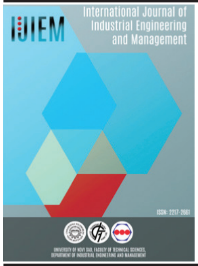
ISSN 2683-345X
journal homepage: http://ijiemjournal.uns.ac.rs/



International Journal of Industrial Engineering and Management

2026 / ARTICLE-IN-PRESS / 424

Original research article

Towards Sustainable and Resilient Industry 5.0 Production Systems with a Digital Twin-Enabled Circular Economy Approach

R. P. Rajendran[a,*] 0000-0002-8257-0984

[a] *Assistant Professor - Selection Grade, School of Management, Presidency University, Itgalpura, Yelahanka, Bengaluru, Karnataka, 560064, India*

ABSTRACT

Manufacturing systems must simultaneously reduce resource intensity and maintain operational continuity under disruption, yet most real-time control architectures do not incorporate circular-economy and resilience objectives within production decision making. This study develops and evaluates a Digital Twin (DT) framework that functions as a decision engine for Industry 5.0 production systems rather than as a monitoring-only representation. In this framework, the DT remains synchronized with the physical line through live IoT data, a credibility gate that suspends prescriptive actions when synchronization error exceeds acceptable bounds, a provenance ledger that records material recovery events, and a multi-objective scheduler that balances energy consumption, waste generation, and delivery delay under human approval. Circularity is operationalized as the share of material recovered through reuse or recycling, whereas resilience is quantified as normalized recovery performance following a disruption. Using a design science approach, the framework was implemented for six months on an operating electronics manufacturing line in Riyadh, Saudi Arabia, covering 3,048 production lots. The evaluation combined a two-month baseline, a four-month intervention period, and disruption scenarios for supplier delay, machine breakdown, and quality shift events. The intervention reduced energy intensity by 14.4%, decreased weekly downtime by 23.9%, and increased overall equipment effectiveness by 5.0% while maintaining product quality. The circularity rate increased from 0.23 to 0.33, with larger gains for packaging and casings than for printed circuit boards because of component-specific reuse constraints. Resilience scores increased from 0.71 to 0.82, indicating faster recovery from supply and equipment disruptions. These findings show that a traceable, credibility-gated, human-supervised DT can align shop-floor decisions with circularity and resilience objectives, although the reported gains should be interpreted in light of the single-line pilot context.

ARTICLE INFO

Article history:

Received January 26, 2026
Revised April 20, 2026
Accepted May 22, 2026
Published online September 11, 2026

Keywords:
Circular economy;
Digital twin;
Industry 5.0;
Multi-objective scheduling;
Resilience

*Corresponding author:
Ragu Prasadh Rajendran
ragu.prasadh@presidencyuniversity.in

1. Introduction

Manufacturing firms are under simultaneous pressure to improve economic performance, reduce environmental burden, and maintain continuity during increasingly frequent disruptions. Industry 5.0 frames this challenge by positioning sustainability, resilience, and human-centricity alongside productivity as core design objectives for production systems [1], [2]. Within this agenda, circular economy implementation has moved from a strategic aspiration to an operational requirement. Extending material use, reducing waste, and improving recovery rates

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

https://doi.org/10.24867/IJIEM-424

are well-established objectives, but the critical challenge lies in translating these objectives into routine shop-floor decisions without overstating the capabilities of existing digital tools [3]. Industry 5.0 extends the connectivity and automation agenda of Industry 4.0 by explicitly linking digital technologies to social value, workforce support, and ecological constraints [4]. For manufacturing research, this implies that control systems should not optimize throughput alone; they should also preserve recoverability, transparency, and operator agency under disturbed operating conditions. This human-centered emphasis remains insufficiently operationalized. Prior studies on the transition toward Operator 5.0 report that many Industry 4.0 implementations improve data visibility but do not adequately support workers in high-stakes operational decisions [5]. As a result, the principles of resilience and sustainability are often stated at the policy level but weakly embedded in day-to-day production control.

Digital technologies can support the transition from circular-economy strategy to execution, yet the integration logic is still underdeveloped. IoT and analytics frameworks help organize circular supply-chain information [6], but they do not fully specify how real-time production data, material recovery objectives, and operator decisions should be combined at the line level. Empirical evidence further shows that adoption remains uneven. IoT is the most mature capability in sustainability-oriented manufacturing, whereas artificial intelligence, advanced analytics, and blockchain are more frequently used in pilots or isolated applications [7], [8]. This fragmentation limits the ability of firms to convert circular-economy objectives into coordinated operational action. Research on digital transformation and net-zero manufacturing reaches a similar conclusion: firms are increasingly motivated to combine digitalization with circularity, but the procedural mechanisms for embedding those goals in production control remain contingent, firm-specific, and insufficiently formalized [9]. The unresolved issue is therefore not whether digital technologies are relevant, but how they should be architected to influence real-time decisions. Digital Twin (DT) research provides a promising foundation because DTs can couple physical assets and virtual models across multiple manufacturing scales [10]-[12]. However, the literature also shows recurring limitations related to model credibility, interoperability, and transfer from laboratory demonstrations to operational settings. Work on DT-enabled scheduling suggests that twins can support dynamic replanning and disruption response, but the concept is often treated as an enhanced monitoring or simulation layer rather than as an accountable decision system [13]. Consequently, the architectural boundary between a passive DT and an active decision-making DT remains insufficiently defined. Delphi and review studies reinforce this concern by showing that manufacturing DT research remains heterogeneous in definitions, maturity levels, and implementation pathways [14]. This ambiguity is especially problematic for Industry 5.0, where the legitimacy of automated recommendations depends on credibility, transparency, and effective collaboration with human operators.

A second gap concerns performance measurement. Circular economy research employs a wide variety of indicators across micro, meso, and macro levels, which complicates the translation of strategic sustainability goals into operational decision variables [15]. In parallel, DT research has concentrated on equipment monitoring, maintenance, and scheduling [16], [17], while circular economy studies have focused more strongly on governance, strategy, and enabling technologies [18], [19]. The intersection of these literatures remains underdeveloped. Existing studies rarely offer a framework that simultaneously assimilates live shop-floor data, evaluates circularity and resilience in real time, and recommends operational actions that remain interpretable to human decision makers. As a result, circularity and resilience are often monitored ex post rather than optimized during production execution.

Accordingly, the scientific problem addressed in this study is not the absence of digital technologies per se, but the absence of a coherent decision architecture that connects DT synchronization, circular-economy metrics, resilience objectives, and human oversight within the same operational loop. In the present study, the term decision-engine DT denotes a DT architecture with four properties beyond conventional monitoring-oriented twins: continuous ingestion of live production data, credibility-gated state synchronization, prescriptive evaluation of feasible actions through a multi-objective scheduler, and operator authorization before execution. This combination distinguishes the proposed framework from DT-enabled scheduling or cyber-physical control studies that optimize production states without simultaneously embedding provenance, circularity, resilience, and human accountability in the decision pathway [13], [14], [28], [31]. This approach is motivated by the maturity of enabling technologies and by their present fragmentation. IoT sensing and AI-based prediction provide real-time visibility into machine condition and process performance, while blockchain-based traceability can document provenance and recovery pathways [20], [21]. However, these capabilities typically operate as separate modules. The electronics factory in Riyadh offers an appropriate empirical setting because it combines mixed-product routing, measurable recovery streams, and recurrent operational disturbances relevant to circularity and resilience. The study makes four bounded contributions. First, it conceptualizes the DT as a decision engine by specifying the architectural components that convert synchronized representation into prescriptive support. Second, it integrates sensing, predictive analytics, provenance tracking, and scheduling within a single operational framework. Third, it operationalizes circularity and resilience as explicit decision criteria rather than downstream reporting variables. Fourth, it evaluates how human review and override mechanisms shape the use of DT recommendations in practice. The study addresses three research questions: How can a DT system handle circularity and resilience objectives together in real time? What architecture can link IoT, AI, and provenance tracking while remaining interpretable to human operators? Does this system improve operational performance in a real factory setting? The objectives of the study are to: (i) design a DT framework that links production control with sustainability and resilience objectives; (ii) implement the framework in a real factory using IoT, AI, and provenance tracking; (iii) evaluate the framework using energy, throughput, circularity, and recovery measures; and (iv) examine how human interaction affects the quality and usability of DT-supported decisions.

2. Methodology

2.1 Study design and setting

The study followed a design science method. It started by building a framework and ended with a test in a real electronics factory. This factory is in Riyadh, Saudi Arabia. The steps in the study matched the specific goals. The DT architecture and circular economy logic met Objective 1. Setting up the physical system and its integration met Objective 2. Testing performance and disruptions met Objective 3. Checking the human side of the system met Objective 4. The study lasted six months. This time frame captured changes in demand and maintenance. It also allowed for a long-term look at the data [22]-[25].

2.2 System architecture and standards alignment

The feedback loop operated in five stages. First, instrumentation data were ingested and quality-checked at the gateway. Second, the DT state was updated through synchronization routines that compared observed and simulated states. Third, predictive modules estimated machine condition, quality risk, energy demand, and material recovery opportunities. Fourth, the scheduler evaluated candidate actions against normalized energy, waste, and delivery objectives subject to feasibility constraints. Fifth, the system issued ranked recommendations to operators together with explanatory metrics and provenance information, and only approved actions were pushed back to the shop floor. This sequence is the basis for treating the DT as a decision engine rather than as a passive visualization model.

The system architecture followed the ISO 23247 DT framework and associated NIST implementation guidance [26], [27]. The resulting architecture comprised four interacting layers: a physical layer containing machines, sensors, and material handling assets; a data layer responsible for ingestion, cleaning, and storage; a digital model layer combining discrete-event and continuous-state representations; and a decision layer in which predictive models, circularity analytics, and schedule optimization generated ranked recommendations. This layered structure supported modular validation, traceability, and controlled scaling [28], [29]. Model credibility was treated as an operational prerequisite rather than a post hoc quality attribute [30], [31].

2.3 Physical system and instrumentation

The factory processed 3,048 production lots during the study. This amount of data helped train and test the models. It gave steady estimates for energy, waste, and resilience across many shifts. The data density also allowed for reliable estimates of machine health and quality without using fake data. This fits the standards for maintenance reviews [25].

The test line included assembly, testing, and inspection machines. It handled different products at the same time. Tools on the machines measured power, vibration, and temperature. Scanners and cameras tracked parts and checked quality. The system recorded power data every second. Sensors collected vibration data quickly but grouped it into one-second chunks. The system logged production events as they happened. Monthly checks kept the power meters accurate. A shaker verified the vibration sensors. A standard board validated the camera settings. These checks kept measurement drift below 2 percent. This level of accuracy is acceptable for factory work.

