive modelling, optimization search, and decentralized coordination in modern HMLV scheduling.

s

#### 4. Data, Infrastructure, and Integration

##### 4.1 Data Requirements and Sources

AI-powered scheduling in HMLV manufacturing hinges on diverse data streams, notably shop-floor IoT sensor data, MES/ERP logs, and rich historical production data (past sequences, lead times, and outcomes) to capture current states and historical patterns that inform predictive and

prescriptive models (Rossit et al., 2019). The integration of these heterogeneous sources through Industry 4.0-enabled architectures provide the data richness needed for accurate forecasting, adaptive dispatching, and scenario-based optimization, enabling schedulers to anticipate bottlenecks and reallocate resources proactively (Rao et al., 2020; Rossit et al., 2019). Digital twins and related CPS/IoT platforms extend these capabilities by offering simulated environments and real-time data streams that support safe training, testing, and validation of AI scheduling policies before deployment in live operations.

Beyond data types, the infrastructure to collect, harmonize, and exploit them is critical. Middleware and integration approaches that connect MES/ERP ecosystems with scheduling engines enable end-to-end visibility and control across planning and execution layers, as demonstrated by MES-AIV integrations and similar middleware solutions in Industry 4.0 settings (Rao et al., 2020). Digital twin-driven scheduling studies further illustrate how data from sensors, enterprise systems, and historical records can be fused to build accurate predictive models and prescriptive policies, while cloud- and platform-enabled data pipelines support scalable, cross-plant data sharing essential for distributed manufacturing contexts (Rossit et al., 2019; Wang et al., 2021). Collectively, these data infrastructures underpin AI-powered scheduling by providing the timely, high-fidelity information that drives adaptive decision making in highly variable HMLV environments (Chang et al., 2022; Rao et al., 2020; Xu et al., 2024).

#### **4.2 Real-Time Scheduling and Edge Computing**

To achieve low-latency decision-making in AI-powered scheduling for HMLV manufacturing, AI models are increasingly deployed at the edge proximate to production equipment to enable immediate perception, inference, and action without routing data to distant clouds (Zhou et al., 2024; Wang et al., 2021; Hamdan et al., 2020). Edge deployment reduces reliance on centralized servers, enabling real-time responses to disruptions such as machine breakdowns and unexpected rush orders, thereby improving the resilience and timeliness of scheduling decisions (Wu et al., 2021). This edge paradigm is frequently augmented by digital twins and related cyber-physical systems, which provide safe, high-fidelity simulators for training and validating Reinforcement Learning (RL) or Machine Learning (ML)-based schedulers before live execution, further enhancing reliability in dynamic HMLV environments (Liu et al., 2024). The convergence of edge computing with Manufacturing Execution Systems (MES), Enterprise Resource Planning (ERP) data streams, and Internet of Things (IoT) sensors supports near-instantaneous data fusion and rapid reallocation of resources, which are essential for responsive scheduling in high-mix variability contexts (Wang et al., 2021; Hamdan et al., 2020; Wu et al., 2021).

Real-time scheduling at the edge relies on a layered infrastructure that integrates data from sensors, shop-floor software, and enterprise systems while leveraging edge-specific orchestration and communication patterns. Middleware that bridges MES/ERP with edge scheduling engines enables end-to-end visibility and control across planning and execution layers, a critical enabler for low-latency decision loops in Industry 4.0 settings (Rao et al., 2020; Godfrey et al., 2024). Edge architectures, from lightweight edge nodes to edge-cloud continuums are designed to deliver inference at millisecond-to-second scales and manage burst workloads through scalable, distributed processing, as summarized in surveys of edge computing architectures and IoT deployments (Wang et al., 2021; Hamdan et al., 2020). Complementary studies highlight the role of edge-enabled reinforcement learning for dynamic task offloading and scheduling in IoT-empowered factories, where privacy-preserving and communication-efficient mechanisms are increasingly important for real-time operations (Wu et al., 2021). Collectively, these infrastructures edge computing platforms, digital twins, and MES/ERP integration, provide the data tempo and computational locality required for robust, real-time scheduling in HMLV manufacturing (Zhou et al., 2024; Liu et al., 2024; Godfrey et al., 2024).

### 4.3 Digital Twins in AI-Powered Scheduling

Digital twins (DTs) provide virtual replicas of the shop floor that mirror physical assets, processes, and workflows, enabling high-fidelity simulation environments for AI-based scheduling research and deployment (Lu et al., 2020; Fuller et al., 2020). In manufacturing contexts, these replicas support simulation-based training of AI models, allowing researchers and practitioners to pre-train and test scheduling policies, dispatch rules, and reinforcement learning agents in a risk-free setting before live execution (Fuller et al., 2020). DTs also facilitate extensive what-if analyses and scenario exploration from demand shifts to disruption events, so that scheduling policies can be iteratively refined and validated against realistic production trajectories (Okeagu et al., 2024; Sit and Lee, 2023). The literature on low-volume high-mix (LVHM) production and mass personalization demonstrates DT-driven optimization frameworks for job allocation and sequencing, underscoring DTs as essential platforms for evaluating new AI-enabled schedulers in complex environments (Sit and Lee, 2023; Chen et al., 2022).

Beyond policy testing, digital twins enable closed-loop learning and continuous improvement by pairing high-fidelity virtual models with real-time data streams from sensors, MES (Manufacturing Execution Systems)/ERP systems, and enterprise networks, thereby supporting iterative reinforcement learning refinement and data-driven decision-making (Lu et al., 2020; Wang et al., 2021). Integrating DTs with scheduling platforms supports end-to-end validation from model training in the DT environment to policy implementation on the shop floor, as shown in shipbuilding and automotive-focused DT scheduling studies (Wang et al., 2021). Moreover, DT architectures often described as a multi-dimensional fusion of physical, digital, and human layers offer a coherent framework for knowledge-driven scheduling where what-if testing, safety checks, and policy evaluation can proceed in parallel with live operations, guided by frameworks that articulate DT-driven scheduling in Industry 4.0 settings. Collectively, DTs anchor AI-powered scheduling by providing safe, scalable, and data-rich environments that accelerate learning, testing, and deployment across HMLV manufacturing ecosystems (Fuller et al., 2020; Okeagu et al., 2024).

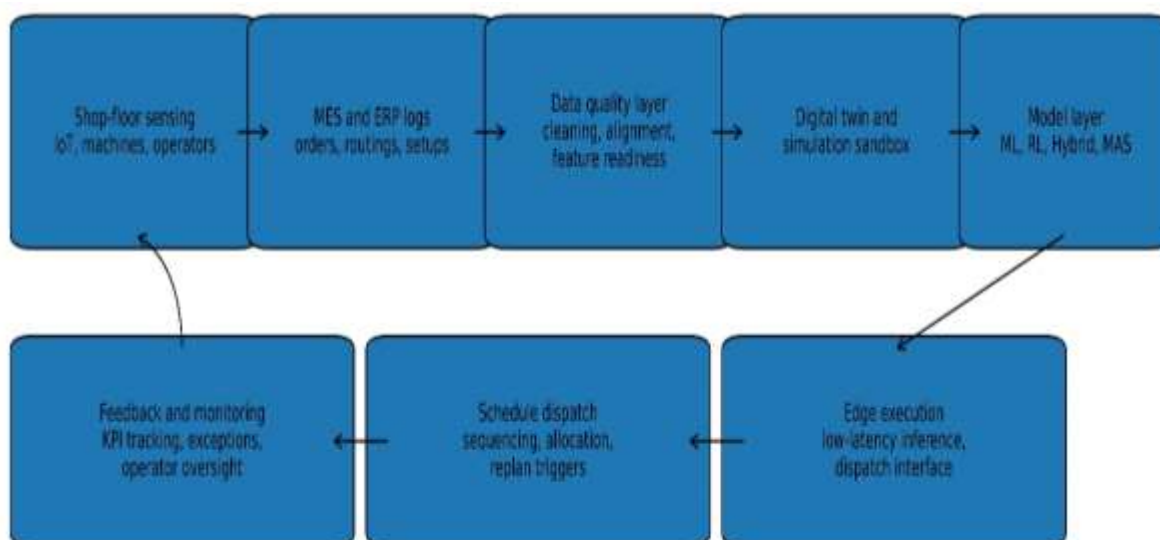


Figure 2. End-To-End Integration architecture for AI-Powered Scheduling in HMLV Production

Figure 2 is a diagram that illustrates closed-loop integration of sensing, enterprise data, digital twins, AI models, edge execution, and feedback. It shows scheduling functions as a continuous cyber-physical control process embedded within digital manufacturing infrastructure.

