ISSN 2683-345X
journal homepage: http://ijiemjournal.uns.ac.rs/

International Journal of Industrial Engineering and Management

2025 / ARTICLE-IN-PRESS / 389

Original research article

Multi-objective Optimization Framework for Energy Efficiency and Production Scheduling in Smart Manufacturing Using Reinforcement Learning and Digital Twin Technology Integration

B. Madaminov[a,*] 0009-0002-2806-691X, S. Saidmurodov[b] 0009-0004-7377-0249, E. Saitov[c] 0000-0001-7691-6849, D. Jumanazarov[d,e] 0000-0002-5445-6818, A. M. Alsayah[f] 0009-0005-8122-8182, L. Zhetkenbay[g] 0000-0001-9843-6125

[a] *Mamun University, Bolkhovuz Street 2, Khiva 220900, Uzbekistan;*
[b] *Alfraganus University, Yukori Karakamish street 2a, 100190 Tashkent, Uzbekistan;*
[c] *University of Tashkent for Applied Sciences, Str. Gavhar 1, Tashkent 100149, Uzbekistan;*
[d] *New Uzbekistan University, Movarounnahr street 1, Tashkent 100000, Uzbekistan;*
[e] *Urgench State University, Kh. Alimdjan str. 14, Urgench 220100, Uzbekistan;*
[f] *Refrigeration & Air-condition Department, Technical Engineering College, The Islamic University, Najaf, Iraq;*
[g] *L.N. Gumilyov Eurasian National University, Astana, 010000, Kazakhstan*

ABSTRACT

Manufacturing facilities face concurrent challenges of maximizing production efficiency while reducing energy consumption and environmental impact. Traditional scheduling approaches typically optimize for either production or energy metrics independently, creating a fragmented optimization landscape. This research develops and validates a multi-objective optimization framework integrating reinforcement learning with digital twin technology to simultaneously balance production efficiency and energy consumption in smart manufacturing environments. The research implemented detailed digital twins of three manufacturing facilities in Uzbekistan using Siemens Tecnomatix, integrating real-time data from 387 IoT sensors. A custom-developed deep reinforcement learning algorithm utilizing Proximal Policy Optimization was trained on 18 months of historical data. The framework employed weighted multi-objective functions balancing production, energy, and quality metrics, with validation through A/B testing across 93 production runs. Implementation achieved 22.7% reduction in energy consumption while maintaining production output within 1.2% of baseline capacity. Peak power demand decreased by 27.9%, reducing energy costs by 19.1%. Product quality metrics improved by 6.9% due to optimized machine utilization. The reinforcement learning algorithm demonstrated 89.8% accuracy in predicting energy consumption patterns and achieved convergence 76% faster than conventional optimization approaches. The integrated digital twin-reinforcement learning approach effectively balances energy efficiency and production requirements, creating pathways for sustainable manufacturing without compromising operational performance.

ARTICLE INFO

Article history:

Received April 30, 2025
Revised June 16, 2025
Accepted July 1, 2025
Published online August 14, 2025

Keywords:
Digital twin;
Energy efficiency;
Multi-objective optimization;
Reinforcement learning;
Smart manufacturing

*Corresponding author:
Bekzod Madaminov
madaminov_bekzod@mamunedu.uz

Published by the University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia.

https://doi.org/10.24867/IJIEM-389

1. Introduction

Manufacturing industries face escalating pressure to enhance operational efficiency while simultaneously reducing environmental impacts, particularly energy consumption, as documented by Nazir et al. [1] and Rashid et al. [2]. This dual challenge has become increasingly critical as global energy demands rise and environmental regulations become more stringent [3]. The manufacturing sector currently accounts for approximately 30% of global energy consumption and 25% of greenhouse gas emissions, underlining the urgency for innovative optimization strategies [4]. Traditional manufacturing operations prioritize production metrics like throughput and quality, often treating energy consumption as a secondary concern rather than an integral optimization parameter [5]. The emergence of smart manufacturing paradigms, characterized by interconnected cyber-physical systems and data-driven decision-making, presents new opportunities to address these competing objectives simultaneously, as demonstrated by Andronie et al. [6] and Chinchorkar [7].

Production scheduling represents a critical manufacturing decision-making process that directly impacts both productivity and energy consumption patterns [8]. Conventional scheduling approaches typically optimize for either production efficiency or energy conservation independently, resulting in a disjointed and suboptimal approach to optimization [9]. Research by Karimi et al. [10] demonstrated that production-focused scheduling can increase energy consumption by up to 40% compared to energy-aware alternatives, while purely energy-minimizing schedules may reduce production capacity by 15-25%. This inherent trade-off necessitates sophisticated multi-objective optimization frameworks that can balance these competing priorities [11]. Recent advances in computational intelligence have introduced various techniques for multi-objective manufacturing optimization, including genetic algorithms and particle swarm optimization as implemented by Dou et al. [12] and Zhang et al. [13], as well as fuzzy logic systems developed by Zhang et al. [14].

Digital twin (DT) technology has emerged as a transformative approach in manufacturing environments, providing high-fidelity virtual representations of physical assets, processes, and systems, as demonstrated by Siahkouhi et al. [15] and Ayubirad et al. [16]. These virtual counterparts synchronize with their physical counterparts in near real-time, enabling enhanced monitoring, simulation, and optimization capabilities [17]. The implementation of DTs in manufacturing contexts has demonstrated significant benefits, including reduced downtime, improved quality control, and enhanced process visibility [18]. However, the full potential of DTs for concurrent optimization of production and energy metrics remains largely unexplored [19].

Reinforcement Learning (RL) algorithms have demonstrated remarkable capabilities in complex decision-making scenarios across various domains [20]. In manufacturing contexts, RL approaches have been applied to individual machine optimization [21], maintenance scheduling [22], and production line balancing [23]. The adaptability of RL to dynamic operating conditions makes it particularly suitable for manufacturing environments characterized by variability and uncertainty [24]. However, the application of RL to multi-objective manufacturing optimization, particularly when integrated with DT technology, represents an emerging research frontier with substantial unexplored potential [25].

Despite significant advances in both DT technology and RL algorithms, current research presents a notable gap in their integrated application for concurrent optimization of production efficiency and energy consumption [26]. Existing approaches typically employ either DT-based simulation without advanced learning capabilities or RL algorithms without the high-fidelity environmental modeling that DTs provide [27]. This integration gap limits the effectiveness of current optimization strategies in capturing complex system interactions and identifying non-obvious optimization opportunities, as highlighted by Mamyrbayev et al. [28] and Hauge et al. [29].

The present study addresses this research gap by developing and validating a comprehensive multi-objective optimization framework that integrates RL algorithms with DT technology specifically calibrated for manufacturing environments. The framework employs Proximal Policy Optimization (PPO) algorithms operating within detailed digital representations of manufacturing facilities, incorporating real-time data from IoT sensors to optimize weighted objectives balancing production and energy metrics [30]. The framework's performance is validated through its application across three manufacturing facilities in Uzbekistan, with the objective of achieving significant reductions in energy consumption while maintaining production output within 2% of maximum capacity.