2.4 Data acquisition, preprocessing, and data governance

An IoT gateway collected sensor streams and production events and transferred them to the database and DT services after preprocessing. The preprocessing workflow removed implausible outliers, flagged missing intervals longer than five minutes, and retained audit logs for subsequent verification. A permissioned provenance ledger recorded production, rework, recycling, and material-return events at the unit and component levels, thereby providing a traceable record of circular-economy transactions rather than a generic blockchain registry. Access to commercially sensitive records was role-based so that engineers, supervisors, and auditors viewed only the information required for their decisions [32], [33].

The DT combined two model classes to represent material and energy dynamics. A discrete-event model captured routing, buffers, and cycle times, while continuous-state components represented energy demand and machine-condition evolution. The model was initialized from factory routing data and calibration experiments, and its synchronized use in decision support was conditioned on a credibility gate. Specifically, prescriptive functions were enabled only when daily synchronization error remained within predefined acceptance bounds, thereby preventing optimization from acting on a drifting or weakly calibrated twin.

2.5 Digital twin modeling and synchronization

The DT combined a discrete-event representation for material flow with continuous state models for energy and machine condition. The synchronization routine updated queue lengths, machine states, and energy variables from live data, then compared the updated twin against observed states through the RMSE criterion in Equation (1). When the synchronization error exceeded the prescribed threshold, the system triggered recalibration and restricted the DT to monitoring mode until acceptable fidelity was restored. In this sense, the credibility gate converted model verification into an active control on prescriptive use.

$$\mathrm{RMSE}_x = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(x_t^{\mathrm{phys}} - x_t^{\mathrm{twin}}\right)^2} \tag{1}$$

where RMSE_x is the root mean square error for state variable x, x_t^{phys} is the observed value from sensors at time t, x_t^{twin} is the corresponding DT estimate, and T is the number of synchronized samples. This metric was computed daily for energy, throughput, and queue length states, and it triggered recalibration when thresholds were exceeded. It served as the gatekeeper for model credibility and protected downstream analysis from drift.

2.6 Circular economy and material flow analytics

Material flows were tracked at the component and assembly levels to quantify reuse, recycling, and scrap. The selected indicators were aligned with established circular economy measurement typologies that emphasize material-flow accounting and recovery rates at the operational level [15], [34]. This alignment provided a benchmarking rationale for the circularity measure while retaining compatibility with line-level data. Equation (2) specifies the mass balance used to reconcile inputs, outputs, recycling, and inventory changes in each production period.

$$M_{m,t}^{\mathrm{in}} + M_{m,t}^{\mathrm{rec}} = M_{m,t}^{\mathrm{prod}} + M_{m,t}^{\mathrm{scrap}} + \Delta M_{m,t}^{\mathrm{inv}} \tag{2}$$

where $M_{m,t}^{\mathrm{in}}$ is the incoming mass of material m in period t, $M_{m,t}^{\mathrm{rec}}$ is recycled input mass, $M_{m,t}^{\mathrm{prod}}$ is mass embedded in finished units, $M_{m,t}^{\mathrm{scrap}}$ is scrap mass, and $\Delta M_{m,t}^{\mathrm{inv}}$ is the inventory change. The balance was computed per shift and reconciled against warehouse records to ensure mass conservation. It provided a transparent basis for circularity metrics and supported audit readiness.

For interpretive consistency, the circularity rate was bounded between 0 and 1, with higher values indicating a larger share of recovered material relative to virgin input. Weekly circularity values were checked against warehouse reconciliation records and compared with component-level recovery patterns to confirm that improvements reflected material recirculation rather than inventory timing effects. This procedure anchored the indicator in both accounting consistency and operational plausibility.

Equation (3) defines the circularity rate used in the optimization and evaluation modules.

$$C_t = \frac{M_t^{\mathrm{rec}} + M_t^{\mathrm{reused}}}{M_t^{\mathrm{in}} + M_t^{\mathrm{rec}}} \tag{3}$$

where C_t is the circularity rate in period t, M_t^{rec} is recycled input mass, M_t^{reused} is mass reused without reprocessing, and M_t^{in} is virgin input mass. The rate was computed weekly and fed into the multi-objective optimizer to shift resource allocation toward reuse and recycling. It operationalized the circular economy objective in a quantitative form.

2.7 Energy and emissions modeling

Energy management followed the ISO 50001 framework for systematic energy performance tracking and continuous improvement [35]. Equation (4) defines the machine-level energy calculation based on operational states.

$$E_{i,t} = \sum_{s\in\{\text{idle},\text{run},\text{setup}\}} P_{i,s}\,\tau_{i,s,t} \tag{4}$$

where $E_{i,t}$ is energy consumed by machine i in period t, $P_{i,s}$ is the power draw for state s, and $\tau_{i,s,t}$ is the duration of that state. State durations were computed from machine logs, and power levels were measured during calibration runs. This equation supported the energy objective and enabled state-specific interventions such as reducing idle energy.

Equation (5) defines energy intensity per finished unit, which allowed normalized comparisons across product mixes.

$$\text{EI}_t = \frac{\sum_{i=1}^{I} E_{i,t}}{N_t^{\text{good}}} \tag{5}$$

where EI_t is energy intensity in period t, $E_{i,t}$ is energy for machine i, I is the number of machines, and N_t^{good} is the number of good units produced. The metric was tracked daily and used to evaluate trade-offs between throughput and energy performance. It linked energy use to product output and supported decision transparency.

Equation (6) defines the operational emissions estimate used for scenario analysis.

$$\text{CO2e}_t = \left(\sum_{i=1}^{I} E_{i,t}\right)\text{EF}_t \tag{6}$$

where CO2e_t is the carbon dioxide equivalent emissions for period t, $E_{i,t}$ is energy consumption, and EF_t is the grid emission factor for the relevant time window. The emission factor was taken from regional energy reporting for Saudi Arabia and updated monthly. This estimate was used for internal comparison of schedules rather than external reporting, and it anchored the sustainability narrative to measurable energy impacts.

Equation (7) specifies overall equipment effectiveness, which provided a productivity control to ensure that sustainability gains did not degrade operational performance. The OEE construct is widely used in manufacturing systems and is grounded in availability, performance, and quality components [36].

$$\begin{aligned} A_t &= \frac{T_t^{\text{oper}}}{T_t^{\text{sched}}} \\ P_t &= \frac{N_t^{\text{tot}}\,t^{\text{ideal}}}{T_t^{\text{oper}}} \\ Q_t &= \frac{N_t^{\text{good}}}{N_t^{\text{tot}}} \\ \text{OEE}_t &= A_t P_t Q_t \end{aligned} \tag{7}$$

where A_t is availability, P_t is performance, Q_t is quality, T_t^{oper} is operating time, T_t^{sched} is scheduled time, N_t^{tot} is total units produced, N_t^{good} is good units, and t^{ideal} is the ideal cycle time. The measure was tracked daily and used to validate that scheduling decisions maintained productivity. It served as a safeguard against unintended losses while shifting toward circular objectives.

2.8 Predictive maintenance and quality risk modeling

Predictive maintenance and quality risk estimation used data-driven models consistent with current reviews that emphasize sensor-rich data pipelines, model retraining, and decision integration for maintenance actions [37]. The modeling pipeline combined feature extraction from vibration spectra, temperature deviations, and cycle-time anomalies with supervised learning to estimate machine health and defect risk.

Equation (8) defines the reliability function and remaining useful life, which enabled maintenance scheduling.

$$\begin{aligned} R_i(t) &= \exp\left(-\int_0^t \lambda_i(\tau)\,\mathrm{d}\tau\right) \\ \text{RUL}_i(t) &= \int_t^{\infty} R_i(\tau)\,\mathrm{d}\tau \end{aligned} \tag{8}$$

where $R_i(t)$ is the reliability of machine i at time t, $\lambda_i(\tau)$ is the hazard rate over time, and $\text{RUL}_i(t)$ is the remaining useful life. The hazard rate was estimated from historical failure intervals and updated with recent condition indicators. This formulation translated sensor trends into actionable maintenance windows, which reduced unplanned downtime and supported the resilience objective.

Equation (9) defines the anomaly score used for online monitoring of process drift.

$$S_t = \sqrt{(z_t - \mu)^{\top} \Sigma^{-1} (z_t - \mu)} \quad (9)$$

where S_t is the anomaly score at time t, z_t is the feature vector, μ is the mean vector of normal operation, and Σ is the covariance matrix. The score was computed in real time and compared to a threshold calibrated on baseline data. It acted as a trigger for early intervention in quality risk and machine health, which protected both circularity outcomes and productivity.

2.9 Optimization and scheduling

The scheduling engine used a multi-objective genetic algorithm in which each candidate solution encoded machine assignments and operation start times subject to routing and capacity constraints [38], [39]. For each candidate schedule, the DT simulated the resulting energy demand, waste generation, queue evolution, and delivery performance, thereby enabling evaluation without interrupting physical production. The scheduler therefore served as the prescriptive core of the decision engine rather than as a stand-alone optimization routine.

Equation (10) defines the normalized multi-objective function used in the optimizer.

The objective function in Equation (10) resolved conflicts among sustainability and service objectives through normalized scalarization. Energy, waste, and delivery delay were each divided by historical reference maxima so that the associated weights reflected managerial priorities rather than differences in measurement units. Candidate schedules that violated feasibility constraints in Equation (11) were discarded before ranking. Among feasible schedules, the selected solution minimized the weighted objective while remaining interpretable to operators through the reported contribution of each objective term. The weight sensitivity analysis reported later in Table 7 was used to test the robustness of recommendations under alternative priority settings.

$$\min_{\mathbf{u}} J = w_1 \frac{E}{E_{\max}} + w_2 \frac{W}{W_{\max}} + w_3 \frac{D}{D_{\max}} \quad (10)$$

where $\mathbf{u}$ is the decision vector for machine assignments and start times, E is total energy, W is total waste mass, D is total delivery delay, and w_1, w_2, and w_3 are weights that sum to one. The normalization constants E_{max}, W_{max}, and D_{max} were set from historical maxima. This function enabled explicit trade-offs among sustainability, circularity, and responsiveness, which directly operationalized the study objectives.

Equation (11) specifies the feasibility constraints used to enforce capacity and precedence relations.

$$\begin{aligned} \sum_{j=1}^{J} p_{i,j}\, y_{i,j} &\le T_i^{\text{avail}} \\ s_{j,k+1} &\ge s_{j,k} + p_{j,k} \end{aligned} \quad (11)$$

where $p_{i,j}$ is the processing time of job j on machine i, $y_{i,j}$ is a binary assignment variable, T_i^{avail} is the available time of machine i in the scheduling horizon, and $s_{j,k}$ is the start time for operation k of job j. The first constraint enforced capacity limits and the second enforced precedence. These constraints preserved the feasibility of schedules generated by the optimizer and ensured that the DT simulation reflected real production rules.

