information

MDPI

Article

A Digital Twin Architecture for Integrating Lean Manufacturing with Industrial IoT and Predictive Analytics

Gulshat Amirkhanova [1], Shyrailym Adilkyzy [1,*], Bauyrzhan Amirkhanov [1], Dina Baizhanova [1] and Siming Chen [2]

[1] Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan; gulshat.amirkhanova@kaznu.edu.kz (G.A.); baur@fax.kz (B.A.); dina.baizhanova@kaznu.edu.kz (D.B.)
[2] School of Data Science, Fudan University, Shanghai 200433, China; simingchen@fudan.edu.cn
* Correspondence: adilkyzy_shyrailym@kaznu.edu.kz

Abstract

The convergence of Lean manufacturing and Industry 4.0 requires digital infrastructures capable of transforming high-frequency telemetry into actionable insights. However, architectures that integrate near real-time data with closed-loop process control remain scarce, particularly in the food-processing industry. This study proposes a "Lean 4.0" framework based on a six-layer Digital Twin (DT) architecture to digitise waste detection and optimise a medium-scale bakery. The methodology integrates a heterogeneous Industrial Internet of Things (IIoT) network comprising 17 ESP32 (Espressif Systems, Shanghai, China)-based monitoring nodes. Data collection is managed via an edge-centric MQTT–InfluxDB (version 2.7, InfluxData, San Francisco, CA, USA) data pipeline. Furthermore, the analytics layer employs discrete-event simulation in Siemens Plant Simulation (version 2302, Siemens Digital Industries Software, Plano, TX, USA), constraint programming with Google OR-Tools (version 9.8, Google LLC, Mountain View, CA, USA), and machine learning models (Isolation Forest and SARIMA). Multi-month validation in a brownfield bakery, including a 60-day continuous monitoring test, demonstrated that the proposed architecture reduced production cycle time by 24.4% and inter-operational waiting time by 51.2%. Moreover, manual planning time decreased by 87.4% through the use of low-code scheduling interfaces. In addition, state-based control of critical ovens reduced energy consumption by 23.06%. These findings indicate that combining deterministic simulation and combinatorial optimisation with data-driven analytics provides a scalable blueprint for implementing cyber-physical systems in food-processing SMEs. This approach effectively bridges the gap between traditional Lean principles and data-driven smart manufacturing.

Keywords: digital twin; lean manufacturing; industrial internet of things (IIoT); predictive analytics; hybrid flow shop; energy optimisation; machine learning; bakery manufacturing; cyber physical systems

Check for updates

Academic Editor: Yu-Cheng Wang

Received: 17 December 2025
Revised: 3 February 2026
Accepted: 5 February 2026
Published: 13 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The paradigm of Industry 4.0 is reshaping manufacturing by tightly coupling physical assets with their digital counterparts. At the core of this transformation lies the concept of the Digital Twin (DT)—a dynamic virtual representation that mirrors the state and behaviour of physical systems and enables continuous monitoring, performance evaluation and optimisation across the life cycle of industrial assets. Compared with traditional production management, which is often based on static models and offline analyses, DT-driven

environments require the seamless integration of high-velocity telemetry, simulation models and decision-support tools to maintain an up-to-date view of complex manufacturing systems [1].

In parallel, production management is moving from paper-based routines and aggregated indicators towards data-oriented shopfloor management. Recent work highlights that modern shopfloors must exploit granular, time-stamped data to support short feedback loops, systematic problem solving and transparent performance tracking [2]. This evolution is closely related to the emergence of "Lean 4.0", in which digital technologies reinforce established Lean Manufacturing principles—flow, pull and waste elimination—rather than replacing them [3]. For small and medium-sized enterprises (SMEs), particularly in emerging economies, this convergence remains challenging due to limited investment capacity, fragmented legacy systems and shortages of digital skills [4].