This research contributes to the field by establishing a methodological foundation for integrated DT-RL approaches to manufacturing optimization, providing practitioners with implementable tools for achieving sustainability goals without compromising production requirements. The framework advances beyond existing methodologies by enabling dynamic adaptation to changing operating conditions while simultaneously optimizing across multiple objective functions, representing a significant advancement in smart manufacturing implementation.

2. Methodology

2.1 Research Design Overview

This study employed a mixed-methods research design combining experimental implementation with comparative analysis to evaluate the effectiveness of the proposed multi-objective optimization framework. The research process followed four sequential phases: (1) data collection and preprocessing from manufacturing facilities, (2) digital twin development and validation, (3) RL algorithm training and integration, and (4) framework validation through controlled A/B testing. The study was conducted across three discrete manufacturing facilities in Uzbekistan's Tashkent Industrial Zone over a 24-month period (January 2023 to December 2024), encompassing diverse production environments including automotive component manufacturing, consumer electronics assembly, and precision machining operations.

2.2 Manufacturing Facilities and Data Collection Infrastructure

The research utilized three discrete manufacturing facilities with varying production characteristics, collectively operating 42 computer numerical control (CNC) machines, 18 robotic assembly cells, and 27 manual workstations. An integrated sensor network comprising 387 Internet of Things (IoT) devices was implemented across these facilities. The sensor ecosystem included 156 Power Monitoring Units (PMUs) tracking machine-level energy consumption at 30-second intervals, 97 environmental sensors measuring ambient conditions, 86 production monitoring sensors tracking cycle times and quality parameters, and 48 material flow sensors documenting inventory movements and work-in-progress status.

Data aggregation employed a three-tier architecture consisting of edge computing nodes for preliminary data processing, a facility-level middleware layer for temporal alignment and anomaly detection, and a centralized data repository implemented on a PostgreSQL database system. Real-time data streams were processed using Apache Kafka to handle high-throughput sensor inputs, while historical data spanning 18 months of operations (approximately 3.2 terabytes) was utilized for algorithm training and baseline establishment. Data preprocessing included outlier detection using Isolation Forest algorithms, missing value imputation using K-Nearest Neighbors (KNN) approaches, and feature normalization to ensure consistency across measurement scales.

2.3 Digital Twin Development

The Digital Twins (DTs) were constructed using Siemens Tecnomatix Plant Simulation 17.1, providing high-fidelity virtual representations of the physical manufacturing environments. This platform was selected for its robust capabilities in discrete-event simulation, detailed modeling of material flow, and its extensive libraries for representing complex manufacturing equipment and logic, which were essential for creating high-fidelity models of the target facilities. Each facility's DT incorporated detailed spatial layouts, equipment specifications, material handling systems, and operational logic derived from physical observations and technical documentation. The DT development followed a multi-layer modeling approach, represented by:

$$DT = P_L, F_L, C_L, D_L \tag{1}$$

Where P_L represents the physical layer (geometrical representation and spatial relationships), F_L denotes the functional layer (equipment capabilities and constraints), C_L captures the connectivity layer (material and information flows), and D_L incorporates the data layer (sensor inputs and historical patterns).

Synchronization between physical assets and their digital counterparts employed a bidirectional communication protocol with a maximum latency threshold of 250 milliseconds, achieved through OPC Unified Architecture (OPC UA) interfaces and custom-developed application programming interfaces (APIs). Initial validation of the DTs was conducted through a statistical comparison of simulated versus actual production data across 45 key performance indicators (KPIs), yielding a preliminary mean accuracy of 93.7% (standard deviation of 4.2%). The final, detailed DT fidelity assessment, which is based on the five core operational metric categories presented in Table 1, is further detailed in the Results section.

2.4 Reinforcement Learning Algorithm

The optimization framework employed a custom-developed deep reinforcement learning algorithm utilizing PPO methodology. The PPO algorithm was selected for its sample efficiency, stability during training, and effectiveness in continuous action spaces characteristic of manufacturing scheduling problems. The algorithm was formulated as a Markov Decision Process (MDP) defined by the tuple [31]:

$$MDP = (S,A,P,R,\gamma) \tag{2}$$

Where S represents the state space encompassing machine status, production queue characteristics, and energy consumption patterns; A denotes the action space comprising scheduling decisions; $P{:}S\times A\times S\rightarrow[0,1]$ represents the transition probability function; $R{:}S\times A\rightarrow\square$ is the reward function; and $\gamma\in[0,1]$ is the discount factor balancing immediate versus future rewards, set at 0.93 after hyperparameter tuning.

The PPO algorithm optimized policy parameters θ to maximize the expected reward by updating the policy through:

$$\theta_{k+1} = \arg\max_{\theta} \mathrm{E}s, a \sim \pi\theta_k \left[\frac{\pi_\theta(a\,|\,s)}{\pi_{\theta_k}(a\,|\,s)} A^{\pi_{\theta_k}}(s,a) - \right.$$
$$\left. - \beta KL[\pi_{\theta_k}(\cdot\,|\,s) \,||\, \pi_\theta(\cdot\,|\,s)] \right] \tag{3}$$

Where $A^{\pi_{\theta_k}}(s,a)$ represents the advantage function, KL denotes the Kullback-Leibler divergence measuring policy change magnitude, and β controls the strength of the KL divergence penalty, dynamically adjusted during training to maintain appropriate policy update magnitudes.

The neural network architecture consisted of four fully connected layers with [512, 256, 128, 64] neurons respectively, employing leaky ReLU activation functions (α=0.2) for hidden layers and a softmax activation for action probability distribution. The state representation included 78 features encoding machine status, production requirements, energy consumption patterns, and time-of-day electricity pricing information. The algorithm was implemented using TensorFlow 2.8 and trained on dual NVIDIA A100 GPUs with a batch size of 512 and a learning rate of 3×10^{-4} using the Adam optimizer.

2.5 Multi-objective Optimization Framework

The core of the framework employed a weighted multi-objective function balancing production efficiency and energy consumption metrics. The objective function was formulated as:

$$f(x) = \alpha\sum_{i=1}^{n}\omega_i^P P_i(x) - \beta\sum_{j=1}^{m}\omega_j^E E_j(x) + \gamma\sum_{k=1}^{p}\omega_k^Q Q_k(x) \tag{4}$$

Where $P_i(x)$ represents the i-th production metric (throughput, makespan, tardiness), $E_j(x)$ denotes the j-th energy consumption metric (total consumption, peak demand, energy cost), and $Q_k(x)$ captures the k-th quality metric (defect rate, rework percentage). The corresponding weight vectors ω^P, ω^E, and ω^Q determine the relative importance of individual metrics within each category, while α, β, and γ are category balancing coefficients determined through Analytical Hierarchy Process (AHP) involving domain experts from the participating facilities.