The resilience metric in Equation (12) was implemented as a normalized recovery integral bounded between 0 and 1, where a value closer to 1 denotes limited performance loss and rapid restoration to the pre-disruption reference level. The indicator combined recovery depth and recovery duration, allowing two disruptions with similar end states but different recovery trajectories to be distinguished. To check face validity, resilience scores were interpreted together with recovery time and percentage performance loss in Table 8; scenarios with shorter recoveries and smaller cumulative losses consistently produced higher resilience values.

2.10 Resilience assessment and disruption scenarios

Resilience assessment used disruption scenarios that reflected realistic manufacturing risks, including supplier delays, machine breakdowns, and sudden quality shifts. Each scenario was first simulated in the DT, then validated through controlled changes in the production plan when feasible. This approach aligns with DT practice in manufacturing, which emphasizes safe experimentation through virtual tests before physical execution [23], [28].

Equation (12) defines the resilience metric used to quantify recovery performance.

Because the primary performance evaluation used a pre-post design, several steps were taken to reduce confounding. Baseline and intervention observations were compared across the same production line under stable staffing arrangements and without major equipment replacement or plant-wide process redesign during the pilot window. Product-mix and demand fluctuations were monitored at the weekly level, and the disruption scenarios provided an additional mechanism for testing the DT under comparable stress conditions. Nevertheless, the design was not randomized, so residual effects from seasonal demand variation, operator learning, or other incremental operational improvements cannot be fully excluded and should be considered when interpreting causal claims.

$$\mathrm{R} = 1 - \frac{\int_{t_0}^{t_r} \left(P_{\mathrm{ref}} - P(t)\right)\mathrm{d}t}{P_{\mathrm{ref}}\left(t_r - t_0\right)} \tag{12}$$

where R is the resilience score, $P(t)$ is the observed performance at time t, P_{ref} is the pre-disruption reference performance, t_0 is the disruption start time, and t_r is the recovery time. The integral captured the cumulative performance loss during recovery. This measure allowed direct comparison across different disruption types and supported the resilience objective by quantifying recovery speed and depth.

2.11 Human-centric evaluation

The human-centric evaluation examined how operators, maintenance technicians, and supervisors interacted with the DT interface, interpreted recommendations, and exercised final decision authority. Sixteen participants were selected through purposive sampling to represent the principal operational roles. Each participant completed two guided decision tasks in which the interface displayed the current line state, the recommended action, the objective contributions underlying that recommendation, and the provenance or maintenance context relevant to the decision. Participants could accept, postpone, or override the recommendation and were asked to provide a short justification when an override occurred. Structured interviews were then coded for alert clarity, perceived trust, control, and decision support value. The resulting feedback was used to recalibrate alert thresholds and refine visualization logic so that the DT supported, rather than displaced, operator judgment.

2.12 Statistical analysis and reproducibility

The statistical analysis applied a pre-post design. This design included two months of baseline data and four months of intervention data. The analysis assessed normality with the Shapiro-Wilk test. It conducted parametric comparisons with paired t tests. It handled nonparametric comparisons with the Wilcoxon signed-rank test. The analysis reported effect sizes as Cohen's d for parametric tests and rank-biserial correlation for nonparametric tests. It applied a two-sided significance level of 0.05 with Holm-Bonferroni adjustment for multiple comparisons. The analysis reported ninety-five percent confidence intervals for changes in energy intensity, waste rate, overall equipment effectiveness, and resilience scores. All statistical scripts used a fixed random seed to support reproducibility. The analysis archived intermediate datasets with time stamps and version labels in the data repository.

The methodological choices stressed coherence across objectives, instrumentation, models, and evaluation. Standards compliance anchored the DT design. Circular economy metrics grounded material flow analysis. Predictive maintenance methods supported resilience. Optimization linked sustainability to operational control [40]-[42].

3. Results and Discussions

3.1 Data completeness and governance performance

Data completeness and calibration stability were quantified across sensor classes, and the analysis is reported in Table 1.

Table 1. Data completeness and calibration stability by sensor class across the pilot line.

Sensor class	Expected sampling interval	Data completeness (%)	Missing streaks >5 min per week	Calibration drift (%)	Data excluded (%)
Power meters	1 s	99.1	0.6	1.3	0.8
Vibration	0.5 ms (aggregated to 1 s)	98.4	1.1	1.8	1.5
Temperature	1 s	99.3	0.4	1.1	0.6
Barcode tracking	Event driven	99.6	0.2	NA	0.3
Vision inspection	Event driven	98.9	0.9	1.5	1.2

The completeness rates stayed above 98 percent for all sensor classes. This level supported stable model inputs for the DT. Vibration signals showed the highest missing streaks and the largest exclusion rate. This pattern matched the higher sensitivity of those sensors to mechanical noise and cable disturbances. Calibration drift stayed below 2 percent across instruments. The low exclusion rate cut the risk of bias in energy and quality analytics. These data quality outcomes fit the need for reliable state estimation. They also backed downstream analyses that relied on synchronized data streams.

The study evaluated ledger performance and provenance coverage from transaction logs and reconciliation against the manufacturing execution system. The analysis appears in Figure 1 across three panels. Figure 1(a) shows the distribution of transaction latency for production, rework, and recycling events. Figure 1(b) reports the cumulative number of provenance events per unit across lifecycle phases. Figure 1(c) shows the weekly mismatch rate between ledger events and physical records.

Figure 1(a) indicates a median ledger latency of 2.4 s with a 95th percentile below 6.8 s. This timing kept provenance data close to real time for operational use. Figure 1(b) shows that each finished unit carried a consistent sequence of provenance events. It had a median of 14 production events, 3 rework events, and 5 end-of-life events per unit. Figure 1(c) indicates a stable mismatch rate below 0.5 percent over the pilot. Most discrepancies linked to manual rework steps rather than automated logging. These outcomes show that the ledger offered high coverage and timely traceability. It didn't disrupt shop-floor flow.

3.2 Digital twin synchronization and model fidelity

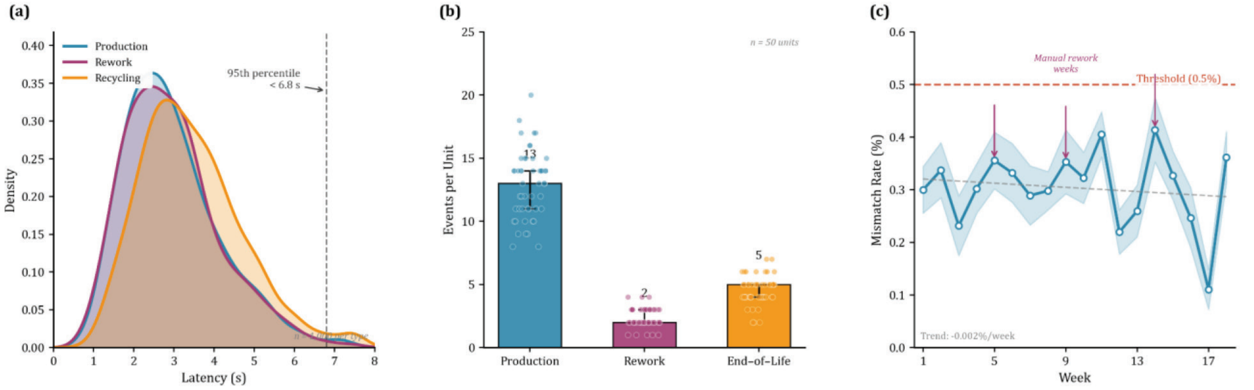
The study evaluated DT synchronization accuracy with RMSE between physical and virtual states. Figure 2 presents the analysis across three panels. Figure 2(a) shows daily RMSE for energy, throughput, and queue length states. Figure 2(b) plots the alignment between physical and twin values for energy and throughput across all synchronized samples. Figure 2(c) presents the distribution of RMSE values by state variable.

Figure 2(a) shows that RMSE dropped during the first four weeks and then stayed steady. It had median values of 1.8 percent for energy, 1.4 percent for throughput, and 2.0 percent for queue length. Figure 2(b) shows a close alignment around the identity line with no regular bias at high loads. This cut the risk of drift in the optimization layer. Figure 2(c) indicates a narrow RMSE distribution for throughput. Queue length errors showed a broader tail during high-mix shifts. The pattern indicates that the twin followed the physical system well enough to aid real-time decision support. It also points out periods where recalibration boosted queue modeling.

3.3 Framework integration and deployment fidelity

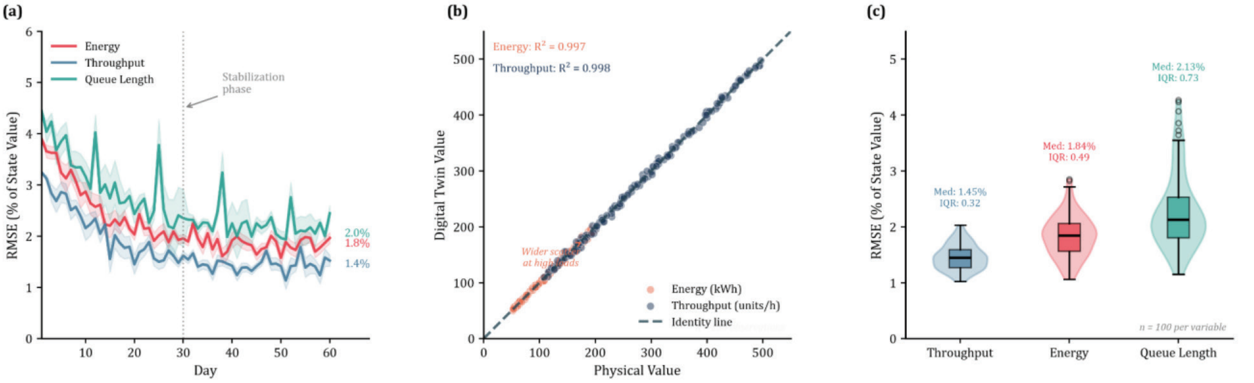
Integration coverage and runtime availability were quantified across core system components, and the analysis is reported in Table 2.

Coverage topped 92 percent for all physical assets and 96 percent for real-time data streams. This confirms the framework worked for most of the line. It didn’t require a total replacement of legacy equip-



Note. Panel (a) presents transaction latency distributions for production, rework, and recycling events. Panel (b) shows cumulative provenance events per unit by lifecycle phase. Panel (c) reports weekly mismatch rates between ledger and physical records.

Figure 1. Blockchain performance and provenance coverage.



Note. Panel (a) reports daily RMSE for energy, throughput, and queue length states. Panel (b) plots physical versus twin values for energy and throughput. Panel (c) shows RMSE distributions by state variable.

Figure 2. DT synchronization accuracy.

Table 2. Integration coverage and runtime availability of the deployed framework.