## 5. Comparative Analysis of AI Approaches

### 5.1 Performance Metrics in HMLV Scheduling

Key metrics used to evaluate scheduling approaches include:

- i. **Makespan** (total time to complete all jobs). In HMLV scheduling, makespan remains a primary objective for assessing how well AI-based methods can coordinate diverse job orders and reconfigurations. Recent multi-objective deep reinforcement learning (DRL)-inspired approaches have demonstrated improved makespan relative to traditional baselines in flexible flow shop settings and hybrid optimization methods have achieved notable reductions in makespan on unrelated job environments. These results collectively indicate that modern AI-driven schedulers can substantially shorten total production time by learning effective sequencing and resource allocations under variability.
- ii. **Tardiness** (delays beyond due dates). Tardiness is a critical concern in HMLV contexts where mass personalization and tight due windows collide with unpredictable disruptions. Recent studies utilizing DRL and hybrid optimization report improvements in tardiness, including reductions via dynamic decision policies and robust scheduling under random disturbances (Wu et al., 2021), and strategies that consider tardiness penalties alongside energy optimization when balancing objective trade-offs. The convergence of learning-based and robust optimization approaches presents a promising pathway to simultaneously manage makespan and tardiness in highly variable manufacturing settings.
- iii. **Machine utilization** (efficiency of resource use). Efficient machine utilization is essential in HMLV plants with scarce and reconfigurable equipment. Evidence from AI-enhanced optimization techniques shows improved utilization alongside makespan optimization, with genetic algorithm (GA)-based tools enhancing machine usage in flexible systems and particle swarm optimization (PSO)-based hybrids demonstrating superior utilization in flow-shop contexts (Onyeka et al., 2024). Energy efficiency considerations often accompany utilization objectives, with multi-objective approaches explicitly balancing throughput, utilization, and energy use to achieve practical, sustainable schedules. Together, these results highlight the capacity of AI-based approaches to enhance machine utilization without sacrificing other performance targets.
- iv. **Flexibility** (ability to adapt to disruptions). Flexibility is central to HMLV, and multi-agent systems (MAS), DRL, and hybrid AI methods have shown promise in adapting to disturbances such as random breakdowns and late changes in demand. Robust scheduling under uncertainty and adaptive heuristics have demonstrated improved resilience over static policies, with DRL-based and hybrid approaches providing real-time adaptation and better responsiveness to disruptions (Ojsteršek et al., 2020; Sajadi et al., 2019). These findings underscore that flexible, learning-enabled schedulers can rapidly reallocate resources and replan sequences, maintaining performance under nonstationary conditions typical of HMLV manufacturing.

### 5.2 Case Studies from Literature

Case studies demonstrate the application of AI in various domains:

- i. **Aerospace: RL for assembly line scheduling.** In aerospace manufacturing, reinforcement learning (RL) has emerged as a viable approach to assembly-line-like scheduling where dynamic sequencing and tight timelines prevail (Chang et al., 2022). Deep reinforcement learning (DRL) has been applied to dynamic flexible job shop scheduling with random job arrivals, illustrating how an autonomous agent can adapt dispatch decisions in the face of unpredictable aerospace-order streams (Chang et al., 2022). The aerospace domain also emphasizes human factors and cognitive aspects in scheduling, underscoring the need for RL policies that harmonize with operators and organizational workflows on the shop floor. Complementary aerospace-focused scheduling work demonstrates the value of intelligent

- optimization in complex aerospace environments, including process planning and scheduling for aerospace components that mirror the intricacy of assembly-line planning in aircraft manufacturing (Yang et al., 2024).
- ii. Automotive prototyping: genetic algorithms for dynamic job sequencing. Automotive prototyping presents a quintessential High-Mix Low-Volume (HMLV) setting where genetic algorithms (GAs) are leveraged to manage dynamic job sequencing under frequent design and variant changes. Multi-objective GA studies address trade-offs among makespan, tardiness, and energy use in flexible job-shop contexts, offering relevant insights for automotive prototyping lines that must rapidly re-sequence tasks as part variants evolve. Earlier GA work in cellular manufacturing demonstrates the benefits of population-based search with elite strategies for handling sequence-dependent setups and diverse product families, a common scenario in automotive prototyping. Hybrid GA approaches tailored to compatibility constraints in flow-shop-like sequences further validate the applicability of GA-based scheduling to automotive prototyping environments where assembly-stage and part variants drive sequencing decisions (Muñoz-Sánchez et al., 2024). Empirical results also show GA-enhanced tool utilization and makespan improvements in flexible manufacturing, highlighting the practical gains achievable in rapid-iteration automotive prototyping settings (Onyeka et al., 2024), with energy-aware GA variants illustrating sustainable scheduling considerations relevant to automotive contexts.
  - iii. Semiconductor fabs: ML for predictive bottleneck detection. In semiconductor manufacturing, ML and related data-driven methods are increasingly deployed to detect and mitigate bottlenecks within fab scheduling. Self-supervised and reinforcement learning-based scheduling studies in semiconductor environments demonstrate how learning-based policies can anticipate delays, reallocate scarce resources, and improve throughput under highly stochastic process flows. Systematic literature reviews reinforce the relevance of AI-enabled scheduling in real industrial settings, including semiconductor contexts, and document the positive impacts of ML-driven scheduling on throughput and reliability in fabs (Gallo et al., 2023). Additionally, graph-based and maintenance-focused ML models are explored to predict failures and maintenance needs, enabling proactive bottleneck management through timely resource reallocation and schedule adaptation (Wang et al., 2021).

### 5.3 Trade-offs: Accuracy vs. Interpretability vs. Computation Time

In AI-powered scheduling for HMLV (High-Mix Low-Volume) environments, a central tension exists among accuracy, interpretability, and computation time. Deep reinforcement learning (DRL) and other high-capacity models often deliver superior predictive and scheduling performance under uncertainty, but their black-box nature hinders interpretability, complicating trust, governance, and regulatory compliance in manufacturing contexts (Minh et al., 2021; Onyeka and Emeka, 2025). Explainable AI (XAI) frameworks and reviews consistently argue that performance gains commonly come at the cost of opaque reasoning, prompting a trade-off analysis where stakeholders demand understandable rationale alongside high-precision decisions (Minh et al., 2021; Kostopoulos et al., 2024). At the same time, literature on scheduling in the industry emphasizes that explainability can enhance adoption and collaboration with human operators, even if it restricts the complexity of the underlying models (Habiba et al., 2024; Gallo et al., 2023). Consequently, practitioners frequently adopt hybrid or two-stage approaches that balance accuracy with explainability to support robust decision-making under real-world constraints (U-Dominic et al., 2025; Thakre et al., 2023).