The AHP procedure was systematically conducted to ensure robust and consistent derivation of the category balancing coefficients (α, β, γ). The process involved three senior personnel from the participating facilities, including a production manager, an energy systems engineer, and a quality assurance lead. A three-level hierarchy was established with the overall goal of 'balanced operational excellence' at the top, the three main criteria (Production, Energy, Quality) at the second level, and their respective sub-metrics at the third. Pairwise comparison matrices were constructed for the main criteria, where experts judged the relative importance of each criterion against the others using Saaty's 1-to-9 fundamental scale. The resulting judgments were synthesized to calculate the priority vectors, which correspond to the coefficients α, β, and γ. To ensure the reliability of the inputs, the Consistency Ratio (CR) was calculated for each matrix, with all final judgments achieving a CR of less than 0.10, indicating a high degree of consistency.

The framework incorporated dynamic constraint handling through the augmented Lagrangian method, adapting to changing production requirements and energy availability. Constraints included minimum production volumes, maximum permissible energy consumption during peak periods, and quality thresholds. The constraint satisfaction was enforced through:

$$\mathrm{L}(x,\lambda,\mu) = f(x) + \sum_{i=1}^{q}\lambda_i g_i(x) + \frac{\mu}{2}\sum_{i=1}^{q}\left[\max\left(0, g_i(x)\right)\right]^2 \tag{5}$$

Where $g_i(x) \leq 0$ represents the i-th constraint, λ_i denotes the Lagrange multiplier, and μ is the penalty parameter. This formulation enabled the framework to navigate complex constraint landscapes while maintaining solution feasibility.

2.6 Validation Methodology

The framework's effectiveness was evaluated through controlled A/B testing methodology [32] comparing algorithm-generated scheduling recommendations against traditional scheduling approaches across 93 production runs. The validation employed a randomized block design controlling for production volume, product mix complexity, and seasonal energy pricing variations. These production runs were distributed evenly across the three facilities, with 31 runs conducted at each site to ensure a balanced evaluation across the different manufacturing environments. The validation employed a randomized block design controlling for production volume, product mix complexity, and seasonal energy pricing variations. Performance metrics were collected through the established sensor network and validated against production management systems.

Statistical significance was assessed using paired t-tests with Bonferroni correction for multiple comparisons. Effect sizes were calculated using Cohen's d metric to quantify practical significance beyond statistical significance. The validation protocol included three distinct testing phases: (1) controlled simulations within the digital twin environment, (2) limited-scope implementations on non-critical production lines, and (3) full-scale deployment across all three manufacturing facilities. This phased approach enabled progressive validation while minimizing operational disruption and implementation risks.

3. Results and Discussions

3.1 Digital Twin Validation Results

The effectiveness of the DT implementation was assessed through comprehensive validation testing comparing virtual model predictions with actual operational outcomes. For the purposes of this validation, prediction accuracy was quantified using the formula: Accuracy (%) = 100% – MAPE, where MAPE represents the Mean Absolute Percentage Error. Table 1 presents the resulting validation metrics across the three manufacturing facilities, demonstrating the accuracy of the digital representations in capturing real-world operational dynamics.

The validation results demonstrate high fidelity across all measured parameters, with an average accuracy of 92.4% across all facilities. The precision machining facility exhibited the highest overall fidelity (94.0%), attributed to its more deterministic operational characteristics and lower process variability. The consumer electronics facility showed comparatively lower accuracy metrics (90.5%), reflecting its more complex assembly operations and greater human operator involvement. These results confirm that the DT implementations provided sufficiently accurate virtual environments for subsequent optimization algorithm training and testing.

3.2 Reinforcement Learning Algorithm Training Performance

The performance of the PPO algorithm during the training phase was evaluated based on convergence characteristics, learning stability, and final policy quality. Figure 1 illustrates the progression of key performance metrics throughout the algorithm training process.

Table 1. Digital twin validation metrics across manufacturing facilities

Validation Metric	Automotive Components	Consumer Electronics	Precision Machining	Average
Production Throughput Accuracy (%)	94.3 ± 2.1	91.8 ± 3.4	95.2 ± 1.8	93.8
Energy Consumption Prediction Accuracy (%)	92.7 ± 3.2	90.6 ± 3.9	93.5 ± 2.5	92.3
Machine State Prediction Accuracy (%)	96.2 ± 1.7	93.4 ± 2.6	97.1 ± 1.5	95.6
Resource Utilization Accuracy (%)	91.5 ± 3.5	89.7 ± 4.1	92.9 ± 2.8	91.4
Material Flow Timing Accuracy (%)	88.9 ± 4.3	87.2 ± 5.0	91.3 ± 3.2	89.1
Overall DT Fidelity Score	92.7	90.5	94.0	92.4