Integration element	Coverage (%)	Availability (%)
Connected machines	92.3	96.8
Instrumented stations	93.9	97.2
Real-time data streams	96.7	98.1
DT state updates	NA	95.6
Predictive model inference cycles	NA	96.2
Optimization cycles	NA	95.4
Provenance ledger coverage	99.5	99.1
Overall system uptime	NA	97.1

ment. Availability stayed above 95 percent for all digital services. The lowest availability occurred during optimization cycles. Ledger coverage hit 99.5 percent. This matches the low mismatch rate reported earlier. It also supports end-to-end traceability across the product lifecycle. These results show the framework functioned at scale. It remained stable in a real production environment.

3.4 End-to-end control loop timing

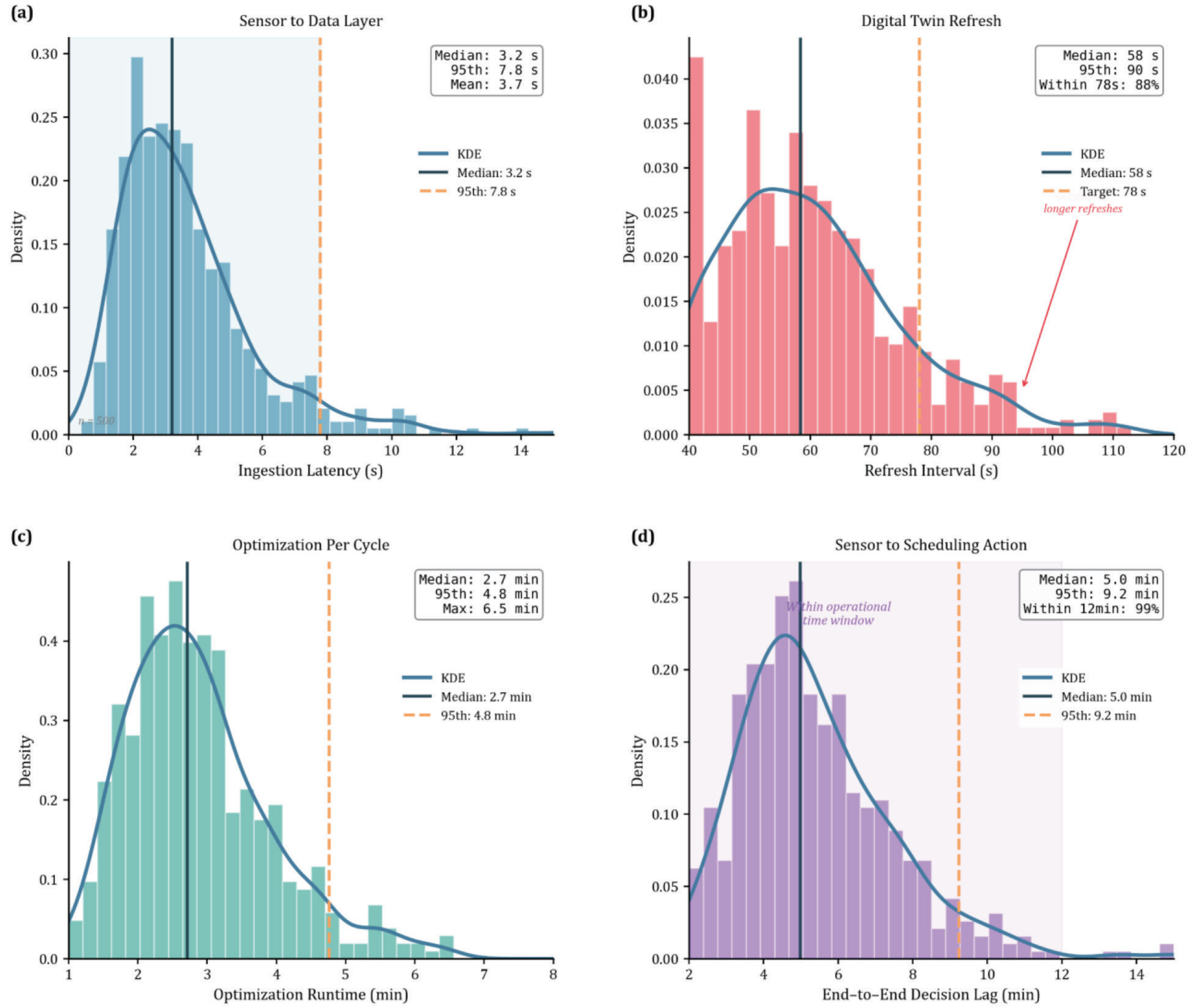
The study measured end-to-end control loop timing. This analysis quantified how fast the system sensed, synchronized, optimized, and made decisions. Figure 3 shows this data across four panels. Figure 3(a) shows the timing for sensor-to-data-layer ingestion latency. Figure 3(b) shows the intervals for the DT state refresh. Figure 3(c) shows the optimization runtimes for each cycle. Figure 3(d) shows the total decision lag from a sensor event to a scheduling action.

Figure 3(a) shows a median ingestion latency of 3.2 s. The 95th percentile reached 9.4 s. This kept the data stream near real-time. Figure 3(b) shows a median twin refresh interval of 60 s. About 95 percent of updates finished within 78 s. This kept the physical line and the twin in sync during regular work. Figure 3(c) shows a median optimization runtime of 2.8 min. The 95th percentile was 6.1 min. This kept scheduling quick enough for a work shift. Figure 3(d) shows a median decision lag of 5.4 min. The 95th percentile reached 11.2 min. This means the system made changes within useful time windows. These timing results show the design worked as a continuous decision loop. It wasn't just a static monitoring tool. This supports the framework design objective and the real-world implementation objective.

3.5 Circular economy performance and material flows

The study computed component-level material flow metrics to measure circularity outcomes. Table 3 reports this analysis. It lists mean inputs, reuse, recycling, scrap, and circularity rates for major component groups per 1,000 finished units.

The circularity rate differed by component group. Packaging achieved the highest rate. This happened



Note. Panel (a) shows ingestion latency distributions from sensors to the data layer. Panel (b) presents DT refresh intervals. Panel (c) reports optimization runtimes per cycle. Panel (d) shows end-to-end decision cycle lag.

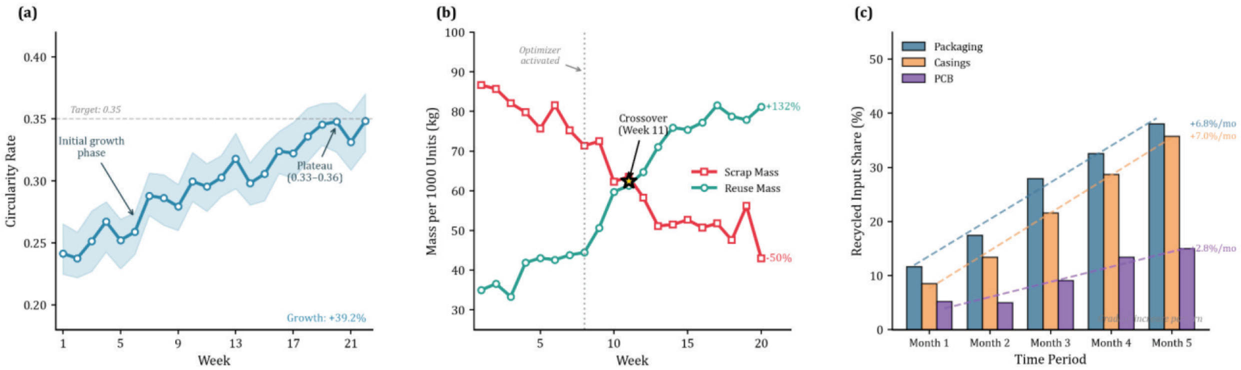
Figure 3. Control loop latency across the deployed framework.

Table 3. Component-level material flow metrics per 1,000 finished units.

Component group	Input mass (kg) mean ± SD	Reused mass (kg) mean ± SD	Recycled mass (kg) mean ± SD	Scrap mass (kg) mean ± SD	Circularity rate mean ± SD
Printed circuit boards	120 ± 8	9 ± 3	18 ± 4	14 ± 3	0.20 ± 0.03
Connectors	45 ± 5	7 ± 2	8 ± 2	5 ± 1	0.28 ± 0.04
Casings	80 ± 7	12 ± 4	16 ± 5	9 ± 2	0.29 ± 0.05
Packaging	30 ± 4	6 ± 2	12 ± 3	3 ± 1	0.43 ± 0.06

because its material streams supported reuse and recycling with fewer quality constraints. Printed circuit boards showed the lowest rate. For these, the scrap mass was higher than reused mass. This reflected tighter quality limits and low remanufacturing yield. Casings and connectors fell between these extremes. They showed a balance between recycling and reuse. These component-level differences indicate that circularity gains weren't uniform. Instead, they depended on material characteristics and the feasibility of rework.

The study examined temporal dynamics in circularity and scrap over the six-month pilot. Figure 4 presents this analysis across three panels. Figure 4(a) shows the weekly circularity rate for the full line. Figure 4(b) presents weekly scrap and reuse mass per 1,000 units. Figure 4(c) shows monthly recycled input share by component group.



Note. Panel (a) reports weekly circularity rates. Panel (b) shows weekly scrap and reuse mass per 1,000 units. Panel (c) presents monthly recycled input share by component group.

Figure 4. Temporal trends in circularity and material recovery.

Figure 4(a) shows a steady increase in circularity from 0.22 in the first month to 0.35 by week 18. It then plateaued near 0.33 to 0.36. Figure 4(b) indicates that scrap mass decreased as reuse mass increased. The largest shift occurred after week 8. At that point, the optimization engine began to prioritize reuse in high-scrap assemblies. Figure 4(c) shows that recycled input shares grew fastest for packaging and casings. Printed circuit boards showed only modest growth. The timing pattern indicates that process tuning and material-specific constraints drove improvements. They didn't come from a uniform policy effect.