Computation time adds another critical dimension to these trade-offs. Real-time or near-real-time scheduling in HMLV settings requires rapid inference, which can be challenging for DRL and other large neural models that demand substantial training and runtime computation (Ngwu et al., 2025; Ezeji et al., 2024). To reconcile speed with quality, researchers propose hybrid AI paradigms,

where learning components guide or prune optimization searches, effectively reducing computation while preserving solution quality (U-Dominic et al., 2025; Ezeji et al., 2024). Additional work demonstrates how process and scheduling decisions can be partitioned into faster, interpretable components or staged learning pipelines to achieve timely responses without sacrificing long-horizon objectives (Ono and Okpala, 2025). Across these strands, the literature converges on a practical takeaway: selecting AI scheduling approaches in HMLV should explicitly trade off accuracy, interpretability, and computation time in line with the decision horizon, risk tolerance, and governance requirements of the manufacturing context (Minh et al., 2021; Kostopoulos et al., 2024; Ngwu et al., 2025; Ono and Okpala, 2025).

## 6. Challenges and Open Research Questions

### 6.1 Data Scarcity and Quality Issues

High-Mix Low-Volume (HMLV) manufacturing often yields sparse datasets because many products are customized, batches are small, and demand signals are volatile, which complicates training robust machine learning (ML) models for scheduling. This scarcity is widely recognized in the literature on HMLV environments, including comprehensive reviews that emphasize variability, limited historical data, and the consequent challenges for data-driven scheduling (Gan et al., 2023). Furthermore, studies underscore information-content and data-quality concerns in low-volume production contexts, highlighting how these factors impact the efficiency of manufacturing processes (Nwamekwe and Chikwendu, 2025). Analyses of HMLV applications where data come from disparate sources such as shop-floor sensors, MES/ERP logs, and historical records illustrate how data fragmentation and inconsistent logging can degrade model performance and impede timely decision making on the shop floor (Sit and Lee, 2023; Ezeanyim et al., 2025; Kocsi et al., 2020). Collectively, these sources illustrate that data scarcity and quality gaps are fundamental barriers to deploying effective ML-based schedulers in HMLV settings (Gan et al., 2023; Nwamekwe and Chikwendu, 2025; Sit and Lee, 2023; Ezeanyim et al., 2025; Kocsi et al., 2020).

To address these data challenges, the literature identifies two principal strategies: transfer learning and synthetic data generation. Transferring knowledge learned in data-rich domains can assist in adapting models to HMLV tasks with limited local data, an approach that aligns with expert-knowledge transfer and cross-domain learning discussed in the context of high-mix scheduling problems. Additionally, digital twin and simulation-enabled data generation present scalable ways to produce diverse training scenarios and labels for reinforcement learning and predictive models, as demonstrated in digital twin-driven scheduling studies (Sit and Lee, 2023; Ezeanyim et al., 2025). Beyond digital twins, physics-informed and data-driven synthetic data generation offers a method to expand the effective training set when real data are scarce, as evidenced in applications of synthetic data that support robust learning under limited samples. Together, these approaches provide practical pathways to mitigate data scarcity and improve the reliability and generalizability of AI-powered scheduling in HMLV environments (Gan et al., 2023; Nwamekwe and Chikwendu, 2025; Sit and Lee, 2023; Ezeanyim et al., 2025; Kocsi et al., 2020).

### 6.2 Human–AI Collaboration in Scheduling Decisions

Operator trust is critical to the effective deployment of AI-powered scheduling in HMLV manufacturing; when trust in automated recommendations is misaligned with system performance, operators may underutilize or override schedules, undermining potential gains in responsiveness and throughput. Trust in automation is shaped by perceptions of usefulness, risk, and reliability, and calibrating this trust across diverse tasks and contexts is essential for durable human-machine collaboration in dynamic shop floors (Kohn et al., 2021). To bridge the gap between algorithmic recommendations and human decision-making, explainable AI tools and interpretable decision aids are needed, enabling transparent rationale, uncertainty disclosure, and auditability of scheduling

choices without sacrificing operational agility (Thakre et al., 2023). Reviews emphasize the trade-offs between accuracy and interpretability in AI-driven planning, underscoring that adoption hinges on explanations that practitioners can understand and act upon (Minh et al., 2021; Onyeka and Emeka, 2025).

Given these dynamics, human–AI collaboration in scheduling should embrace human-in-the-loop designs, adjustable levels of automation, and user-centred explanations that support governance and trust calibration across diverse operator populations (Nwamekwe et al., 2025). Practical approaches include presenting interpretable justifications and confidence signals, along with actionable overlays or rule-based levers that allow timely overrides in emergencies, enabling safe yet autonomous decision-making in routine and exception scenarios (Habiba et al., 2024; Thakre et al., 2023). Recognition that trust is not uniform across individuals or cultures necessitates explanation strategies that adapt to different user profiles and organizational contexts to maintain effective cooperation over time, including when automation errs or encounters novel disruptions. These considerations point to a research agenda where trust, transparency, and human oversight are integral components of AI-powered scheduling systems in HMLV environments (Minh et al., 2021; Kostopoulos et al., 2024; Thakre et al., 2023).

### 6.3 Scalability and Generalizability

Scaling AI-driven scheduling from laboratory or single-facility experiments to industrial-scale High-Mix Low-Volume (HMLV) systems remains a significant bottleneck. Reviews of AI-enabled scheduling in modern manufacturing highlight that while reinforcement learning (RL) and deep RL (DRL) can handle large state spaces and nonstationary demand, their performance often degrades when moving from controlled environments to real multi-plant deployments, revealing challenges in scalability and generalizability (Ngwu et al., 2025). Systematic literature reviews show that real-world adoption lags behind lab results due to issues related to integration, data quality, and orchestration across disparate facilities. This underscores the necessity for architectures that facilitate scalable deployment across plants (Gallo et al., 2023).

To achieve timely decision-making at scale, researchers have explored edge and distributed computing paradigms, which include hierarchical dispatchers for edge devices and edge–cloud data management frameworks. These help mitigate latency and bandwidth constraints while extending the geographic footprint of schedulers. Moreover, case studies in cloud manufacturing indicate that DRL and ML-based schedulers can scale with appropriate infrastructure but encounter persistent generalization issues when faced with novel task mixes and disruptions (Chen et al., 2022).

Addressing these challenges requires multi-pronged strategies that encompass data, algorithms, and infrastructure. Transfer learning and synthetic data generation are proposed as approaches to adapt data-hungry AI models to data-sparse HMLV contexts while diversifying training scenarios without incurring excessive data collection costs (Gan et al., 2023). Hybrid approaches that combine ML predictions with optimization techniques such as metaheuristics or ML-guided search provide practical routes for reducing search spaces while maintaining solution quality, thereby improving scalability in complex Job Shop Scheduling (JSP) and Flexible Job Shop Scheduling (FJSP) contexts (Lu et al., 2020). Decentralized scheduling through multi-agent systems (MAS) and agent-based coordination offers scalability via parallel decision-making and localized communication, which is particularly beneficial for modular or distributed HMLV lines (Okorocho et al., 2022). Finally, robust evaluation frameworks for scheduling under uncertainty support the assessment of generalizability across different disruption scenarios and plant heterogeneity, thereby informing more transferable policy design. Collectively, these directions indicate a research agenda aimed at developing scalable, generalizable AI-powered schedulers that emerge from integrated edge/cloud architectures, learning-with-optimization hybrids, and human-machine interfaces

conducive to deployment across diverse HMLV environments (Lu et al., 2020; Okorochoa et al., 2022; Ngwu et al., 2025; Gallo et al., 2023).