The food-processing sector adds further complexity. Under the emerging "Food Sustainability 4.0" paradigm, food manufacturers are expected not only to increase productivity and responsiveness but also to improve resource efficiency and environmental performance along their value chains [5]. Unlike many discrete manufacturing domains, food production—such as bakery lines—is characterised by continuous or hybrid flows, strict perishability windows and tight coupling between material handling and energy-intensive operations. As a consequence, production planning and control must simultaneously address throughput, quality and energy consumption while respecting product-specific time constraints and hygiene regulations. Recent studies emphasise the transition towards "Food Sustainability 4.0", where energy-aware scheduling becomes a critical component of Lean operations. In this context, detecting anomalies in industry-specific power data is a complex but necessary prerequisite for robust energy management in non-stationary food-processing environments [6]. Furthermore, the democratisation of Digital Twins for SMEs in emerging economies is gaining momentum, moving away from expensive proprietary systems towards modular, low-code architectures that bypass traditional infrastructure barriers [7,8].

Energy management is therefore a critical dimension of smart food manufacturing. Industrial power profiles in such environments are highly non-stationary and influenced by batch sizes, recipe changes and operator behaviour. Detecting anomalies and inefficiencies in such industry-specific power data is far from trivial and requires tailored algorithms that account for process knowledge and operating regimes [9]. At the same time, transmitting all high-frequency measurements to a central server can introduce latency and communication overhead. Edge-computing approaches perform signal segmentation and anomaly detection directly on embedded devices. They are increasingly recognised as a means to reduce latency, enhance resilience and alleviate bandwidth constraints [10]. Yet these solutions are seldom integrated into unified architectures that also address production planning, Lean performance indicators and closed-loop control.

In operations management and Lean implementation, combining simulation with improvement tools has proved effective for analysing bottlenecks, evaluating "what-if" scenarios and supporting redesign decisions. Frameworks such as SimLean extend traditional simulation by explicitly embedding Lean tools and problem-solving routines, providing structured guidance for waste identification and process redesign [11]. However, most existing DT- and SimLean-type applications focus on discrete manufacturing sectors and on offline analysis rather than near-real-time integration with Industrial Internet of Things (IIoT) infrastructures [12]. While our approach adopts the SimLean philosophy, it differs significantly by replacing heuristic logic with deterministic CP-SAT optimisation to resolve scheduling conflicts mathematically, providing better scalability for parallel-machine environments. There is a lack of comprehensive architectures that (i) continuously fuse IIoT data

with discrete-event simulation models, (ii) exploit combinatorial optimisation for complex scheduling problems and (iii) feed decisions back into the shopfloor through closed-loop control—especially for continuous or hybrid flow processes such as bakery lines.

Current work often focuses on conceptual DT architectures or Lean 4.0 adoption factors [13]. However, few studies provide an integrated, energy-aware DT that combines IIoT sensing, discrete-event simulation, Hybrid Flow Shop optimisation and machine learning in a food-processing SME. The comparison of the proposed framework with existing studies is presented in Table 1.

Table 1. Positioning of the proposed Lean 4.0 Digital Twin relative to representative related work.

Study	Domain/Scope	Real IIoT/DT Implementation	Simulation/ Optimisation	Energy Analytics and Control	SME/Low-Code and Lean 4.0 Focus
Tao et al.	Generic industrial DT frameworks	Conceptual/ review	Conceptual discussion of DT capabilities	Not specifically addressed	Not SME- or low-code-specific
Eichenseer & Winkler	Data-oriented shopfloor management	Conceptual framework + cases	Focus on KPIs and management routines	Energy treated at KPI/indicator level	SME challenges discussed conceptually
Qureshi et al.	Lean 4.0 adoption factors in SMEs	Survey/ modelling	No detailed DT or scheduling models	Not explicitly modelled	Strong focus on SME Lean 4.0 readiness
Thenarasu et al.—SimLean-TRIZ	Cold-room door production (manufact.)	Simulation-based improvement	SimLean + TRIZ simulation for waste	No integrated energy analytics or control	Not focused on food-processing or low-code
This work (Lean 4.0 DT for bakery)	Food-processing SME (bakery line)	Deployed six-layer DT with IIoT	DES-based bottleneck analysis + HFSP via CP-SAT	Integrated anomaly detection, forecasting and state-based energy control for ovens	Explicit SME case with low-code planning apps and Lean 4.0 KPIs