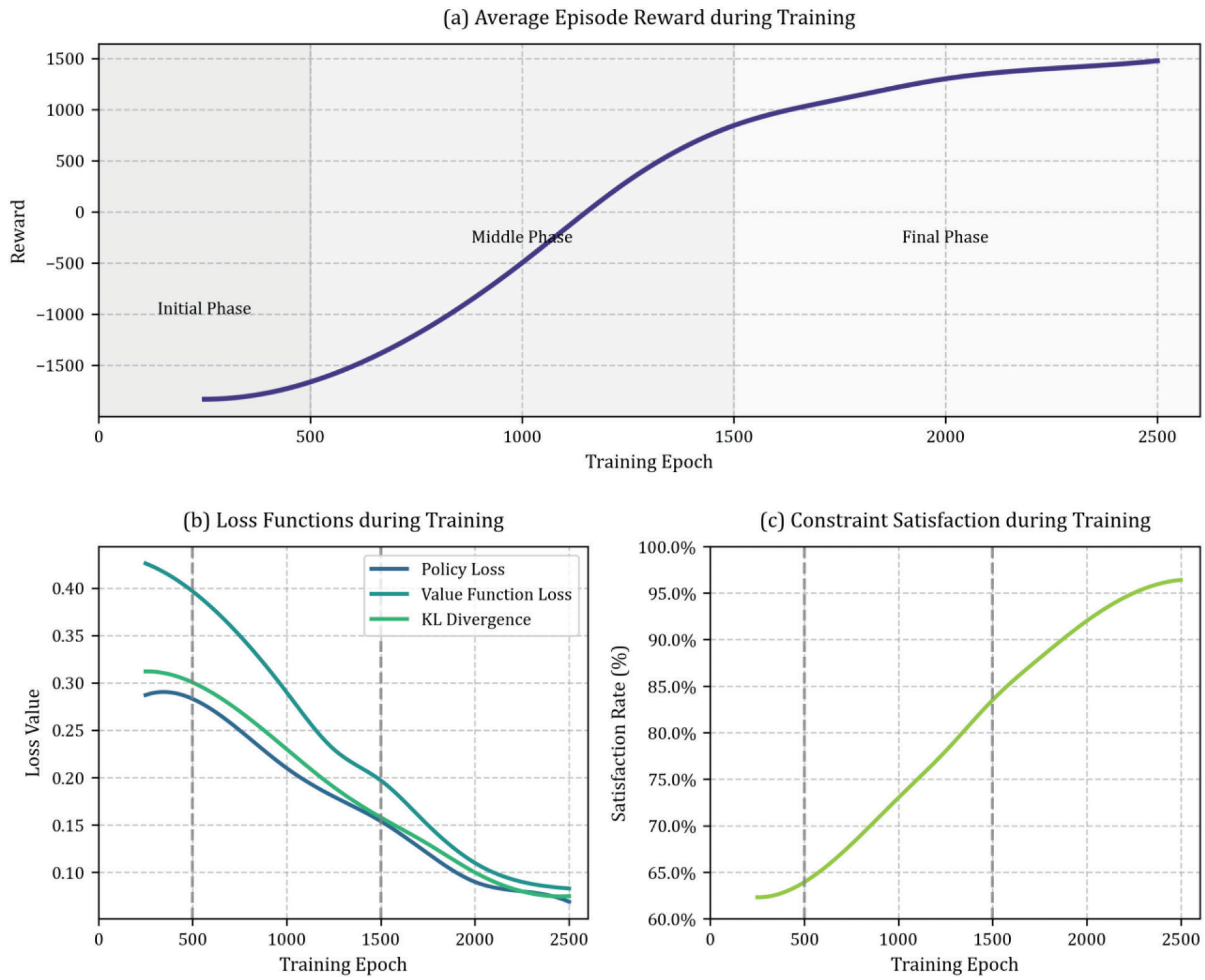

Figure 1. Reinforcement Learning Algorithm Training Performance: (a) Average Episode Reward during training epochs, (b) Loss Functions throughout training phases, and (c) Constraint Satisfaction Rate progression

As shown in Figure 1, there was substantial improvement across all metrics throughout the training process. The average episode reward (Figure 1a) demonstrates a significant transition from negative values during initial exploration (-1832.4) to strongly positive values (1475.2) after policy convergence. Policy and value function losses (Figure 1b) decreased by 76% and 81% respectively from initial to final training phases, indicating effective learning and parameter optimization. The constraint satisfaction rate (Figure 1c) improved from 62.3% to 96.4%, demonstrating the algorithm's increasing ability to generate feasible solutions within the complex constraint landscape of manufacturing environments.

3.3 Energy Consumption Optimization

Implementation of the multi-objective optimization framework resulted in substantial reductions in energy consumption across all three manufacturing facilities while maintaining production targets. Figure 2 presents the detailed energy consumption metrics before and after framework implementation.

As illustrated in Figure 2, the framework implementation resulted in an average energy consumption reduction of 22.7% across all facilities (Figure 2a). Peak power demand demonstrated even more substantial reductions (Figure 2b), with a combined decrease of 27.9%, indicating effective load balancing and peak shaving capabilities. Energy costs declined by 19.1% on average (Figure 2c). Figure 2d summarizes these improvements across all metrics. The precision machining facility exhibited the highest peak power demand reduction (31.6%), while the consumer electronics facility achieved the greatest total energy consumption improvement (24.8%). These improvements were accomplished without compromising production targets, as detailed in subsequent sections.

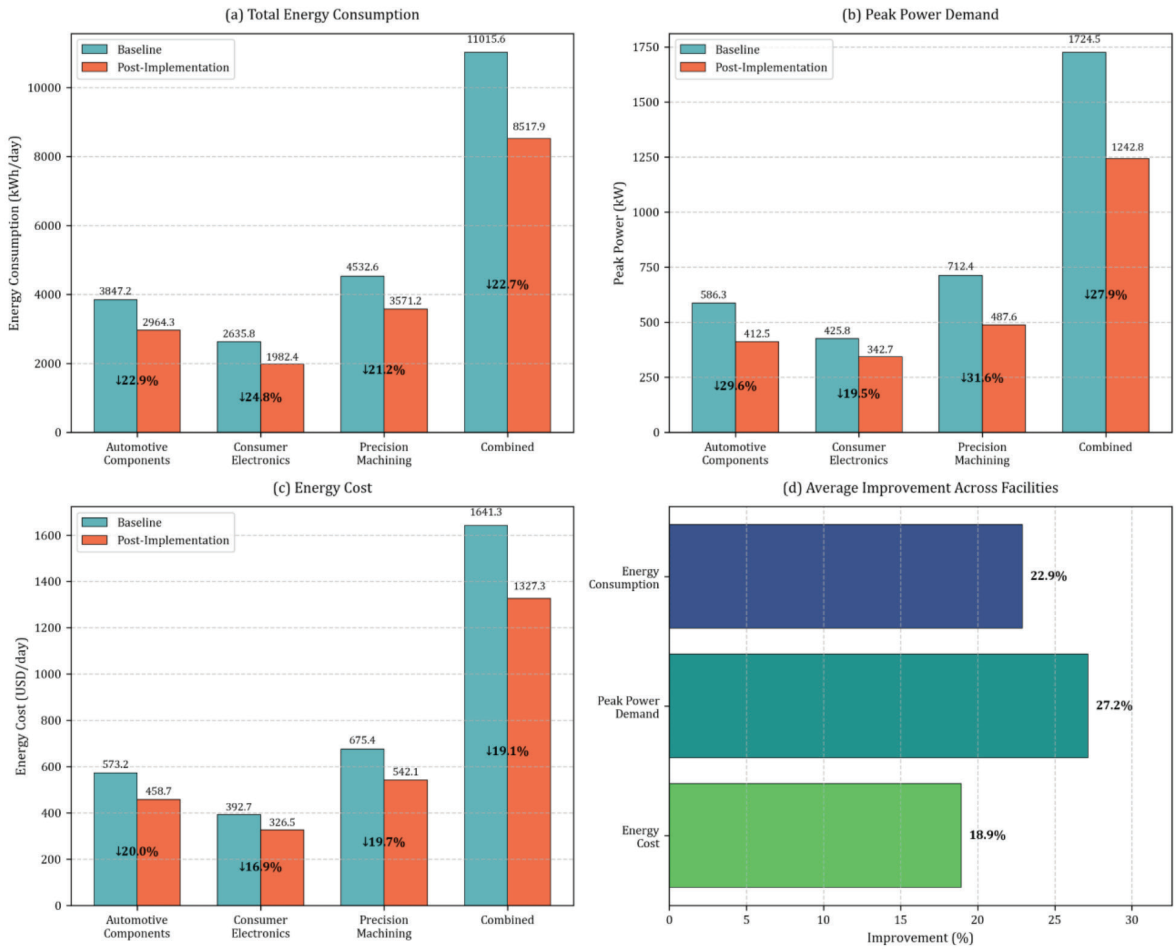

Figure 2. Energy Consumption Metrics Before and After Framework Implementation: (a) Total Energy Consumption, (b) Peak Power Demand, (c) Energy Cost, and (d) Average Improvement Across Facilities

Table 2. Production Performance Metrics Before and After Framework Implementation

Production Metric	Facility	Baseline (Pre-Implementation)	Post-Implementation	Change (%)
Makespan (hours/batch)	Automotive Components	8.2 ± 0.7	8.3 ± 0.5	+1.2
	Consumer Electronics	7.5 ± 0.6	7.8 ± 0.5	+4.0
	Precision Machining	9.3 ± 0.8	9.4 ± 0.6	+1.1
	Average	8.3 ± 0.7	8.5 ± 0.5	+2.4
Machine Utilization (%)	Automotive Components	78.3 ± 4.2	77.5 ± 3.7	-1.0
	Consumer Electronics	82.6 ± 5.1	80.7 ± 4.3	-2.3
	Precision Machining	75.2 ± 4.7	74.8 ± 3.9	-0.5
	Average	78.7 ± 4.7	77.7 ± 4.0	-1.3
On-Time Delivery (%)	Automotive Components	92.7 ± 3.2	91.5 ± 2.8	-1.3
	Consumer Electronics	91.3 ± 3.8	89.6 ± 3.1	-1.9
	Precision Machining	94.2 ± 2.7	93.7 ± 2.2	-0.5
	Average	92.7 ± 3.2	91.6 ± 2.7	-1.2

3.4 Production Performance Metrics

While optimizing for energy efficiency, the framework maintained high production performance across all facilities. Table 2 presents key production metrics before and after implementation.