#### **6.4 Ethical and Sustainability Considerations**

AI models used for scheduling in HMLV environments are often energy-intensive, particularly during training and continued inference at scale, which raises sustainability concerns for manufacturing operations seeking to minimize environmental footprints (Ribeiro et al., 2020). Life-cycle assessment (LCA) frameworks provide a structured means to quantify environmental impacts, including energy consumption and emissions across the product and process life cycles, enabling researchers and practitioners to compare AI-enabled scheduling approaches not only on performance but also on environmental costs (Ribeiro et al., 2020). Recent work highlights the integration of sustainability considerations into digital transformation and Industry 4.0 initiatives, advocating for governance that accounts for energy use, waste, and social impacts as AI systems expand within manufacturing ecosystems. Collectively, these areas underscore the necessity to embed LCA-informed designs and energy-aware optimization into AI-based schedulers to prevent shifting environmental burdens across the value chain (Ribeiro et al., 2020).

Beyond energy metrics, ethical and social dimensions particularly fairness in workforce scheduling are increasingly recognized as essential for responsible AI deployment on the shop floor. Fairness concerns centre on how shifts, overtime, and task allocations are distributed among workers, affecting job satisfaction, fatigue, and morale while also influencing regulatory and reputational risk for manufacturers (Habiba et al., 2024). Explainable AI (XAI) tools and transparent decision-making rationales are seen as critical enablers for human-AI collaboration, helping operators understand, trust, and appropriately override or negotiate AI-generated schedules when necessary (Minh et al., 2021; Habiba et al., 2024). Trust calibration is also a focal point: miscalibrated reliance on automation can undermine effectiveness, underscoring the need for interpretable explanations, confidence signals, and governance mechanisms that accommodate diverse operator preferences and organizational cultures (Kohn et al., 2021; Nwamekwe et al., 2025). Thus, advancing ethical and sustainability considerations in AI-powered scheduling requires a holistic approach that combines energy-aware, life-cycle thinking with human-centered, transparent governance to ensure equitable and responsible deployment across HMLV settings (Minh et al., 2021; Habiba et al., 2024; Kohn et al., 2021; Nwamekwe et al., 2025).

### **7. Future Directions**

#### **7.1 Large Language Models (LLMs) for Scheduling Knowledge**

The integration of large language models (LLMs) into high-mix low-volume (HMLV) scheduling offers a compelling pathway to encode domain expertise directly into the planning loop. By leveraging ontology-driven and symbolically grounded representations, LLMs can capture expert heuristics, constraints, and routing rules as part of a schedule-aware knowledge base. For example, Ludwig et al. (2022), describe "Speech-to-Jobshop," an ontology-driven digital assistant that translates scheduling constraints and operations into executable planning steps, effectively encoding expert knowledge into natural-language interfaces that support simulation-based decision-making. Complementarily, approaches that fuse LLMs with symbolic reasoning, such as TRACE-CS, demonstrate how an LLM can process user requests into logical clauses while a symbolic engine provides verifiable constraints and explanations, enabling adaptive yet auditable scheduling decisions (Vasileiou and Yeoh, 2025). Real-world scheduling scenarios, including aviation-like dynamic sequencing, have also been explored with LLMs to handle complex constraint interactions and dynamic arrivals, underscoring the potential for LLMs to internalize scheduling heuristics and constraints in practice (Zhou et al., 2024). In parallel, studies applying LLMs to logistics and warehouse sequencing illustrate how language-informed reasoning can augment traditional search

with expert heuristics, offering practical templates for LLM-assisted genetic algorithm (GA)-based scheduling in distributed settings. Finally, planning-oriented works on code generation with LLMs indicate the feasibility of automatically producing constrained, policy-guided scheduling logic that can be validated and refined *in silico* before deployment.

Beyond encoding, the value of LLMs in scheduling hinges on their ability to support adaptive and explainable policies. The literature emphasizes that explanations are critical for trust, governance, and human-AI collaboration in scheduling contexts, guiding operators on when and how to intervene and ensuring that AI-driven recommendations align with operational policies (Minh et al., 2021; Habiba et al., 2024; Thakre et al., 2023). The synergy of LLMs with symbolic or rule-based components as exemplified by TRACE-CS and similar efforts facilitates interpretable scheduling rationales while preserving the flexibility of data-driven adaptation, a combination particularly suited to the nonstationary demands of HMLV environments (Thakre et al., 2023; Vasileiou and Yeoh, 2025). Moreover, studies on domain-specific applications including disruptive, multi-criteria flight scheduling and industrial logistics illustrate that LLMs can produce interpretable justifications and confidence signals that support human oversight and governance without sacrificing responsiveness (Zhou et al., 2024). Collectively, these directions suggest a future where LLMs function as scalable reservoirs of domain knowledge, constraints, and heuristics, integrated with symbolic and human-in-the-loop mechanisms to yield adaptive, auditable, and trustworthy scheduling in HMLV manufacturing (Minh et al., 2021; Thakre et al., 2023; Vasileiou and Yeoh, 2025; Zhou et al., 2024).

## 7.2 Federated Learning Across Manufacturing Sites

Federated learning (FL) provides a pathway for cross-site scheduling knowledge without exposing proprietary data, enabling collaborative AI training across multiple manufacturing sites while protecting intellectual property and enhancing model robustness (Yang et al., 2024; Yan et al., 2024). In scheduling contexts, FL has been applied in settings that integrate digital twins and distributed learning to share learned scheduling policies without data leakage, as demonstrated in literacy-based federated deep reinforcement learning for software-defined factories (Ahn et al., 2024). Personalized FL approaches, such as FedSeq, address client heterogeneity by adapting models to local data while working towards a shared global objective, which is particularly relevant for geographically dispersed plants with varying product mixes (Jang and Choi, 2024). Techniques that reduce communication burden, such as over-the-air federated averaging with privacy budgets, illustrate feasible methods to scale FL across factory networks without compromising data privacy (Yan et al., 2024). Furthermore, research into cloud- and edge-enabled scheduling has shown the potential of federated approaches for multi-site resource allocation and task sequencing under diverse constraints (George et al., 2024; Tang et al., 2023). Targeted studies are also exploring client selection strategies to optimize learning efficiency in heterogeneous manufacturing ecosystems (Yang et al., 2024). The promise of FL in maintaining privacy is supported by differential privacy frameworks and governance models that preserve confidentiality while enabling cross-site improvements (Yhdego et al., 2024; as well as FL-based multi-task scheduling mechanisms designed for trusted computing environments (Liu et al., 2024).

Despite the promise of FL, scaling it across manufacturing sites introduces important challenges related to heterogeneity, privacy budgets, and governance. Non-IID (independent and identically distributed) data distributions, variable device capabilities, and changing network conditions present difficulties in model generalization and convergence, necessitating ongoing research into personalized FL and robust aggregation tactics suitable for industrial scheduling (Jang and Choi, 2024; Yang et al., 2024; Ahn et al., 2024; George et al., 2024). Privacy constraints can affect accuracy; thus, methods like differential privacy and privacy-preserving aggregation are being researched to balance confidentiality with performance (Yan et al., 2024; Yhdego et al., 2024). Communication efficiency and synchronization remain practical challenges in multi-site

deployments, driving advancements such as OTA-FedAvg, selective updates, and edge-cloud collaboration to maintain real-time responsiveness (Yan et al., 2024; Liu et al., 2024). Collectively, the literature suggests a future where federated learning, enhanced by federated digital twins, edge-cloud orchestration, and robust governance frameworks, enables scalable, privacy-preserving, and effective scheduling across diverse highly flexible low-volume (HMLV) manufacturing networks (Yang et al., 2024; Ahn et al., 2024; Jang and Choi, 2024; Yan et al., 2024; George et al., 2024; Tang et al., 2023; Yang et al., 2024; Yhdego et al., 2024; Liu et al., 2024).