A further gap concerns the practical deployment of such systems in Central Asia, where many industrial enterprises still operate as "islands of automation" with heterogeneous equipment and limited interoperability. In this context, an end-to-end Digital Twin is not only a technical challenge but also a strategic enabler for Lean 4.0 transformation. Bridging deterministic simulation models with stochastic, real-world data streams can provide decision-makers with both predictive insights and robust optimisation recommendations, thereby supporting more systematic waste reduction, energy efficiency and planning reliability.

To address these gaps, this study presents a hybrid Digital Twin architecture that integrates Lean Manufacturing principles with IIoT-enabled data acquisition and predictive analytics in a medium-scale bakery facility in Kazakhstan. The proposed "Lean 4.0" framework combines a heterogeneous sensor network (ESP32-based nodes and Saiman (Saiman Corporation, Almaty, Kazakhstan) energy meters) with a six-layer DT stack covering data acquisition, edge processing, cloud-level storage, analytics, simulation and visualisation. Discrete-event simulation in Siemens Plant Simulation is used for bottleneck analysis and capacity planning, while a CP-SAT solver from Google OR-Tools formulates and solves a Hybrid Flow Shop Problem (HFSP) to optimise production schedules under resource and sequence-dependent constraints. Near-real-time telemetry further drives state-based control logic for critical assets, enabling dynamic adjustment of operating modes according

to demand and energy conditions. As summarised in Table 2, to the best of our knowledge there is currently no six-layer Lean 4.0 Digital Twin for bakery manufacturing that integrates IIoT telemetry, DES-based bottleneck analysis, HFSP scheduling, energy-aware oven control and low-code planning tools in a single, validated architecture.

Table 2. Technological stack of the six-layer Digital Twin architecture implemented in the bakery case study.

Layer	Component	Implementation in the Bakery DT
Physical	Bakery equipment	Batch dough mixers; bun- and bread-moulding stations; six industrial ovens (Wachtel, MIWE, BlackTechnik/Wachtel); six proofers (incl. Wachtel Aeromat); cooling conveyors and stations; ice generators.
Sensing	Monitoring nodes	Seventeen ESP32-based IIoT nodes with SAIMAN Dala CA4-E720 R three-phase meters and Orman single-phase meters, connected via RS-485/Modbus RTU; optional DHT22 climate sensors on selected assets.
Network	Communication	Modbus RTU between meters and ESP32; MQTT over Wi-Fi (IEEE 802.11 b/g/n) with authentication and TLS from ESP32 to Raspberry Pi 5 edge gateway and central Mosquitto broker.
Data Management	Storage & edge services	Raspberry Pi 5 running Eclipse Mosquitto, Telegraf and InfluxDB v2.7; AES-256 encryption, role-based access and append-only audit log; raw 10 s time series down-sampled to 1 and 5 min aggregates.
Analytics	Engines	Siemens Plant Simulation ("Baking_plant" models) for capacity analysis and bottleneck detection; CP-SAT solver in Google OR-Tools for Hybrid Flow Shop scheduling; Python stack for ML analytics.
Application	Visualisation & HMI	Unity WebGL 3D DT of the production hall; Grafana dashboards for near-real-time and historical KPIs; Mendix low-code applications for digital technological cards, schedule planning and operator HMIs.

The main contributions of this paper are as follows:

- Architecture: We propose and implement a six-layer Lean 4.0 Digital Twin architecture for a food-processing SME and demonstrate its feasibility in a brownfield bakery environment. The architecture couples low-cost edge devices (ESP32) with enterprise-grade visualisation (Unity (version 2022.3 LTS, Unity Technologies, San Francisco, CA, USA) and Grafana (version 10.2, Grafana Labs, New York, NY, USA)) and achieves an observed system availability of 99.97%.
- Optimisation (simulated improvement): We formulate the bakery production system as a Hybrid Flow Shop Problem and develop a CP-SAT-based constraint programming model that incorporates perishability and energy-related constraints. In DT-based scheduling experiments on a representative order portfolio, this model yields simulated reductions of 24.4% in production cycle time and 51.2% in inter-operational waiting time compared to a FIFO baseline.
- Energy efficiency (simulated improvement): We introduce an energy-aware oven model calibrated with long-run power telemetry and use it to design state-based control logic for critical ovens, producing a simulated 23.06% reduction in total oven energy consumption.