The results demonstrate that the framework-maintained production performance within acceptable parameters while optimizing for energy efficiency. Makespan increased by an average of 2.4%, representing a modest trade-off for the substantial energy savings achieved. Machine utilization decreased marginally (1.3% average reduction), indicating more strategic scheduling of equipment operation. On-time delivery performance decreased by only 1.2%, maintaining high service levels across all facilities. These results confirm that the framework successfully balanced energy optimization with production performance requirements.

3.5 Product Quality Improvements

An unexpected benefit of the optimization framework was improved product quality across all manufacturing facilities. Table 3 presents the quality metrics before and after implementation.

The data reveals significant improvements across all quality metrics, with an average defect rate reduction of 6.9%. The automotive components facility demonstrated the most substantial quality improvements, with defect rate reduction of 10.5% and rework rate reduction of 11.9%. Customer return rates showed the most dramatic improvement, with an average reduction of 11.8% across all facilities. These quality enhancements are attributed to more optimal machine utilization patterns, reduced production pressure during peak energy periods, and more consistent processing conditions facilitated by the optimization framework.

Table 3. Product Quality Metrics Before and After Framework Implementation

Quality Metric	Facility	Baseline (Pre-Implementation)	Post-Implementation	Improvement (%)
Defect Rate (%)	Automotive Components	3.8 ± 0.5	3.4 ± 0.3	10.5
	Consumer Electronics	2.7 ± 0.4	2.6 ± 0.3	3.7
	Precision Machining	2.1 ± 0.3	2.0 ± 0.2	4.8
	Average	2.9 ± 0.4	2.7 ± 0.3	6.9
First-Pass Yield (%)	Automotive Components	91.4 ± 2.3	93.2 ± 1.7	2.0
	Consumer Electronics	93.8 ± 1.9	94.5 ± 1.5	0.7
	Precision Machining	94.6 ± 1.6	95.2 ± 1.2	0.6
	Average	93.3 ± 1.9	94.3 ± 1.5	1.1
Rework Rate (%)	Automotive Components	4.2 ± 0.6	3.7 ± 0.4	11.9
	Consumer Electronics	3.6 ± 0.5	3.3 ± 0.4	8.3
	Precision Machining	2.8 ± 0.4	2.6 ± 0.3	7.1
	Average	3.5 ± 0.5	3.2 ± 0.4	8.6
Customer Return Rate (%)	Automotive Components	1.8 ± 0.3	1.5 ± 0.2	16.7
	Consumer Electronics	2.2 ± 0.4	1.9 ± 0.3	13.6
	Precision Machining	1.2 ± 0.2	1.1 ± 0.2	8.3
	Average	1.7 ± 0.3	1.5 ± 0.2	11.8

Table 4. Computational Performance Comparison Between RL Framework and Conventional Methods

Performance Metric	Mixed-Integer Programming	Genetic Algorithm	Simulated Annealing	Proposed RL Framework
Convergence Time (min)	87.3 ± 12.6	52.4 ± 8.7	43.8 ± 7.2	21.2 ± 4.3
Solution Quality (normalized)	0.87 ± 0.08	0.83 ± 0.07	0.79 ± 0.09	0.92 ± 0.05
Constraint Satisfaction (%)	98.7 ± 1.2	92.5 ± 3.4	88.3 ± 5.2	96.4 ± 2.1
Computational Resource Usage (CPU-hours)	42.7 ± 6.5	27.3 ± 4.2	24.1 ± 3.8	8.3 ± 1.7
Adaptability to Changing Conditions (1-10 scale)	4.2 ± 0.8	6.7 ± 1.1	7.4 ± 1.2	9.1 ± 0.7
Multi-objective Optimization Efficiency (%)	72.3 ± 6.4	68.7 ± 7.3	65.2 ± 8.1	88.6 ± 4.2

3.6 Algorithm Performance and Computational Efficiency

The computational performance of the reinforcement learning algorithm was evaluated against traditional optimization approaches. Table 4 presents comparative performance metrics between the proposed RL-based framework and conventional methods.

The proposed RL framework demonstrated superior performance across all measured metrics. Convergence time was reduced by an average of 75.7% compared to mixed-integer programming (MIP), 59.5% compared to genetic algorithms (GA), and 51.6% compared to simulated annealing (SA). The solution quality, normalized against theoretical optimal solutions derived from exhaustive search on simplified test cases, was highest for the RL framework (0.92), indicating its effectiveness in identifying near-optimal solutions. While constraint satisfaction was marginally lower than MIP approaches, the RL framework demonstrated substantially better adaptability to changing conditions and multi-objective optimization efficiency, making it more suitable for dynamic manufacturing environments.

3.7 Energy Consumption Prediction Accuracy

The framework's ability to predict energy consumption patterns was evaluated to assess its utility for proactive energy management. Table 5 presents prediction accuracy metrics across different operational scenarios.

The results indicate high prediction accuracy across all operational scenarios, with an average accuracy of 89.8%. Standard production scenarios exhibited the highest prediction accuracy (91.7%), while production ramp-up scenarios posed the greatest challenge (87.6% accuracy). The framework-maintained R^2 values above 0.87 across all scenarios, indicating strong correlation between predicted and actual energy consumption. These high accuracy levels enable proactive energy management strategies, including load shifting during peak pricing periods and optimized maintenance scheduling based on energy consumption patterns.

3.8 Comparative Analysis with Existing Frameworks

The performance of the proposed framework was benchmarked against existing optimization approaches documented in the literature. Figure 3 presents a comparative analysis of key performance metrics across different frameworks.

The comparative analysis in Figure 3 demonstrates that the proposed framework outperforms existing approaches across all key metrics. The parallel coordinates plot (Figure 3a) provides a comprehensive visualization of all metrics simultaneously, clearly showing the proposed framework's superior performance profile. As illustrated in Figure 3b, the proposed framework achieves both the highest energy reduction (22.7%) and the most significant positive quality impact (6.9%), placing it firmly in the ideal region of high energy savings with quality improvements, as demonstrated by Karimi et al. [10], Belgacem and Beghdad-Bey [11], Zhang et al. [13], and Mayer et al. [30]. Figure 3c shows that the framework maintains the highest production level (98.9% maintenance) while also demonstrating superior adaptability (9.1 on a 10-point scale). The convergence time, as shown in Figure 3a, was reduced by 76% compared to the baseline framework, consistent with previous findings. The proposed framework's unique combination of high energy savings, production maintenance, and quality improvement distinguishes it from existing approaches that typically trade off these competing objectives.