### 7.3 Integration with Autonomous Manufacturing Systems

Future AI schedulers are expected to be embedded directly within Industry 4.0/5.0 factories, coordinating in real time with autonomous machines and robots to orchestrate production across diverse lines and sites (Rossit et al., 2019). This integration is empowered by digital twins, which enable AI agents to train, validate, and refine scheduling policies in a risk-free virtual environment before deployment, supporting self-optimization, adaptive scheduling, and resilience in dynamic shop floors (Okpala et al., 2024; Serrano-Ruiz et al., 2021). The convergence of AI, digital twins, and autonomous systems is being explored in Industry 5.0 contexts, where autonomous agents and robots contribute to measurable gains in task completion, maintenance efficiency, and overall productivity, highlighting the viability of end-to-end autonomous scheduling in complex manufacturing ecosystems. Complementary work on cloud- and digital twin-enabled manufacturing platforms demonstrates how AI agents can be trained and deployed at scale, leveraging digital twins to bridge planning and execution while maintaining necessary human oversight (Nwamekwe and Igbokwe, 2024). Together, these developments point to a future where AI-powered schedulers operate as core components of autonomous manufacturing systems, coordinating with programmable robots, collaborative robots (cobots), and automated transport systems to realize mass personalization with high reliability (Rossit et al., 2019; Okpala et al., 2025; Nwamekwe et al., 2024).

The realization of autonomous, AI-driven scheduling across multiple plants hinges on architectures that support seamless machine-machine and human-machine coordination, governance, and safety. Decentralized, multi-agent coordination can distribute scheduling decisions to specialized agents embedded in robots, automated guided vehicles (AGVs), and production cells, enabling scalable, fault-tolerant orchestration in heterogeneous lines (Serrano-Ruiz et al., 2021). Effective integration will also demand robust explainability and governance to foster operator trust and appropriate intervention, leveraging explainable AI and human-in-the-loop designs to balance autonomy with oversight in routine and exception scenarios (Minh et al., 2021; Thakre et al., 2023; Kohn et al., 2021; Nwamekwe et al., 2025). Additionally, edge-cloud orchestration will be critical for achieving low-latency decision cycles while maintaining data privacy and interoperability across sites, as demonstrated by edge-distributed scheduling and digital twin-enabled frameworks that keep computation near the data sources and support real-time reconfiguration (Rossit et al., 2019). Altogether, these directions underscore a practical roadmap: embedding AI schedulers in autonomous factory architectures, empowering cross-site coordination with digital twins and multi-agent systems, and implementing governance and edge-cloud infrastructures that sustain reliable, auditable, and scalable scheduling in Industry 4.0/5.0 settings (Okpala et al., 2025; Nwamekwe et al., 2024; Serrano-Ruiz et al., 2021; Minh et al., 2021; Thakre et al., 2023; Kohn et al., 2021).

### 7.4 Towards Resilient and Self-Adaptive Scheduling

The long-term vision for AI-powered scheduling in HMLV manufacturing is the emergence of self-adaptive schedulers that continually learn from disruptions, evolving product mixes, and shifting customer demands. Digital twins (DTs) provide safe, high-fidelity environments for training, validating, and refining scheduling policies and reinforcement learning agents before live deployment, thereby enabling rapid adaptation to a wide range of disturbance scenarios and

contributing to resilient operations on dynamic shop floors (Fuller et al., 2020). Federated learning (FL) and federated digital twins further facilitate cross-site knowledge sharing of scheduling strategies without exposing proprietary data, addressing plant-to-plant heterogeneity while preserving data privacy and improving generalizability across distributed manufacturing networks (Ahn et al., 2024; George et al., 2024; Tang et al., 2023). Together with Industry 4.0/5.0-aligned architectures, edge-cloud orchestration, and autonomous systems, these approaches collectively enable scalable, self-improving scheduling that can accommodate diverse product mixes and evolving demand patterns (Rossit et al., 2019; Okpala et al., 2025).

Realizing resilient, self-adaptive scheduling requires architectures that support decentralized coordination, governance, and trust. Decentralized, multi-agent systems coordination and autonomous scheduling across cells, robots, and automated guided vehicles (AGVs) can yield scalable, fault-tolerant orchestration in heterogeneous HMLV environments, while DTs supply a sandbox for experimenting with risk-aware policies and validating safety-critical decisions before execution (Serrano-Ruiz et al., 2021; Rossit et al., 2019). To ensure responsible deployment, explainability and governance must accompany autonomy, providing interpretable justifications, confidence signals, and override mechanisms that sustain human oversight and trust in AI-driven decisions across diverse operator profiles and organizational cultures (Minh et al., 2021; Thakre et al., 2023; Kohn et al., 2021). Edge-cloud orchestration will be central to achieving low-latency decision cycles and scalable privacy-preserving collaboration across sites, with DT-driven learning enabling continuous improvement and rapid adaptation to nonstationary demand and disruptions (Rossit et al., 2019; Ahn et al., 2024; George et al., 2024). In this synthesis, future scheduling systems will integrate DTs, FL, multi-agent systems (MAS), and edge-cloud architectures to realize resilient, self-adaptive, and auditable scheduling across Industry 4.0/5.0 manufacturing ecosystems (Fuller et al., 2020; Ahn et al., 2024; Jang and Choi, 2024; George et al., 2024; Minh et al., 2021; Thakre et al., 2023).

## 8. Conclusion

AI-driven scheduling stands as a structural response to the operational realities of High-Mix Low-Volume production. The review demonstrates that data-driven models improve schedule quality by learning processing patterns, anticipating disruptions, and reallocating resources with greater timing accuracy than static rule systems. Reinforcement learning studies show adaptive sequencing under nonstationary demand, while hybrid optimization frameworks reduce search effort and improve feasibility in large scheduling spaces. Multi-agent coordination further strengthens resilience through localized negotiation and distributed decision execution. Together, these strands confirm that intelligent scheduling shifts planning from reactive control toward predictive orchestration across machines, products, and routing alternatives.

The analysis also establishes clear conditions for industrial deployment. Effective adoption depends on reliable data pipelines, integration with MES and digital twin platforms, and edge-enabled execution capable of near-real-time response. Persistent barriers include sparse historical data, limited transparency of deep models, and the computational demands of large-scale learning architectures. Future work therefore requires explainable decision layers, sustainability-aware optimization objectives, and continuous learning pipelines supported by federated and digital-twin infrastructures. Progress along these directions will position AI scheduling as a core coordination mechanism inside smart factories, strengthening operational stability, responsiveness, and long-term manufacturing competitiveness.