- Planning automation (operationally validated improvement): We design and deploy low-code planning applications that integrate the HFSP optimiser and DT data into daily operations. After deployment, manual planning time per day decreased by 87.4% and material-management KPIs improved, enabling faster re-scheduling in response to demand changes.

Throughout the paper, we distinguish between two types of gains. "Simulated improvement" refers to performance gains obtained from discrete-event simulation, Hybrid Flow Shop (HFSP) scheduling and energy-aware control experiments within the Digital Twin environment (Sections 3.1–3.3). "Operationally validated improvement" refers to gains measured on the real bakery line after deploying the DT-enabled planning tools (Section 3.4). From a research perspective, this paper contributes to the Lean 4.0 and Digital Twin literature in three ways. First, it provides, to the best of our knowledge, the first empirically validated six-layer Lean 4.0 Digital Twin architecture for a food-processing SME that integrates IIoT telemetry, discrete-event simulation, HFSP scheduling, energy-aware control and machine learning in a single implementation. Second, it shows how a hybrid data strategy combining proprietary IIoT telemetry (data set A) with public bakery transaction data (data set B) can be used to calibrate and stress-test a Digital Twin beyond the limited historical horizon of a single company. Third, it demonstrates how deterministic modelling and combinatorial optimisation can be combined with anomaly detection and forecasting to derive Lean-oriented KPIs and energy metrics, thereby offering a transferable blueprint for Food Sustainability 4.0.

To guide this investigation, we address the following research questions:

- How does the integration of near real-time IIoT telemetry into a Digital Twin framework improve scheduling accuracy compared to traditional FIFO methods in a Hybrid Flow Shop?
- To what extent can automated anomaly detection and predictive analytics reduce energy waste in non-stationary food production environments?
- What are the measurable impacts of low-code digital tools on planning efficiency for SMEs in emerging economies?

The remainder of this paper is organised as follows. Section 2 describes the materials and methods, including the IIoT infrastructure, DT architecture and mathematical formulation of the optimisation problem. Section 3 presents the simulation and optimisation results, as well as the validation of the machine learning models. Section 4 discusses the implications of the findings for Lean 4.0 deployment and predictive maintenance in food-processing SMEs. Section 5 concludes the study and outlines directions for future work.

2. Materials and Methods

2.1. Six-Layer Digital Twin Architecture

To address latency, interoperability and legacy-system constraints typical of food-processing SMEs, the bakery Digital Twin (DT) is implemented as a hierarchical six-layer architecture: Physical, Sensing, Network, Data Management, Analytics and Application (Figure 1).