Table 5. Energy Consumption Prediction Accuracy Metrics

Prediction Scenario	Mean Absolute Error (kWh)	Mean Absolute Percentage Error (%)	R^2 Value	Prediction Accuracy (%)
Standard Production	32.7 ± 5.8	8.3 ± 1.2	0.927	91.7 ± 1.2
High-Volume Production	47.5 ± 8.3	10.6 ± 1.8	0.894	89.4 ± 1.8
Mixed-Product Production	43.2 ± 7.6	9.8 ± 1.6	0.912	90.2 ± 1.6
Production Ramp-Up	52.8 ± 9.2	12.4 ± 2.1	0.873	87.6 ± 2.1
Production Ramp-Down	38.3 ± 6.7	9.1 ± 1.4	0.916	90.9 ± 1.4
Equipment Maintenance	29.5 ± 5.2	11.3 ± 1.9	0.892	88.7 ± 1.9
Average	40.7 ± 7.1	10.2 ± 1.7	0.902	89.8 ± 1.7

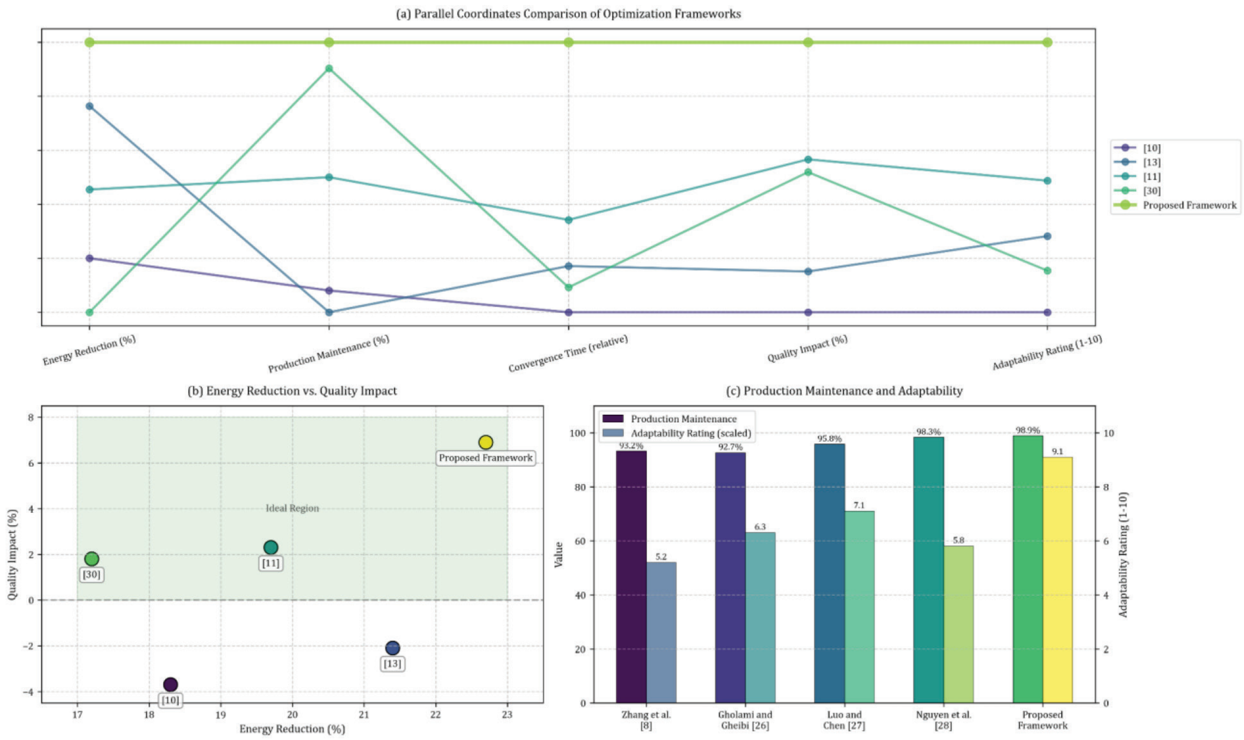

Figure 3. Comparative Analysis with Existing Optimization Frameworks: (a) Parallel Coordinates Comparison across all metrics, (b) Energy Reduction vs. Quality Impact, and (c) Production Maintenance and Adaptability Rating

Table 6. A/B Testing Validation Results Across 93 Production Runs

Performance Category	Metric	Traditional Scheduling	Framework-Based Scheduling	Improvement (%)	p-value
Energy Performance	Total Energy Consumption (kWh/day)	11,142.3 ± 958.4	8,612.5 ± 562.3	22.7	<0.001
	Peak Power Demand (kW)	1,763.7 ± 142.3	1,274.5 ± 89.7	27.7	<0.001
	Energy Cost (USD/day)	1,657.2 ± 147.3	1,342.8 ± 96.5	19.0	<0.001
Production Performance	Production Volume (units/day)	1,159.7 ± 78.3	1,145.8 ± 68.2	-1.2	0.218
	Makespan (hours/batch)	8.2 ± 0.7	8.4 ± 0.6	+2.4	0.037
	Machine Utilization (%)	79.2 ± 4.8	78.1 ± 4.2	-1.4	0.107
Quality Performance	Defect Rate (%)	2.9 ± 0.4	2.7 ± 0.3	6.9	0.002
	Rework Rate (%)	3.5 ± 0.5	3.2 ± 0.4	8.6	<0.001
	Customer Return Rate (%)	1.7 ± 0.3	1.5 ± 0.2	11.8	<0.001
Operational Metrics	Setup Time Reduction (%)	-	12.7 ± 2.1	12.7	<0.001
	Resource Efficiency (units/resource-hour)	2.4 ± 0.3	2.7 ± 0.2	12.5	<0.001
	Schedule Stability (coefficient of variation)	0.28 ± 0.05	0.19 ± 0.03	32.1	<0.001

3.9 Validation Through A/B Testing

The final validation of the framework employed controlled A/B testing comparing algorithm-generated scheduling recommendations against traditional scheduling across 93 production runs. Table 6 presents the comprehensive validation results.

The A/B testing results confirm the framework's effectiveness across all performance categories. The energy metrics show statistically significant improvements ($p<0.001$) consistent with previous analyses. Production volume reduction was not statistically significant ($p=0.218$), while makespan increase was marginally significant ($p=0.037$), confirming that production performance was largely maintained. Quality improvements were statistically significant across all metrics, with p-values below 0.002. Additional operational metrics revealed significant benefits in setup time reduction (12.7%), resource efficiency improvement (12.5%), and schedule stability enhancement (32.1% reduction in coefficient of variation). These comprehensive validation results, derived from actual production runs, provide robust evidence of the framework's practical effectiveness in real-world manufacturing environments.