## REFERENCES

- Ahn, J., Yun, S., Kwon, J., and Kim, W. (2024). Literacy deep reinforcement learning-based federated digital twin scheduling for the software-defined factory. *Electronics*, 13(22), 4452. <https://doi.org/10.3390/electronics13224452>
- Castañé, G., Dolgui, A., Kousi, N., Meyers, B., Thevenin, S., Vyhmeister, E., ... and Östberg, P. (2022). The assistant project: ai for high level decisions in manufacturing. *International Journal of Production Research*, 61(7), 2288-2306. <https://doi.org/10.1080/00207543.2022.2069525>
- Chang, X., Jia, X., Fu, S., Hu, H., and Liu, K. (2022). Digital twin and deep reinforcement learning enabled real-time scheduling for complex product flexible shop-floor. *Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture*, 237(8), 1254-1268. <https://doi.org/10.1177/09544054221121934>
- Chen, Z., Zou, J., and Wang, W. (2022). Digital twin oriented multi-objective flexible job shop scheduling model and its hybrid particle swarm optimization. *Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture*, 237(8), 1269-1282. <https://doi.org/10.1177/09544054221121921>
- Chidiebube, I. N., Nwamekwe, C. O., Chukwuemeka, G. H., and Wilfred, M. (2025). OPTIMIZATION OF OVERALL EQUIPMENT EFFECTIVENESS FACTORS IN A FOOD MANUFACTURING SMALL AND MEDIUM ENTERPRISE. *Journal of Research in Engineering and Applied Sciences*, 10(1), 836-845.
- Emeka, U. C., Okpala, C., and Nwamekwe, C. O. (2025). CIRCULAR ECONOMY PRINCIPLES IMPLEMENTATION IN ELECTRONICS MANUFACTURING: WASTE REDUCTION STRATEGIES IN CHEMICAL MANAGEMENT. *International journal of industrial and production engineering*, 3(2), 29-42.
- Ezeanyim, O. C., Ewuzie, N. V., Aguh, P. S., Nwabueze, C. V., and Nwamekwe, C. O. (2025). Effective Maintenance of Industrial 5-Stage Compressor: A Machine Learning Approach. *Gazi University Journal of Science Part A: Engineering and Innovation*, 12(1), 96-118. <https://dergipark.org.tr/en/pub/gujsa/issue/90827/1646993>
- Ezeanyim, O.C., Nwabunwanne, E.C., Igbokwe, N.C. and Nwamekwe, C.O. (2025b) 'Patient flow and service efficiency in public hospitals: data-driven approaches, strategies, challenges, and future directions', *Journal Health of Indonesian*, 3(02), pp. 104–124. <https://doi.org/10.58471/health.v3i02.228>
- Ezeji, I., Adigun, M., and Oki, O. (2024). Computational complexity in explainable decision support system: a review. *Journal of Intelligent and Fuzzy Systems*. <https://doi.org/10.3233/jifs-219407>
- Fuller, A., Fan, Z., Day, C., and Barlow, C. (2020). Digital twin: enabling technologies, challenges and open research. *Ieee Access*, 8, 108952-108971. <https://doi.org/10.1109/access.2020.2998358>
- Gallo, M., Mazzuto, G., Ciarapica, F., and Bevilacqua, M. (2023). Artificial intelligence to solve production scheduling problems in real industrial settings: systematic literature review. *Electronics*, 12(23), 4732. <https://doi.org/10.3390/electronics12234732>
- Gan, Z., Musa, S., and Yap, H. (2023). A review of the high-mix, low-volume manufacturing industry. *Applied Sciences*, 13(3), 1687. <https://doi.org/10.3390/app13031687>
- George, N., Anoop, B., and Vijayan, V. (2024). Multicriteria generalized regressive neural federated learning for cloud computing task scheduling and resource allocation. *E3s Web of Conferences*, 529, 04017. <https://doi.org/10.1051/e3sconf/202452904017>
- Godfrey, O. C., Chukwuemeka, G. H., Edith, M. C., and Daniel, E. C. (2024). Stochastic process assessment for XP600 printhead failures: A Weibull method study. *UNIZIK Journal of Engineering and Applied Sciences*, 3(1), 445-456.
- Habiba, U., Habib, M., Bogner, J., Fritzsche, J., and Wagner, S. (2024). How do ml practitioners perceive explainability? an interview study of practices and challenges. *Empirical Software Engineering*, 30(1). <https://doi.org/10.1007/s10664-024-10565-2>
- Hamdan, S., Ayyash, M., and Almajali, S. (2020). Edge-computing architectures for internet of things applications: a survey. *Sensors*, 20(22), 6441. <https://doi.org/10.3390/s20226441>
- Igbokwe, N. C., and Nwamekwe, C. O. (2025). Application of Machine Learning in Predicting Emergency Obstetric Cases in Sub-Saharan Africa: An Early Appraisal. *International Journal of Industrial Engineering, Technology and Operations Management*, 3(1), 13-22.

- Igbokwe, N. C., Christiana, C., Nweke, C. O. N., and Onyeka, C. (2025). Data-Driven Solutions for Shuttle Bus Travel Time Prediction: Machine Learning Model Evaluation at Nnamdi Azikiwe University. *African Journal of Computing, Data Science and Informatics (AJCDSI)*, 1(1), 31-55.
- Igbokwe, N. C., Okeagu, F. N., Onyeka, N. C., Onwuliri, J. B., and Godfrey, O. C. (2024). MACHINE LEARNING-DRIVEN MAINTENANCE COST OPTIMIZATION: INSIGHTS FROM A LOCAL INDUSTRIAL COMPRESSOR CASE STUDY. *Jurnal Inovasi Teknologi dan Edukasi Teknik*, 4(11), 2.
- Igbokwe, N.C., Emmanuel, U.N. and Nwamekwe, C.O. (2025) ‘Advances in post-harvest fish processing: an appraisal of traditional and modern smoking techniques for improved quality and efficiency’, *Jurnal Integrasi Dan Harmoni Inovatif Ilmu-Ilmu Sosial*, 5 (9), pp. 1-13. <https://philarchive.org/rec/IGBAIP>
- Jang, J. and Choi, B. (2024). Fedseq: personalized federated learning via sequential layer expansion in representation learning. *Applied Sciences*, 14(24), 12024. <https://doi.org/10.3390/app142412024>
- Julaiti, J., Oh, S., DAS, D., and Kumara, S. (2022). Stochastic parallel machine scheduling using reinforcement learning. *Journal of Advanced Manufacturing and Processing*, 4(4). <https://doi.org/10.1002/amp2.10119>
- Kocsi, B., Matonya, M., Pusztai, L., and Budai, I. (2020). Real-time decision-support system for high-mix low-volume production scheduling in industry 4.0. *Processes*, 8(8), 912. <https://doi.org/10.3390/pr8080912>
- Kohn, S., Visser, E., Wiese, E., Lee, Y., and Shaw, T. (2021). Measurement of trust in automation: a narrative review and reference guide. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.604977>
- Kostopoulos, G., Davrazos, G., and Kotsiantis, S. (2024). Explainable artificial intelligence-based decision support systems: a recent review. *Electronics*, 13(14), 2842. <https://doi.org/10.3390/electronics13142842>
- Liu, X., Xie, T., Wu, B., Li, M., and Li, H. (2024). Construction of digital twin workshop integrated with edge computing and deep learning. *Journal of Physics Conference Series*, 2816(1), 012008. <https://doi.org/10.1088/1742-6596/2816/1/012008>
- Lu, Y., Liu, C., Wang, K., Huang, H., and Xu, X. (2020). Digital twin-driven smart manufacturing: connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing*, 61, 101837. <https://doi.org/10.1016/j.rcim.2019.101837>
- Ludwig, S., Aschenbrenner, D., Scharle, M., Klessig, H., Karrenbauer, M., Wu, H., ... and Fitzek, F. (2022). Reference network and localization architecture for smart manufacturing based on 5g.. <https://doi.org/10.48550/arxiv.2204.00581>
- Minh, D., Wang, H., Li, Y., and Nguyen, T. (2021). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, 55(5), 3503-3568. <https://doi.org/10.1007/s10462-021-10088-y>
- Muñoz-Sánchez, R., Martínez-Salazar, I., González-Velarde, J., and Solis, Y. (2024). Two hybrid flow shop scheduling lines with assembly stage and compatibility constraints. *Plos One*, 19(6), e0304119. <https://doi.org/10.1371/journal.pone.0304119>
- Ngwu, C., Liu, Y., and Wu, R. (2025). Reinforcement learning in dynamic job shop scheduling: a comprehensive review of ai-driven approaches in modern manufacturing. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-025-02585-6>
- Normand, A. and Bradley, T. (2024). An experimental investigation of lean six sigma philosophies in a high-mix low-volume manufacturing environment. *Plos One*, 19(5), e0299498. <https://doi.org/10.1371/journal.pone.0299498>
- Nwamekwe C. O., Ewuzie Nnamdi Vitalis, Igbokwe Nkemakonam Chidiebube, and Nwabueze Chibuzo Victoria. (2025). Evaluating Advances in Machine Learning Algorithms for Predicting and Preventing Maternal and Foetal Mortality in Nigerian Healthcare: A Systematic Approach. *International Journal of Industrial and Production Engineering*, 3(1), 1-15. <https://journals.unizik.edu.ng/ijipe/article/view/5161>
- Nwamekwe C. O., Ezeanyim O. C., and Igbokwe N. C. (2025). Resilient Supply Chain Engineering in the Era of Disruption: An Appraisal. *International Journal of Innovative Engineering, Technology and Science (IJIETS)*, 9(1), 11-23. <https://hal.science/hal-05061524/>
- Nwamekwe, C. O., and Chikwendu, O. C. (2025). Circular economy strategies in industrial engineering: From theory to practice. *International Journal of Multidisciplinary Research and Growth Evaluation*, 6(1):