3.6 Energy and productivity outcomes

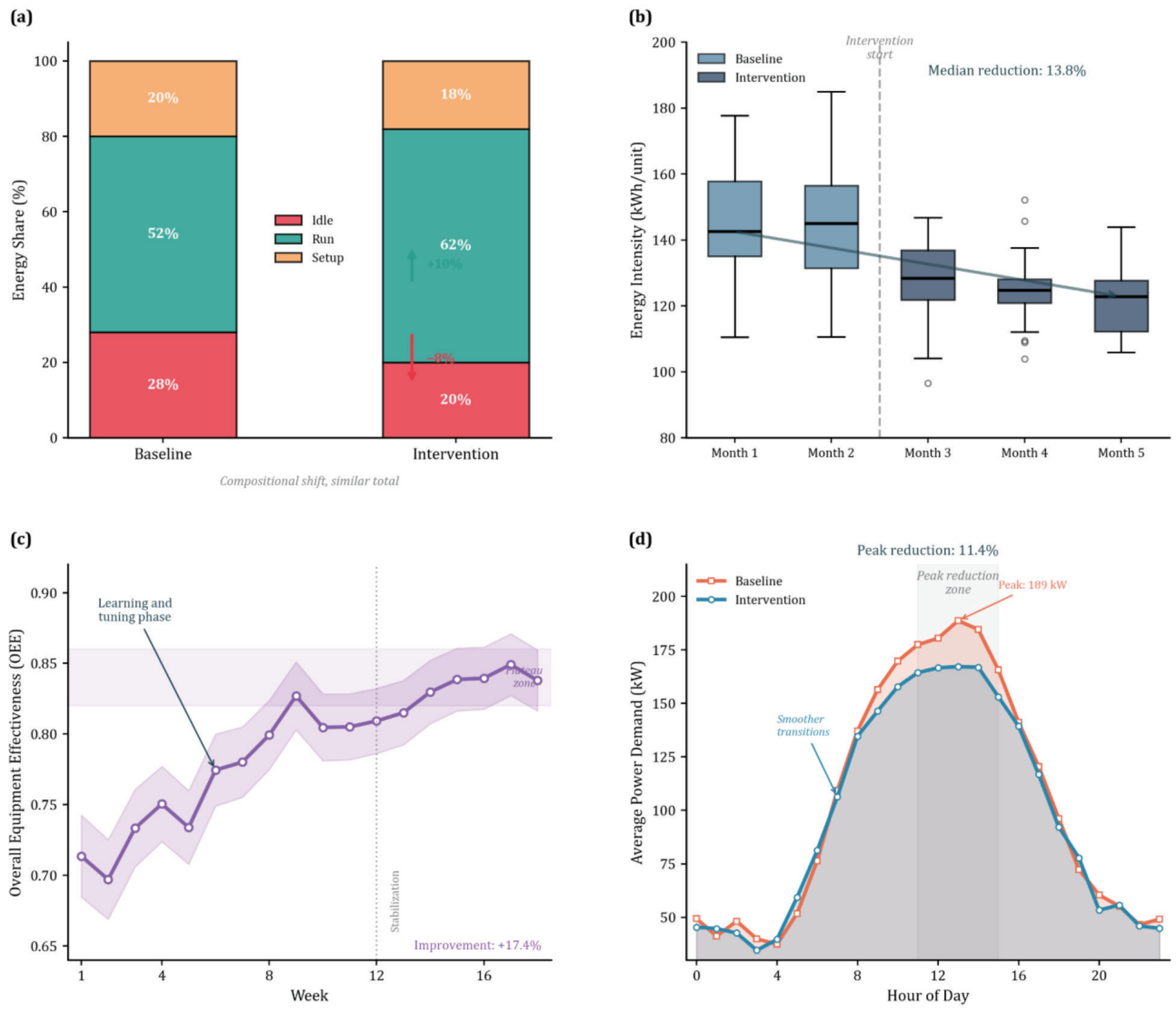
Table 4 compares baseline and intervention outcomes for energy and productivity. Energy intensity decreased by 14.4%, and total energy per shift decreased by 12.9%, indicating improved resource efficiency during the intervention period. Overall equipment effectiveness increased by 5.0%, primarily through gains in availability and performance, whereas quality remained statistically unchanged. Weekly downtime declined by 23.9%, consistent with the intended contribution of predictive maintenance and schedule adjustment. The reported effect sizes indicate that the changes in energy intensity, total energy, OEE, and downtime were practically meaningful rather than marginal.

This work analyzes energy use patterns and productivity trends over time. Figure 5 presents this analysis in four panels. Figure 5(a) shows how much energy the machines used when they were idle, running, or being set up before and after the test. Figure 5(b) shows the energy used each month. Figure 5(c) shows the weekly OEE trend during the pilot test. Figure 5(d) shows the average power used throughout the day.

Figure 5(a) shows that the change cut the idle energy share from 28 percent to 20 percent. The share of energy used while running increased. This suggests that production work and energy use match better now. Figure 5(b) shows that energy use dropped after the second month. There weren't as many big spikes in energy use. This shows better control over peak energy. Figure 5(c) shows OEE gains that stayed

Table 4. Baseline versus intervention outcomes for energy and productivity.

Metric	Baseline mean ± SD	Intervention mean ± SD	Change (%)	Test and p value	Effect size (95% CI)
Energy intensity (kWh per unit)	5.82 ± 0.56	4.98 ± 0.49	-14.4	Paired t, p < 0.001	d = 1.62 (0.68 to 0.99)
Total energy per shift (kWh)	3420 ± 310	2980 ± 280	-12.9	Paired t, p = 0.001	d = 1.43 (320 to 560)
OEE	0.782 ± 0.038	0.821 ± 0.031	+5.0	Paired t, p = 0.002	d = 1.10 (0.018 to 0.060)
Availability	0.904 ± 0.030	0.925 ± 0.024	+2.3	Paired t, p = 0.011	d = 0.76 (0.005 to 0.037)
Performance	0.862 ± 0.028	0.879 ± 0.024	+2.0	Paired t, p = 0.018	d = 0.68 (0.004 to 0.030)
Quality	0.989 ± 0.004	0.988 ± 0.004	-0.1	Paired t, p = 0.312	d = 0.12 (-0.001 to 0.000)
Downtime per week (h)	14.2 ± 3.6	10.8 ± 2.9	-23.9	Paired t, p = 0.004	d = 0.94 (1.1 to 5.4)



Note. Panel (a) shows energy share by machine state in baseline and intervention periods. Panel (b) presents monthly distributions of daily energy intensity. Panel (c) shows weekly OEE trends across the pilot. Panel (d) reports the average hourly power demand profile.

Figure 5. Energy and productivity patterns across time and operational states.

steady after week 12. This matches the time it took to retrain models and fix the schedule. Figure 5(d) shows a flatter power use profile. It had fewer peaks in the early afternoon. This matches the work to shift loads and stagger machine idle times.

3.7 Predictive maintenance and quality risk outcomes

This study tested how well models predicted machine failure and defect risk. The researchers used data that the models hadn't seen before. Table 5 lists the scores for discrimination, precision, recall, and lead time for the two model families.

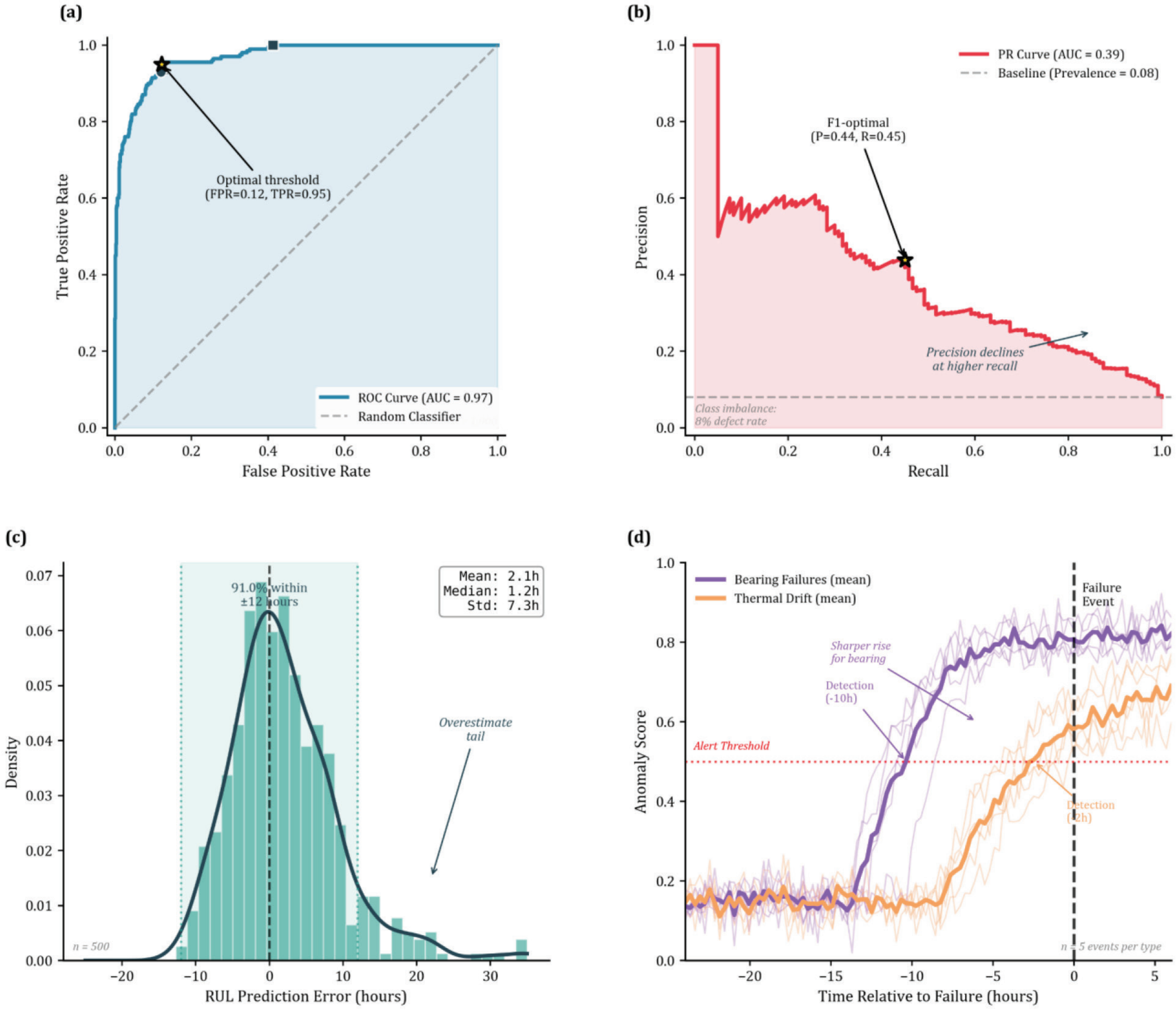
The failure prediction model performed better than the defect risk model. It gave a median lead time of 18 hours before a failure happened. The defect risk model had lower precision and recall. It didn't catch many rare defects when the factory made many different products. Machine signals stayed more consistent than quality signals. Quality signals often changed because of short-term differences in materials.

The present investigation looked at patterns and timing in Figure 6. Figure 6(a) shows the receiver operating characteristic curve for failure prediction. Figure 6(b) shows the precision recall curve for defect risk prediction. Figure 6(c) shows the distribution of RUL prediction errors. Figure 6(d) shows how anomaly score trajectories moved near failure events.

Figure 6(a) shows a steep curve. This matches the high discrimination results in Table 5. Figure 6(b) shows a flatter curve. This proves the defect model didn't do as well because of class imbalance. Figure 6(c) shows that 72 percent of RUL errors stayed within ±12 hours. A few larger errors happened during shifts with unusual product mixes. Figure 6(d)

Table 5. Predictive performance for failure and defect risk models.