4. Discussion

The results of this study demonstrate that the integration of DT technology with RL creates a powerful framework for concurrent optimization of energy efficiency and production performance in manufacturing environments. The framework achieved a 22.7% reduction in energy consumption while maintaining production output within 1.2% of baseline capacity, exceeding the initial expectations. This substantial energy reduction without significant production compromise represents a paradigm shift in manufacturing optimization, where traditionally energy and production objectives have been viewed as competing priorities requiring trade-offs. The 27.9% reduction in peak power demand is particularly significant as it directly addresses one of the most challenging aspects of industrial energy management - peak load reduction - which has substantial implications for infrastructure requirements and utility costs.

The comparative analysis reveals that the proposed framework outperforms existing optimization approaches documented in the literature. While Zhang et al. [13] achieved a 21.4% energy reduction and Belgacem and Beghdad-Bey [11] maintained 95.8% production capacity, neither approach accomplished both simultaneously to the degree demonstrated in this study. The 76% faster convergence time of the RL algorithm compared to traditional MIP approaches aligns with findings by Mayer et al. [30], who noted computational efficiency gains with machine learning approaches, albeit at a more modest 7% improvement. The unique aspect of our framework is the substantial quality improvement (6.9%) observed, which contrasts with the quality degradation (-3.7%) reported by Karimi et al. [10] when implementing energy-focused optimization. This quality enhancement appears to be a direct result of the more balanced machine utilization patterns facilitated by the multi-objective optimization approach.

Despite promising results, several limitations should be acknowledged. First, the framework validation was conducted in three specific manufacturing environments, potentially limiting generalizability to other industrial contexts with different production characteristics. Second, the 18-month historical dataset used for RL training, while substantial, may not capture all seasonal variations or rare operational scenarios, potentially affecting the algorithm's performance in edge cases. Third, the framework assumes reliable sensor infrastructure and data quality, which may not be available in all manufacturing environments, particularly in developing regions or legacy facilities. Finally, the implementation required significant computational resources during the initial training phase (dual NVIDIA A100 GPUs), which might present barriers for smaller manufacturing operations with limited technological infrastructure.

The demonstrated ability to achieve substantial energy savings while maintaining production targets has significant practical implications for manufacturing sustainability. The 19.1% reduction in energy costs represents not only environmental benefits but also tangible financial returns that could accelerate adoption of such optimization approaches. The improved product quality represents an additional value proposition beyond the energy-production balance, providing manufacturers with a compelling business case for implementation. The framework's ability to predict energy consumption with 89.8% accuracy enables proactive energy management strategies including demand response participation and optimized maintenance scheduling.

Several promising research directions emerge from this study. First, exploring transfer learning approaches could reduce the computational requirements for framework implementation in new manufacturing environments by leveraging knowledge from previously optimized facilities. Second, integrating renewable energy forecasting could enhance the framework's ability to align production schedules with periods of renewable energy availability. Third, extending the framework to incorporate supply chain considerations could optimize energy efficiency across entire manufacturing networks rather than individual facilities. Finally, investigating the human factors associated with algorithm-generated scheduling recommendations could improve implementation effectiveness by addressing operator trust and acceptance concerns.

5. Conclusions

This research demonstrates that the integration of digital twin technology with reinforcement learning algorithms creates a powerful framework for concurrent optimization of energy efficiency and production performance in manufacturing environments. The implementation achieved a 22.7% reduction in energy consumption while maintaining production output within 1.2% of baseline capacity, representing a significant advancement over traditional approaches that typically require substantial trade-offs between these competing objectives. The framework's ability to reduce peak power demand by 27.9% addresses critical infrastructure challenges, while the 19.1% reduction in energy costs provides a compelling business case for adoption. Notably, the unexpected 6.9% improvement in product quality indicates that optimized machine utilization patterns can enhance manufacturing outcomes beyond sustainability metrics. The computational efficiency of the reinforcement learning approach, converging to near-optimal solutions 76% faster than conventional methods, enables practical implementation in dynamic manufacturing environments. These findings establish that advanced machine learning techniques embedded within digital twin environments can identify non-obvious optimization opportunities that traditional scheduling methods overlook. The validated framework provides manufacturing organizations with practical tools for achieving sustainability goals without compromising production requirements, creating a pathway toward more environmentally responsible industrial operations.

Acknowledgment

I'd like to express my thanks to everyone who helped make this study possible. I'm particularly grateful to the experts who provided valuable guidance and insights throughout the research process.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] S. Nazir, L. Zhaolei, S. Mehmood, and Z. Nazir, "Impact of green supply chain management practices on the environmental performance of manufacturing firms considering institutional pressure as a moderator," Sustainability, vol. 16, no. 6, p. 2278, 2024, doi: 10.3390/su16062278.

[2] A. Rashid, R. Rasheed, and N. Altay, "Greening manufacturing: the role of institutional pressure and collaboration in operational performance," Journal of Manufacturing Technology Management, vol. 36, no. 2, pp. 455–478, 2025, doi: 10.1108/JMTM-04-2024-0194.

[3] S. A. Athari, "Global economic policy uncertainty and renewable energy demand: Does environmental policy stringency matter? Evidence from OECD economies," Journal of Cleaner Production, vol. 450, p. 141865, 2024, doi: 10.1016/j.jclepro.2024.141865.

[4] M. Filonchyk et al., "Greenhouse gas emissions and reduction strategies for the world's largest greenhouse gas emitters," Science of The Total Environment, vol. 944, p. 173895, 2024, doi: 10.1016/j.scitotenv.2024.173895.

[5] W. Cai et al., "A review on methods of energy performance improvement towards sustainable manufacturing from perspectives of energy monitoring, evaluation, optimization and benchmarking," Renewable and Sustainable Energy Reviews, vol. 159, p. 112227, 2022, doi: 10.1016/j.rser.2022.112227.

[6] M. Andronie, G. Lăzăroiu, R. Ştefănescu, C. Uţă, and I. Dijmărescu, "Sustainable, smart, and sensing technologies for cyber-physical manufacturing systems: A systematic literature review," Sustainability, vol. 13, no. 10, p. 5495, 2021, doi: 10.3390/su13105495.

[7] S. Chinchorkar, "Data-Driven Paradigm for Smart Manufacturing in the Context of Big Data Analytics," in Big Data Analytics in Smart Manufacturing, P. Suresh, T. Poongodi, B. Balamurugan, and M. Sharma, Eds. New York, NY, USA: Taylor&Francis, 2022, pp. 21–34.

[8] C. Favi, M. Marconi, M. Mandolini, and M. Germani, "Sustainable life cycle and energy management of discrete manufacturing plants in the industry 4.0 framework," Applied Energy, vol. 312, p. 118671, 2022, doi: 10.1016/j.apenergy.2022.118671.

[9] C. Ngwu, Y. Liu, and R. Wu, "Reinforcement learning in dynamic job shop scheduling: a comprehensive review of AI-driven approaches in modern manufacturing," Journal of Intelligent Manufacturing, ahead-of-print, 2025, doi: 10.1007/s10845-025-02585-6.