- 1773-1782. [https://www.allmultidisciplinaryjournal.com/uploads/archives/20250212103754\\_MGE-2025-1-288.1.pdf](https://www.allmultidisciplinaryjournal.com/uploads/archives/20250212103754_MGE-2025-1-288.1.pdf)
- Nwamekwe, C. O., and Nwabunwanne, E. C. (2025). Immersive Digital Twin Integration in the Metaverse for Supply Chain Resilience and Disruption Management. *Journal of Engineering Research and Applied Science*, 14(1), 95-105.
- Nwamekwe, C. O., and Igbokwe, N. C. (2024). Supply Chain Risk Management: Leveraging AI for Risk Identification, Mitigation, and Resilience Planning. *International Journal of Industrial Engineering, Technology and Operations Management*, 2(2), 41–51. <https://doi.org/10.62157/ijietom.v2i2.38>
- Nwamekwe, C. O., Chidiebube, I. N., Godfrey, O. C., Celestine, N. E., and Sunday, A. P. (2025). Resilience and Risk Management in Social Robot Systems: An Industrial Engineering Perspective. *Culture education and technology research (Cetera)*, 2(2), 1-12.
- Nwamekwe, C. O., Chidiebube, I. N., Godfrey, O. C., Celestine, N. E., and Aguh, P. S. (2025). Human-Robot Collaboration in Industrial Engineering: Enhancing Productivity and Safety. *Journal of Industrial Engineering and Management Research*, 6(5), 1-20.
- Nwamekwe, C. O., Chinwuko, C. E. and Mgbemena, C. E. (2020). Development and Implementation of a Computerised Production Planning and Control System. *UNIZIK Journal of Engineering and Applied Sciences*, 17(1), 168-187. <https://journals.unizik.edu.ng/ujeas/article/view/1771>
- Nwamekwe, C. O., Edokpia, R. O., and Igbinosa, E. C. (2025). Exploring the Role of Artificial Intelligence in Enhancing Lean Manufacturing and Six Sigma for Smart Factories. *International Journal of Industrial Engineering, Technology and Operations Management*, 3(1), 1-12.
- Nwamekwe, C. O., Ewuzie, N. V., Okpala, C. C., Ezeanyim, C., Nwabueze, C. V., Nwabunwanne, E. C. (2025). Optimizing Machine Learning Models for Soil Fertility Analysis: Insights from Feature Engineering and Data Localization. *Gazi University Journal of Science Part A: Engineering and Innovation*, 12(1), 36-60. <https://dergipark.org.tr/en/pub/gujisa/issue/90827/1605587>
- Nwamekwe, C. O., Ewuzie, N.V., Igbokwe, N. C., Nwabunwanne, E. C., and Ono, C. G. (2025). Digital Twin-Driven Lean Manufacturing: Optimizing Value Stream Flow. *Letters in Information Technology Education (LITE)*, 8 (1), pp.1-13. <https://hal.science/hal-05127340/>
- Nwamekwe, C. O., Nwabunwanne, E. C., Okeagu, F. N., and Ono, C. G. (2025). Lean Manufacturing Principles in the Design and Production of Social Robots. *International Journal of Industrial Engineering, Technology and Operations Management*, 3(1), 23-34.
- Nwamekwe, C. O., Okpala, C. C., and Nwabunwanne, E. C. (2025). Design Principles and Challenges in Achieving Zero-Energy Manufacturing Facilities. *Journal of Engineering Research and Applied Science*, 14(1), 1-21.
- Nwamekwe, C. O., Okpala, C. C., and Okpala, S. C., (2024). Machine Learning-Based Prediction Algorithms for the Mitigation of Maternal and Fetal Mortality in the Nigerian Tertiary Hospitals. *International Journal of Engineering Inventions*, 13(7), PP: 132-138. <https://www.ijejournal.com/papers/Vol13-Issue7/1307132138.pdf>
- Nwamekwe, C., Ewuzie, N., Igbokwe, N., Okpala, C., and U-Dominic, C. (2024). Sustainable Manufacturing Practices in Nigeria: Optimization and Implementation Appraisal. *Journal of Research in Engineering and Applied Sciences*, 9(3). <https://qtanalytics.in/journals/index.php/JREAS/article/view/3967>
- Ojsteršek, R., Acko, B., and Buchmeister, B. (2020). Simulation study of a flexible manufacturing system regarding sustainability. *International Journal of Simulation Modelling*, 19(1), 65-76. <https://doi.org/10.2507/ijstimm19-1-502>
- Okeagu, F., Nwamekwe, C., and Nnamani, B. (2024). Challenges and Solutions of Industrial Development in Anambra State, Nigeria. *Iconic Research and Engineering Journals*, 7(11), 467-472. <https://www.irejournals.com/formatedpaper/1705825.pdf>
- Okorochoa, I. T., Chinwuko, C. E., Mgbemena, C. O., Godfrey, O. C., and Mgbemena, C. E. (2022). Production optimization using gas lift incorporated with artificial neural network. *UNIZIK Journal of Engineering and Applied Sciences*, 21(1), 842-858.
- Okpala C. C., Chukwudi Emeka Udu, and Charles Onyeka Nwamekwe. (2025). Sustainable HVAC Project Management: Strategies for Green Building Certification. *International Journal of Industrial and Production Engineering*, 3(2), 14-28. <https://journals.unizik.edu.ng/ijipe/article/view/5595>.
- Okpala, C. C., Ezeanyim, O. C., and Nwamekwe, C. O. (2024). The Implementation of Kaizen Principles in Manufacturing Processes: A Pathway to Continuous Improvement. *International Journal of*