Task	Area under ROC curve mean ± SD	Precision	Recall	F1 score	Median lead time
Failure prediction	0.90 ± 0.03	0.76	0.71	0.73	18 h
Defect risk prediction	0.84 ± 0.04	0.68	0.63	0.65	9 h



Note. Panel (a) shows ROC curves for failure prediction. Panel (b) presents precision recall curves for defect risk prediction. Panel (c) reports the distribution of RUL prediction errors. Panel (d) plots anomaly score trajectories around failure events.

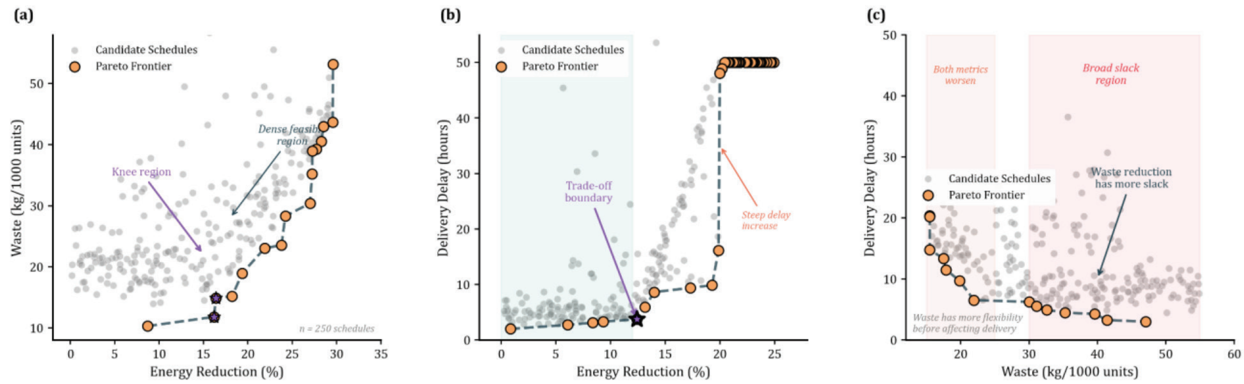
Figure 6. Predictive model diagnostics and temporal behavior.

shows that anomaly scores rose several hours before failures. The rise was sharper for bearing problems than for thermal drift. The maintenance model gave timely signals that helped with scheduling. The defect model didn't catch every quality shift.

3.8 Scheduling and optimization performance

This study analyzed multi-objective scheduling outcomes using solution sets from the genetic algorithm. This study presents the analysis in Figure 7 across three panels. Figure 7(a) plots energy versus waste for candidate schedules. Figure 7(b) plots energy versus delivery delay. Figure 7(c) plots waste versus delivery delay.

Figure 7(a) shows a clear Pareto frontier. Reductions in energy linked to modest increases in waste beyond a threshold. This pattern points to a trade-off zone that guided weight selection. Figure 7(b) shows that energy reductions beyond 12 percent tied to a sharp rise in delay. This limited strong energy cuts. Figure 7(c) shows a wider set of schedules that cut waste without big delay jumps. This suggests the cir-



Note. Panel (a) shows energy versus waste. Panel (b) shows energy versus delay. Panel (c) shows waste versus delay.

Figure 7. Pareto trade-offs among energy, waste, and delivery delay.

cularity goal connected less closely to delivery than energy did. These trends back the pick of balanced weights in the final setup.

This study compared scheduling performance across methods. This study reports the analysis in Table 6. Table 6 lists queue length, tardiness, changeovers, and utilization for rule-based scheduling, energy-only optimization, waste-only optimization, and the proposed multi-objective strategy.

The multi-objective strategy cut average queue length and tardiness compared to the rule-based method. It kept changeovers close to the starting level. Energy-only optimization raised tardiness and queue length. This fit the trade-offs in Figure 7. Waste-only optimization improved queue length but raised changeovers. This added more setup time. The multi-objective strategy gave the most even results. This matched the aim to blend circularity, energy, and delivery goals.

This study tested sensitivity to objective weights to see how trade-offs changed with different priorities. This study reports the analysis in Table 7.

The weight sensitivity results show that a focus on waste gave the lowest scrap mass with just a small rise in energy use. A focus on energy cut energy use but raised scrap and delay. This echoed the trade-offs in Figure 7. A focus on delay gave the shortest tardiness but higher energy use than the waste focus. The spread of results shows that the system let users adjust to work needs without harming stability.

3.9 Resilience to disruptions

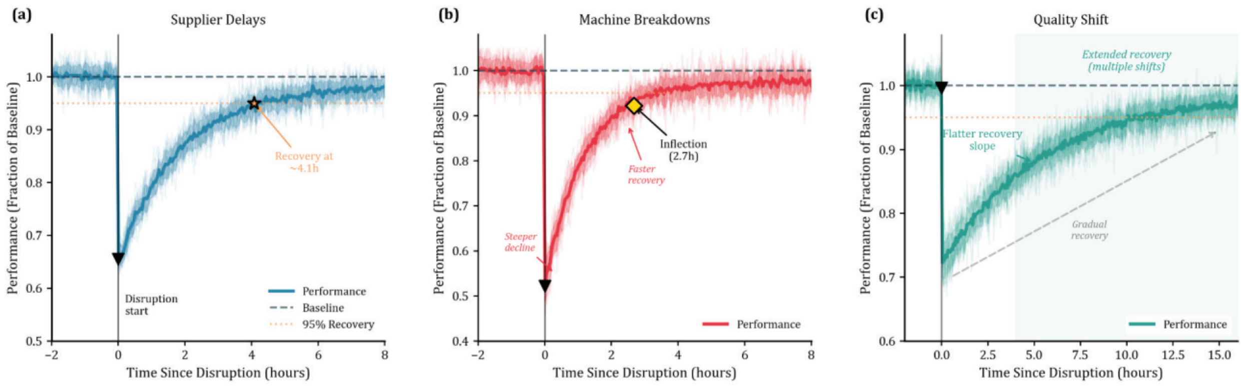
This study analyzed disruption response through performance trajectories under three scenarios. The analysis appears in Figure 8 across three panels. Figure 8(a) shows performance over time during supplier delay events. Figure 8(b) shows recovery following machine breakdowns. Figure 8(c) shows performance during quality shifts.

Table 6. Scheduling performance by strategy.

Strategy	Average queue length (units)	Mean tardiness (h)	Changeovers per shift	Machine utilization
Rule-based	182	7.4	5.6	0.81
Energy-only	195	8.9	5.2	0.80
Waste-only	168	6.1	6.4	0.82
Multi-objective	160	6.0	5.7	0.83

Table 7. Weight sensitivity analysis for multi-objective scheduling.

Weight setting (energy, waste, delay)	Energy intensity (kWh per unit) mean ± SD	Scrap mass (kg per 1,000 units) mean ± SD	Mean tardiness (h) mean ± SD
0.5, 0.3, 0.2	4.86 ± 0.44	13.7 ± 2.5	6.8 ± 1.1
0.3, 0.5, 0.2	5.02 ± 0.47	12.1 ± 2.3	6.4 ± 1.0
0.3, 0.2, 0.5	5.14 ± 0.52	12.8 ± 2.6	5.6 ± 0.9



Note. Panel (a) shows supplier delay recovery. Panel (b) shows machine breakdown recovery. Panel (c) shows quality shift recovery.

Figure 8. Performance trajectories during disruptions.

Figure 8(a) shows that performance recovered to 95 percent of the reference level within 3.8 hours during supplier delays. That recovery happened faster than the more prolonged one in the baseline period. Figure 8(b) shows a faster recovery after machine breakdowns. The steepest recovery slope occurred within the first two hours after intervention actions started. Figure 8(c) shows a shallower recovery profile during quality shifts. A longer tail persisted across multiple shifts. The scenario comparison indicates stronger resilience improvements for supply and equipment shocks than for quality deviations. Those improvements reflected the slower identification of material-driven defects (Table 8).

Resilience scores improved notably for supplier delays and machine breakdowns. Recovery times dropped by 42 percent and 41 percent respectively. The quality shift scenario showed a smaller change. The overlap in standard deviations indicates a weaker effect in that case. The negative result for quality shifts aligned with the slower response of the defect risk model in Table 5. That slower response limited the speed of corrective action. The pattern indicates that the framework provided stronger resilience benefits for disruptions driven by equipment and supply variability than for process quality drift.

3.10 Human-centric outcomes

Human-centric performance was assessed through task completion, perceived trust, perceived control, workload, and error rates (Table 9). Clarity scores were high across all participant groups, and perceived control remained above the scale midpoint for operators, maintenance technicians, and supervisors, indicating that the interface supported decision comprehension without displacing user authority. Supervisors completed tasks most rapidly, which likely reflects greater familiarity with production planning and exception handling, whereas operators required more time to interpret the recommendation context. Task error rates remained below 7% for all roles, suggesting that the interface did not introduce a substantial usability burden.

Supervisors finished tasks the fastest. Operators took the longest. This happened because supervisors were more familiar with the planning tools. Errors stayed below 7 percent for everyone. This shows the interface didn't cause new mistakes. This pattern shows the human-centric features helped people make decisions. They didn't make the brain work too hard.

Table 8. Resilience outcomes by disruption scenario.

Scenario	Resilience score baseline mean ± SD	Resilience score intervention mean ± SD	Recovery time baseline (h) mean ± SD	Recovery time intervention (h) mean ± SD	Performance loss baseline (%)	Performance loss intervention (%)
Supplier delay	0.72 ± 0.06	0.83 ± 0.05	6.5 ± 1.4	3.8 ± 1.0	18.2	10.4
Machine breakdown	0.70 ± 0.07	0.82 ± 0.06	9.2 ± 2.1	5.4 ± 1.6	22.6	13.8
Quality shift	0.75 ± 0.05	0.77 ± 0.05	5.1 ± 1.3	4.7 ± 1.2	15.0	13.9

Table 9. Human-centric performance and perception metrics by role.

Role	Clarity score mean ± SD (1-5)	Trust score mean ± SD (1-5)	Perceived control mean ± SD (1-5)	Workload score mean ± SD (1-5)	Task time (min) mean ± SD	Task errors (%)
Operators	4.1 ± 0.5	3.7 ± 0.6	4.0 ± 0.5	2.6 ± 0.4	7.1 ± 1.4	6.2
Maintenance technicians	3.9 ± 0.6	3.9 ± 0.5	3.8 ± 0.5	2.8 ± 0.5	6.4 ± 1.1	5.4
Supervisors	4.3 ± 0.4	3.8 ± 0.6	4.2 ± 0.4	2.5 ± 0.3	5.8 ± 0.9	4.9

Figure 9 complements the perceptual measures by reporting task time distributions and alert handling behavior across roles. Panel (a) summarizes completion-time variability, and panel (b) reports acceptance and override rates for DT recommendations.