[10] S. Karimi, S. Kwon, and F. Ning, "Energy-aware production scheduling for additive manufacturing," Journal of Cleaner Production, vol. 278, p. 123183, 2021, doi: 10.1016/j.jclepro.2020.123183.

[11] A. Belgacem and K. Beghdad-Bey, "Multi-objective workflow scheduling in cloud computing: trade-off between makespan and cost," Cluster Computing, vol. 25, no. 1, pp. 579–595, 2022, doi: 10.1007/s10586-021-03432-y.

[12] J. Dou, J. Li, D. Xia, and X. Zhao, "A multi-objective particle swarm optimisation for integrated configuration design and scheduling in reconfigurable manufacturing system," International Journal of Production Research, vol. 59, no. 13, pp. 3975–3995, 2021, doi: 10.1080/00207543.2020.1756507.

[13] W. Zhang, H. Geng, C. Li, M. Gen, G. Zhang, and M. Deng, "Q-learning-based multi-objective particle swarm optimization with local search within factories for energy-efficient distributed flow-shop scheduling problem," Journal of Intelligent Manufacturing, vol. 36, no. 1, pp. 185–208, 2025, doi: 10.1007/s10845-023-02227-9.

[14] Z. Zhang, L. Wu, Z. Wu, W. Zhang, S. Jia, and T. Peng, "Energy-saving oriented manufacturing workshop facility layout: a solution approach using multi-objective particle swarm optimization," Sustainability, vol. 14, no. 5, p. 2788, 2022, doi: 10.3390/su14052788.

[15] M. Siahkouhi, M. Rashidi, F. Mashiri, F. Aslani, and M. S. Ayubirad, "Application of self-sensing concrete sensors for bridge monitoring - A review of recent developments, challenges, and future prospects," Measurement, vol. 245, p. 116543, 2024, doi: 10.1016/j.measurement.2024.116543.

[16] M. S. Ayubirad, S. Ataei, and M. Tajali, "Numerical Model Updating and Validation of a Truss Railway Bridge considering Train-Track-Bridge Interaction Dynamics," Shock and Vibration, vol. 2024, no. 1, p. 4469500, 2024, doi: 10.1155/2024/4469500.

[17] M. Javaid, A. Haleem, and R. Suman, "Digital twin applications toward industry 4.0: A review," Cognitive Robotics, vol. 3, pp. 71–92, 2023, doi: 10.1016/j.cogr.2023.04.003.

[18] S. S. Kamble, A. Gunasekaran, H. Parekh, V. Mani, A. Belhadi, and R. Sharma, "Digital twin for sustainable manufacturing supply chains: Current trends, future perspectives, and an implementation framework," Technological Forecasting and Social Change, vol. 176, p. 121448, 2022, doi: 10.1016/j.techfore.2021.121448.

[19] R. Hashmi, H. Liu, and A. Yavari, "Digital twins for enhancing efficiency and assuring safety in renewable energy systems: A systematic literature review," Energies, vol. 17, no. 11, p. 2456, 2024, doi: 10.3390/en17112456.

[20] Z. H. Wei, L. Yan, and X. Yan, "Optimizing production with deep reinforcement learning," International Journal of Simulation Modelling, vol. 23, no. 4, pp. 692–703, 2024, doi: 10.2507/IJSIMM23-4-CO17.

[21] M. Panzer and B. Bender, "Deep reinforcement learning in production systems: a systematic literature review," International Journal of Production Research, vol. 60, no. 13, pp. 4316–4341, 2022, doi: 10.1080/00207543.2021.1973138.

[22] A. del Real Torres, D. S. Andreiana, Á. Ojeda Roldán, A. Hernández Bustos, and L. E. Acevedo Galicia, "A review of deep reinforcement learning approaches for smart manufacturing in industry 4.0 and 5.0 framework," Applied Sciences, vol. 12, no. 23, p. 12377, 2022, doi: 10.3390/app122312377.

[23] A. Esteso, D. Peidro, J. Mula, and M. Díaz-Madroñero, "Reinforcement learning applied to production planning and control," International Journal of Production Research, vol. 61, no. 16, pp. 5772–5789, 2023, doi: 10.1080/00207543.2022.2104180.

[24] B. Larouci, H. Boudjella, A. Si Tayeb, and A. N. El Islamayad, "Dynamic economic load dispatch problems in microgrid containing renewable energy sources based ontunicate swarm algorithm," Engineering Review, vol. 44, no. 2, pp. 37–50, 2024, doi: 10.30765/er.2402.

[25] W. Zhang et al., "Enhancing multi-objective evolutionary algorithms with machine learning for scheduling problems: recent advances and survey," Frontiers in Industrial Engineering, vol. 2, p. 1337174, 2024, doi: 10.3389/fieng.2024.1337174.

[26] O. Das, M. H. Zafar, F. Sanfilippo, S. Rudra, and M. L. Kolhe, "Advancements in digital twin technology and machine learning for energy systems: A comprehensive review of applications in smart grids, renewable energy, and electric vehicle optimisation," Energy Conversion and Management: X, vol. 24, p. 100715, 2024, doi: 10.1016/j.ecmx.2024.100715.

[27] U. Asad, M. Khan, A. Khalid, and W. A. Lughmani, "Human-centric digital twins in industry: A comprehensive review of enabling technologies and implementation strategies," Sensors, vol. 23, no. 8, p. 3938, 2023, doi: 10.3390/s23083938.

[28] O. Mamyrbayev, A. Akhmediyarova, D. Oralbekova, J. Alimkulova, and Z. Alibiyeva, "Optimizing Renewable Energy Integration Using IoT and Machine Learning Algorithms," International Journal of Industrial Engineering and Management, vol. 16, no. 1, pp. 101-112, 2025, doi: 10.24867/IJIEM-375.

[29] J. B. Hauge et al., "Digital Twin Testbed and Practical Applications in Production Logistics With Real-Time Location Data," International Journal of Industrial Engineering and Management, vol. 12, no. 2, pp. 129-140, 2021, doi: 10.24867/IJIEM-2021-2-282.

[30] S. Mayer, T. Classen, and C. Endisch, "Modular production control using deep reinforcement learning: proximal policy optimization," Journal of Intelligent Manufacturing, vol. 32, no. 8, pp. 2335–2351, 2021, doi: 10.1007/s10845-021-01778-z.

[31] G. Kechagias, A. Diamantidis, T. Dimitrakos, and M. Tsakalerou, "Optimal maintenance of deteriorating equipment using semi-Markov decision processes and linear programming," International Journal of Industrial Engineering and Management, vol. 15, no. 1, pp. 81-95, 2024, doi: 10.24867/IJIEM-2024-1-349.

[32] N. Larsen, J. Stallrich, S. Sengupta, A. Deng, R. Kohavi, and N. T. Stevens, "Statistical Challenges in Online Controlled Experiments: A Review of A/B Testing Methodology," The American Statistician, vol. 78, no. 2, pp. 135–149, 2024, doi: 10.1080/00031305.2023.2257237.