- Engineering Inventions, 13(7), 116-124. <https://www.ijejournal.com/papers/Vol13-Issue7/1307116124.pdf>
- Okpala, C. C., Udu, C. E., and Nwamekwe, C. O. (2025). Artificial Intelligence-Driven Total Productive Maintenance: The Future of Maintenance in Smart Factories. *International Journal of Engineering Research and Development (IJERD)*, (21)1, 68-74. <https://www.ijerd.com/paper/vol21-issue1/21016874.pdf>
- Ono, C. G. and Okpala, C. C. (2025). Smart and Resilient Agriculture for Sustainable Food Systems under Climate Change: Global Lessons for Food Security. *International Journal of Engineering Research and Development*, 21(12), 111-123
- Onyeka, N. C., and Emeka, N. (2025). Circular Economy and Zero-Energy Factories: A Synergistic Approach to Sustainable Manufacturing. *Journal of Research in Engineering and Applied Sciences*, 10(1), 829-835.
- Onyeka, N. C., Vitalis, E. N., Chidiebube, I. N., U-Dominic, C. M., and Chibuzo, N. (2024). Adoption of Smart Factories in Nigeria: Problems, Obstacles, Remedies and Opportunities. *International journal of industrial and production engineering*, 2(2), 68-81. <https://journals.unizik.edu.ng/ijipe/article/view/4167>
- Pantano, M. (2022). Capability-based frameworks for industrial robot skills: a survey. <https://doi.org/10.48550/arxiv.2203.00538>
- Rao, M., Lynch, L., Coady, J., Toal, D., and Newe, T. (2020). Integration of an mes and aiv using a labview middleware scheduler suitable for use in industry 4.0 applications. *Applied Sciences*, 10(20), 7054. <https://doi.org/10.3390/app10207054>
- Ribeiro, I., Matos, F., Jacinto, C., Salman, H., Cardeal, G., Carvalho, H., ... and Peças, P. (2020). Framework for life cycle sustainability assessment of additive manufacturing. *Sustainability*, 12(3), 929. <https://doi.org/10.3390/su12030929>
- Rossit, D., Tohmé, F., and Frutos, M. (2019). A data-driven scheduling approach to smart manufacturing. *Journal of Industrial Information Integration*, 15, 69-79. <https://doi.org/10.1016/j.jii.2019.04.003>
- Sajadi, S., Alizadeh, A., Zandieh, M., and Tavan, F. (2019). Robust and stable flexible job shop scheduling with random machine breakdowns: multi-objectives genetic algorithm approach. *International Journal of Mathematics in Operational Research*, 14(2), 268. <https://doi.org/10.1504/ijmor.2019.097759>
- Sajadieh, S. and Noh, S. (2025). From simulation to autonomy: reviews of the integration of artificial intelligence and digital twins. *International Journal of Precision Engineering and Manufacturing-Green Technology*. <https://doi.org/10.1007/s40684-025-00750-z>
- Serrano-Ruiz, J., Mula, J., and Poler, R. (2021). Smart digital twin for zdm-based job-shop scheduling., 510-515. <https://doi.org/10.1109/metroind4.0iot51437.2021.9488473>
- Sit, S. and Lee, C. (2023). Design of a digital twin in low-volume, high-mix job allocation and scheduling for achieving mass personalization. *Systems*, 11(9), 454. <https://doi.org/10.3390/systems11090454>
- Tahmina, T., Garcia, M., Geng, Z., and Bidanda, B. (2022). A survey of smart manufacturing for high-mix low-volume production in defense and aerospace industries., 237-245. [https://doi.org/10.1007/978-3-031-18326-3\\_24](https://doi.org/10.1007/978-3-031-18326-3_24)
- Tang, Y., Du, G., Long, J., and Zhao, Y. (2023). Integration of online deep reinforcement learning and federated learning: a dynamic energy management strategy for industrial applications. *Iet Communications*, 17(19), 2162-2177. <https://doi.org/10.1049/cmu2.12686>
- Thakre, P., Sahu, P., Soni, P., Bisen, D., and Sahu, A. (2023). Evaluating transparency and explainability in ai-driven planning and scheduling: a comprehensive literature review. *International Journal of Innovative Research in Science Engineering and Technology*, 12(06), 9008-9014. <https://doi.org/10.15680/ijirset.2023.1206155>
- U-Dominic, C. M., Orji, I. J., Nkemakonam, C. I., Onyeka, N. C., and Nwifo, M. A. (2025). A Decision Methodology for Six-Sigma Implementation in the Nigerian Small and Medium Scale Enterprise (SME). *Unizik Journal of Technology, Production and Mechanical Systems*, 5(1), 186-202.
- Vasileiou, S. and Yeoh, W. (2025). Trace-cs: a synergistic approach to explainable course scheduling using llms and logic. *Proceedings of the Aaai Conference on Artificial Intelligence*, 39(28), 29706-29708. <https://doi.org/10.1609/aaai.v39i28.35374>
- Vitalis, E. N., Nwamekwe, C. O., Chidiebube, I. N., Chibuzo, N., Nwabunwanne, E. C., and Ono, C. G. (2024). APPLICATION OF MACHINE-LEARNING-BASED HYBRID ALGORITHM FOR

- PRODUCTION FORECAST IN TEXTILE COMPANY. *Jurnal Inovasi Teknologi dan Edukasi Teknik*, 4(12), 1-9.
- Wang, L., Pan, Z., and Wang, J. (2021). A review of reinforcement learning based intelligent optimization for manufacturing scheduling. *Complex System Modeling and Simulation*, 1(4), 257-270. <https://doi.org/10.23919/csms.2021.0027>
- Wu, J., Zhang, G., Nie, J., Peng, Y., and Zhang, Y. (2021). Deep reinforcement learning for scheduling in an edge computing-based industrial internet of things. *Wireless Communications and Mobile Computing*, 2021(1). <https://doi.org/10.1155/2021/8017334>
- Xu, L., Yu, C., Wu, B., and Gao, M. (2024). A hybrid genetic algorithm for ground station scheduling problems. *Applied Sciences*, 14(12), 5045. <https://doi.org/10.3390/app14125045>
- Yan, N., Wang, K., Pan, C., Chai, K., Shu, F., and Wang, J. (2024). Over-the-air federated averaging with limited power and privacy budgets. *Ieee Transactions on Communications*, 72(4), 1998-2013. <https://doi.org/10.1109/tcomm.2023.3338738>
- Yang, D., Shu, X., Yu, Z., Lu, G., Ji, S., Wang, J., ... and He, K. (2024). Dynamic flexible job shop scheduling based on deep reinforcement learning. *Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture*. <https://doi.org/10.1177/09544054241272855>
- Yhdego, T., Wang, H., Chi, H., and Yu, Z. (2024). Ontology-guided data sharing and federated quality control with differential privacy in additive manufacturing. *Journal of Computing and Information Science in Engineering*, 25(1). <https://doi.org/10.1115/1.4067086>
- Zanchettin, A. (2021). Robust scheduling and dispatching rules for high-mix collaborative manufacturing systems. *Flexible Services and Manufacturing Journal*, 34(2), 293-316. <https://doi.org/10.1007/s10696-021-09406-x>
- Zhang, M., Lü, Y., Hu, Y., Amaitik, N., and Xu, Y. (2022). Dynamic scheduling method for job-shop manufacturing systems by deep reinforcement learning with proximal policy optimization. *Sustainability*, 14(9), 5177. <https://doi.org/10.3390/su14095177>
- Zhao, A., Liu, P., Gao, X., Huang, G., Yang, X., Ma, Y., ... and Li, Y. (2022). Data-mining-based real-time optimization of the job shop scheduling problem. *Mathematics*, 10(23), 4608. <https://doi.org/10.3390/math10234608>
- Zhou, W., Wang, J., Zhu, L., Wang, Y., and Ji, Y. (2024). Flight arrival scheduling via large language model. *Aerospace*, 11(10), 813. <https://doi.org/10.3390/aerospace11100813>