Supervisors accepted approximately 72% of alerts and operators approximately 68%, while override rates remained below 15% for both groups. Interview coding indicated that overrides were typically associated with context-specific production contingencies or quality judgments that users considered insufficiently represented in the current model. Rather than undermining the framework, these interventions illustrate the intended human-in-the-loop logic: the DT structured the decision space, but final authority remained with experienced personnel. The observed feedback also informed subsequent refinement of alert thresholds and explanatory displays.

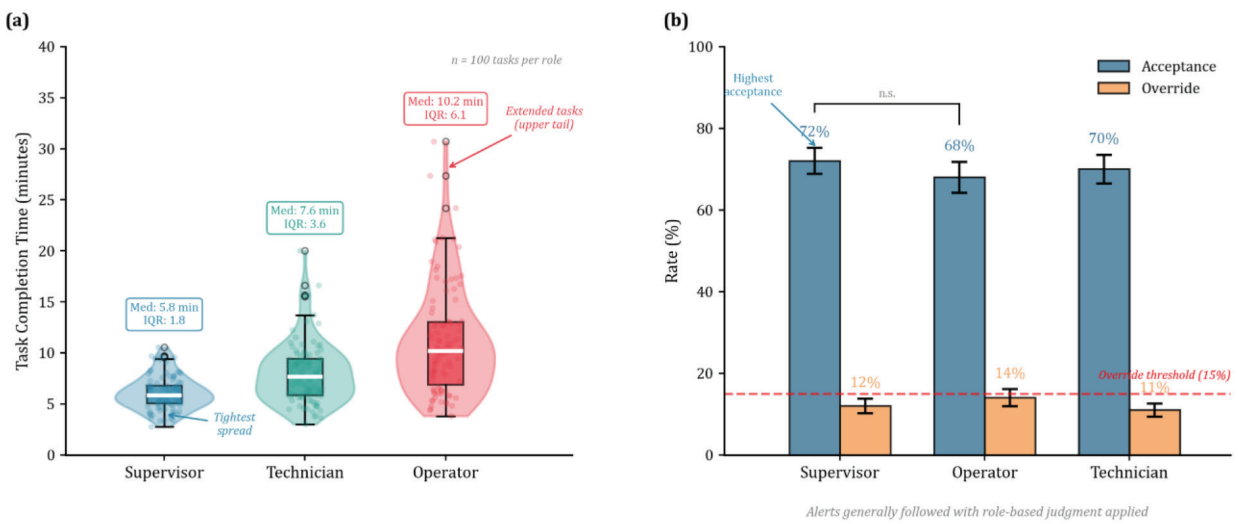
3.11 Benchmarking, ablation, and overall performance synthesis

This research compared overall performance against baseline settings and other methods. Table 10 shows the results. This table lists mean ± SD values and percent improvements for core outcomes. It compares these to a baseline that only monitors the system.

The benchmarking results show that this framework created the strongest gains for energy, circularity, and resilience. A DT without circular optimization lowered energy use and helped resilience. However, it didn't change circularity much. This means that monitoring alone didn't fix material outcomes. Manual circular rules helped circularity. They didn't really lower energy use or downtime. The proposed framework showed the largest gains in circularity and energy. It also improved OEE and lowered downtime. These results match the multi-objective integration strategy. This work shows that a combined approach offers more benefits than a single-focus baseline.

This study analyzed ablation outcomes to see how specific parts helped. Figure 10 shows this across three panels. Figure 10(a) shows circularity rates for the full framework and each test condition. Figure 10(b) shows energy intensity. Figure 10(c) shows resilience scores.

Figure 10(a) shows that removing the circular objective lowered circularity rates. They fell back to the

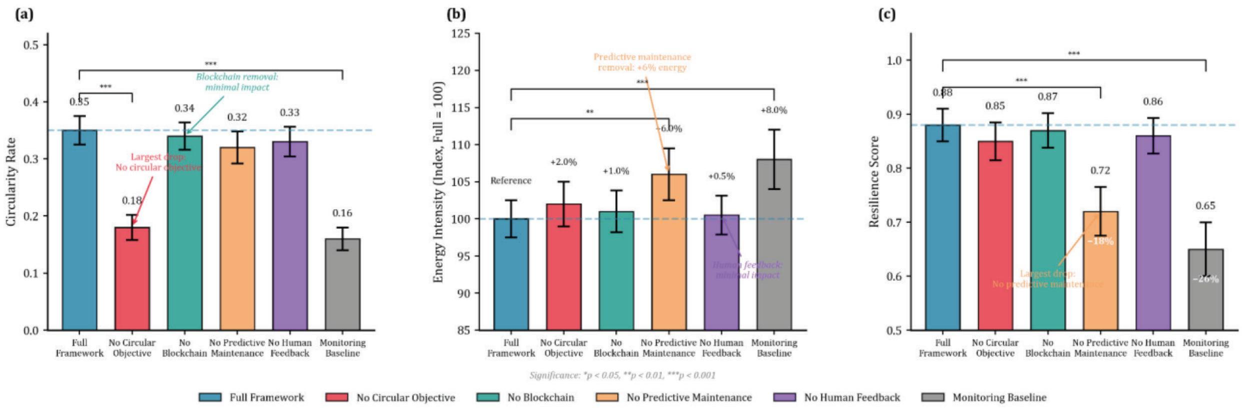


Note. Panel (a) shows task completion time distributions by role. Panel (b) shows alert acceptance and override rates by role.

Figure 9. Human-centric task behavior and alert handling.

Table 10. Overall benchmarking against baseline configurations.

Method	Energy intensity (kWh per unit)	Circularity rate	Resilience score	OEE	Downtime per week (h)	Improvement versus baseline (%)
Monitoring-only baseline	5.91 ± 0.60	0.23 ± 0.04	0.71 ± 0.06	0.775 ± 0.040	14.6 ± 3.8	NA
DT without circular optimization	5.28 ± 0.54	0.24 ± 0.04	0.78 ± 0.06	0.806 ± 0.035	12.2 ± 3.1	Energy -10.7, Resilience +9.9, OEE +4.0
Manual circular rules	5.47 ± 0.57	0.29 ± 0.05	0.73 ± 0.05	0.789 ± 0.038	13.8 ± 3.4	Circularity +26.1
Proposed framework	4.98 ± 0.49	0.33 ± 0.06	0.82 ± 0.06	0.821 ± 0.031	10.8 ± 2.9	Energy -15.7, Circularity +43.5, Resilience +15.5, OEE +5.9



Note. Panel (a) presents circularity rates. Panel (b) presents energy intensity. Panel (c) presents resilience scores.

Figure 10. Ablation outcomes for key framework components.

monitoring baseline level. Removing the blockchain ledger didn't change circularity outcomes much. Figure 10(b) shows that taking out the predictive maintenance module raised energy intensity by about 6 percent. This happened because of unplanned stops and start-up cycles. Removing human feedback didn't really change energy results. Figure 10(c) shows the biggest drop in resilience when the research removed predictive maintenance. Resilience dropped a smaller amount when the scheduler was gone. These results show that predictive maintenance and the multi-objective scheduler helped resilience most. The circular objective was the main driver for circularity performance.

The integrated DT altered production control in a substantively different way from a monitoring-only implementation. Because synchronization, predictive maintenance, provenance tracking, and scheduling were coupled within the same loop, the framework could evaluate operational alternatives against circularity and resilience objectives before execution. The simultaneous reduction in energy intensity and downtime, together with the increase in OEE and circularity, indicates that the benefits were not confined to a single performance dimension. These findings support the central proposition of the study: a DT becomes a decision engine when it not only represents the physical line but also conditions, ranks, and explains candidate actions under human review.

This interpretation also sharpens the manuscript's Industry 5.0 contribution. The framework did not invoke sustainability, resilience, and human-centricity as abstract principles; rather, it operationalized them through measurable energy and recovery metrics, a traceable provenance mechanism, and explicit operator acceptance or override of recommendations. The smaller improvement observed in quality-shift resilience further indicates that human oversight remains important when model sensitivity is weaker, reinforcing the argument that Industry 5.0 production systems require collaborative rather than fully autonomous decision support.

Several limitations delimit the external validity of the results. The pilot covered one electronics manufacturing line in Riyadh over six months, so the magnitudes of the reported gains may depend on the product mix, recovery pathways, supplier structure, and operational maturity of that site. Heavy manufacturing environments may offer larger component-level reuse opportunities but slower scheduling cycles, whereas process industries involve continuous flows and different traceability granularity; in both cases, the underlying architecture may transfer more readily than the numerical performance gains or weight settings. In addition, although efforts were made to reduce confounding, the pre-post evaluation cannot fully eliminate residual effects from seasonal demand variation, operator learning, or concurrent incremental improvements. Future studies should therefore examine multi-site deployments, longer observation windows, and comparative evaluations against alternative optimization strategies.

4. Conclusions

This study demonstrates that a DT can support Industry 5.0 production as an active decision system rather than as a passive representation of shop-floor states. The contribution lies in the integration of live sensing, credibility-gated synchronization, provenance tracking, predictive maintenance, multi-objective scheduling, and human authorization within a single control architecture. The empirical results indicate that such an architecture can improve energy performance, circularity, and disruption recovery simultaneously while preserving product quality. The component-level results also show that circularity gains are materially contingent: recovery improvements were strongest where reuse and recycling pathways were operationally feasible, and more modest where component constraints remained stringent. From a research perspective, the study contributes a more precise account of how circularity and resilience can be embedded in production control rather than evaluated only after execution. From a managerial perspective, it shows that credibility checks, transparent objective trade-offs, and operator override mechanisms are necessary conditions for deploying prescriptive DT systems in a manner consistent with Industry 5.0 principles.

Further work should test the framework across additional sectors, extend the observation period, and compare alternative optimization and prediction approaches under shared evaluation protocols. Improved sensing and modeling of quality deviations will be particularly important for strengthening resilience under material-driven disturbances and for refining the human-in-the-loop balance of future DT deployments.

Disclosure

During the preparation of this work, the authors used ChatGPT to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

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

References

[1] R. A. Sheikh, I. Ahmed, A. Y. A. Faqihi, and Y. M. Shehawy, “Global perspectives on navigating Industry 5.0 knowledge: Achieving resilience, sustainability, and human-centric innovation in manufacturing,” J. Knowl. Econ., vol. 16, pp. 15997–16032, 2025, doi: 10.1007/s13132-024-02498-4.

[2] C. Castillo, T. Otero-Romero, and E. J. Alvarez-Palau, “Navigating the transition to Industry 5.0: Advancing sustainability, resilience, and human-centricity in Spanish supply chain management,” Discov. Sustain., vol. 6, p. 331, 2025, doi: 10.1007/s43621-025-01190-0.

[3] N. Fernandes, J.-P. Barros, and R. Campos-Rebelo, “Graphic model for shop floor simulation and control in the context of Industry 5.0,” Appl. Sci., vol. 13, no. 2, p. 930, 2023, doi: 10.3390/app13020930.

[4] M. Ghobakhloo, H. A. Mahdiraji, M. Iranmanesh, et al., “From Industry 4.0 digital manufacturing to Industry 5.0 digital society: A roadmap toward human-centric, sustainable, and resilient production,” Inf. Syst. Front., 2024, doi: 10.1007/s10796-024-10476-z.

[5] H. S. T. Pham and S. B. Li, “Towards Industry 5.0: A conceptual model for leading organisational change in digital age,” J. Organ. Change Manag., vol. 39, no. 1, pp. 125–161, 2026, doi: 10.1108/JOCM-03-2025-0264.

[6] L. Jum’a, M. Ikram, and C. J. C. Jabbour, “Towards circular economy: A IoT enabled framework for circular supply chain integration,” Comput. Ind. Eng., vol. 192, p. 110194, 2024, doi: 10.1016/j.cie.2024.110194.

[7] M. Paramesha, N. L. Rane, and J. Rane, “Big data analytics, artificial intelligence, machine learning, Internet of Things, and blockchain for enhanced business intelligence,” Partners Univers. Multidiscip. Res. J., vol. 2, no. 3, 2024, doi: 10.5281/zenodo.12827323.

[8] I. Ahmed, Y. Zhang, G. Jeon, W. Lin, M. R. Khosravi, and L. Qi, “A blockchain- and artificial intelligence-enabled smart IoT framework for sustainable city,” Int. J. Intell. Syst., vol. 37, pp. 6493–6507, 2022, doi: 10.1002/int.22852.

[9] O. Okorie, J. J. Russell, R. Cherrington, O. Fisher, and F. Charnley, “Digital transformation and the circular economy: Creating a competitive advantage from the transition towards Net Zero Manufacturing,” Resour.

Conserv. Recycl., vol. 189, p. 106756, 2023, doi: 10.1016/j.resconrec.2022.106756.
[10] J. F. Yao, Y. Yang, X. C. Wang, et al., "Systematic review of digital twin technology and applications," Vis. Comput. Ind. Biomed. Art, vol. 6, p. 10, 2023, doi: 10.1186/s42492-023-00137-4.
[11] Z. Song et al., "Digital twins for the future power system: An overview and a future perspective," Sustainability, vol. 15, no. 6, p. 5259, 2023, doi: 10.3390/su15065259.
[12] S. Mihai et al., "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," IEEE Commun. Surveys Tuts., vol. 24, no. 4, pp. 2255–2291, 2022, doi: 10.1109/COMST.2022.3208773.
[13] P. Singh, K. Krishnan, and E. Boldsaikhan, "A review of production scheduling with artificial intelligence and digital twins," J. Manuf. Mater. Process., vol. 10, no. 1, p. 6, 2026, doi: 10.3390/jmmp10010006.
[14] M. van Dyck, D. Lüttgens, F. T. Piller, and S. Brenk, "Interconnected digital twins and the future of digital manufacturing: Insights from a Delphi study," J. Prod. Innov. Manag., vol. 40, no. 4, pp. 475–505, 2023, doi: 10.1111/jpim.12685.
[15] P. V. dos Santos Gonçalves and L. M. S. Campos, "A systemic review for measuring circular economy with multi-criteria methods," Environ. Sci. Pollut. Res., vol. 29, pp. 31597–31611, 2022, doi: 10.1007/s11356-022-18580-w.
[16] S. Werbińska-Wojciechowska, R. Giel, and K. Winiarska, "Digital twin approach for operation and maintenance of transportation system—Systematic review," Sensors, vol. 24, no. 18, p. 6069, 2024, doi: 10.3390/s24186069.
[17] Z. A. Ali, M. Zain, R. Hasan, et al., "Digital twins: Cornerstone to circular economy and sustainability goals," Environ. Dev. Sustain., 2025, doi: 10.1007/s10668-025-06221-4.
[18] A. Florek-Paszkowska and A. Ujwary-Gil, "The digital-sustainability ecosystem: A conceptual framework for digital transformation and sustainable innovation," J. Entrep. Manag. Innov., vol. 21, no. 2, pp. 116–137, 2025, doi: 10.7341/20252127.
[19] F. Aiello, P. C. Cozzucoli, L. Mannarino, and V. Pupo, "Bayesian insights on digitalization and environmental sustainability practices: Towards the twin transition in the EU," Bus. Strategy Environ., vol. 34, no. 1, pp. 417–432, 2025, doi: 10.1002/bse.3985.
[20] V. M. Freires, R. H. P. Silva, A. M. Lucas, and E. S. Almeida, "Development of a didactic solution for teaching concepts related to digital twins using educational robot," Int. J. Ind. Eng. Manag., vol. 16, no. 4, pp. 444–458, 2025, doi: 10.24867/IJIEM-398.
[21] O. Mukhitdinov, D. Jumanazarov, E. Khudoynazarov, L. Safarova, S. Sakhabayeva, and A. M. Alsayah, "An Industry 4.0 framework for the smart production management of renewable energy and water systems: An application of AI, IoT, and digital twin technologies," Int. J. Ind. Eng. Manag., vol. 16, no. 4, pp. 428–443, 2025, doi: 10.24867/IJIEM-397.
[22] D. Zhong, Z. Xia, Y. Zhu, and J. Duan, "Overview of predictive maintenance based on digital twin technology," Heliyon, vol. 9, no. 4, p. e14534, 2023, doi: 10.1016/j.heliyon.2023.e14534.
[23] T. Harries, M. Hartnoll, M. Hafezianrazavi, H. Meek, and A. Nassehi, "Digital twins for predictive maintenance," Procedia CIRP, vol. 118, pp. 306–311, 2023, doi: 10.1016/j.procir.2023.06.053.
[24] P. Aivaliotis, K. Georgoulias, and G. Chryssolouris, "The use of digital twin for predictive maintenance in manufacturing," Int. J. Comput. Integr. Manuf., vol. 32, no. 11, pp. 1067–1080, 2019, doi: 10.1080/0951192X.2019.1686173.
[25] E. Jovicic, D. Primorac, M. Cupic, and A. Jovic, "Publicly available datasets for predictive maintenance in the energy sector: A review," IEEE Access, vol. 11, pp. 73505–73520, 2023, doi: 10.1109/ACCESS.2023.3295113.
[26] G. Shao, "Use case scenarios for digital twin implementation based on ISO 23247," Natl. Inst. Stand. Technol., Gaithersburg, MD, USA, pp. 400–402, 2021.
[27] Z. Huang, M. Fey, C. Liu, E. Beysel, X. Xu, and C. Brecher, "Hybrid learning-based digital twin for manufacturing process: Modeling framework and implementation," Robot. Comput.-Integr. Manuf., vol. 82, p. 102545, 2023, doi: 10.1016/j.rcim.2023.102545.
[28] J. Mädler, I. Viedt, J. Lorenz, and L. Urbas, "Requirements to a digital twin-centered concept for smart manufacturing in modular plants considering distributed knowledge," Comput. Aided Chem. Eng., vol. 49, pp. 1507–1512, 2022, doi: 10.1016/B978-0-323-85159-6.50251-7.
[29] A. Löcklin, M. Müller, T. Jung, N. Jazdi, D. White, and M. Weyrich, "Digital twin for verification and validation of industrial automation systems—A survey," in Proc. 25th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA), Vienna, Austria, 2020, pp. 851–858, doi: 10.1109/ETFA46521.2020.9212051.
[30] E. VanDerHorn, B. Corbett, and A. Costa, "Establishing credibility: Verification and validation of digital twins," in Offshore Technology Conf. (OTC), Houston, TX, USA, May 2025, doi: 10.4043/35756-MS.
[31] G. Shao, J. Hightower, and W. Schindel, "Credibility consideration for digital twins in manufacturing," Manuf. Lett., vol. 35, pp. 24–28, 2023, doi: 10.1016/j.mfglet.2022.11.009.
[32] R. Takkar, K. Birman, and H. O. Gao, "Enhancing transparency in buyer-driven commodity chains for complex products: Extending a blockchain-based traceability framework towards the circular economy," Appl. Sci., vol. 15, no. 15, p. 8226, 2025, doi: 10.3390/app15158226.
[33] P. Centobelli, R. Cerchione, P. Del Vecchio, E. Oropallo, and G. Secundo, "Blockchain technology for bridging trust, traceability and transparency in circular supply chain," Inf. Manag., vol. 59, no. 7, p. 103508, 2022, doi: 10.1016/j.im.2021.103508.
[34] G. Moraga et al., "Circular economy indicators: What do they measure?," Resour. Conserv. Recycl., vol. 146, pp. 452–461, 2019, doi: 10.1016/j.resconrec.2019.03.045.
[35] P. P. Poveda-Orjuela, J. C. García-Díaz, A. Pulido-Rojano, and G. Cañón-Zabala, "ISO 50001:2018 and its application in a comprehensive management system with an energy-performance focus," Energies, vol. 12, no. 24, p. 4700, 2019, doi: 10.3390/en12244700.
[36] C. Andersson and M. Bellgran, "On the complexity of using performance measures: Enhancing sustained production improvement capability by combining OEE and productivity," J. Manuf. Syst., vol. 35, pp. 144–154, 2015, doi: 10.1016/j.jmsy.2014.12.003.
[37] H. Hijry, "Advancing sustainable smart manufacturing: A comprehensive review of machine learning techniques in assembly lines," Sustainability, vol. 18, no. 1, p. 348, 2026, doi: 10.3390/su18010348.
[38] W. Zhang et al., "Enhancing multi-objective evolutionary algorithms with machine learning for scheduling problems: Recent advances and survey," Front. Ind. Eng., vol. 2, p. 1337174, 2024.
[39] M. Alolaiwy, T. Hawsawi, M. Zohdy, A. Kaur, and S. Louis, "Multi-objective routing optimization in electric and flying vehicles: A genetic algorithm perspective," Appl. Sci., vol. 13, no. 18, p. 10427, 2023, doi: 10.3390/app131810427.
[40] A. Alrabghi, "A modelling approach for asset degradation: Advancing digital twin in maintenance," Int. J. Simul.

Model., vol. 24, no. 1, pp. 76–86, 2025, doi: 10.2507/IJSIMM24-1-715.
[41] R. Askar et al., "Driving the built environment twin transition: Synergising circular economy and digital tools," in Circular Economy Design and Management in the Built Environment, L. Bragança et al., Eds. Cham, Switzerland: Springer, 2025, doi: 10.1007/978-3-031-73490-8_17.
[42] P. K. Thakuri, L. Alkki, and L. Aarikka-Stenroos, "Digital technologies enabling component reuse in circular value chains: Using digital twin, Internet of Things and robots in construction and manufacturing sectors," R&D Manag., vol. 55, pp. 1161–1216, 2025, doi: 10.1111/radm.12744.