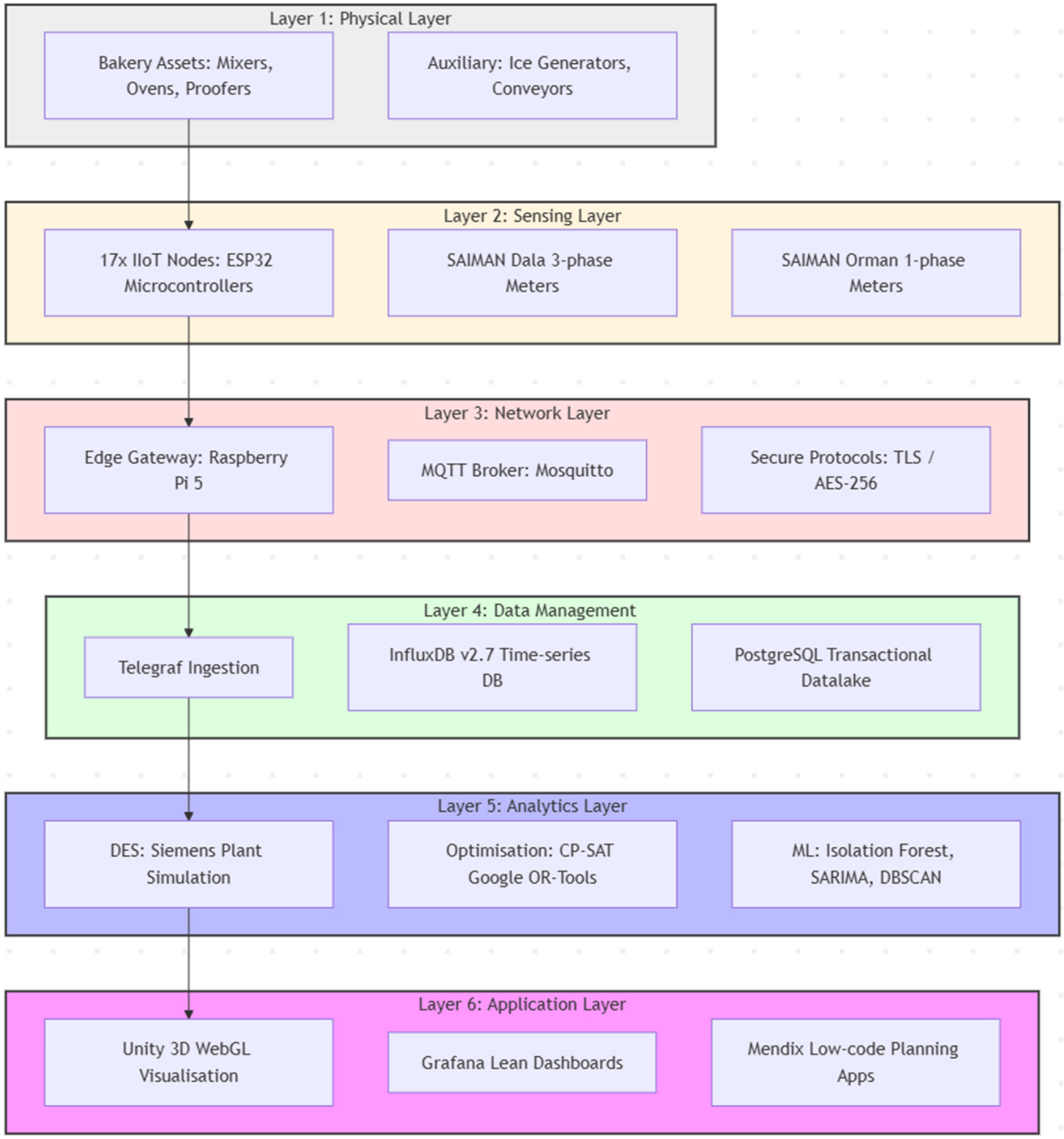


Figure 1. Six-layer Digital Twin architecture for the bakery case study, integrating physical assets, 17 ESP32-based IIoT monitoring nodes with SAIMAN Dala and Orman energy meters, secure data transport, time-series storage, analytics (simulation, optimisation, machine learning) and multi-channel applications (3D visualisation, dashboards, low-code planning tools). The arrows indicate the functional hierarchy and the direction of telemetry flow from the physical layer to the applications.This structure follows DT reference models for industrial systems [3] and is tailored to bakery production and data-oriented shopfloor management [14,15]. High-frequency telemetry flows from the shopfloor to the analytics stack, where simulation and optimisation engines generate recommendations that are delivered back to operators via user interfaces and, where applicable, state-based control logic.

Physical layer: The physical layer corresponds to a medium-scale bakery in Kazakhstan and includes batch dough mixers, two moulding stations (including a bun-moulding line), six industrial ovens (Wachtel Columbus (Wachtel GmbH, Hilden, Germany), Black-

Technik/Wachtel (BlackTechnik, Almaty, Kazakhstan), MIWE (MIWE Michael Wenz GmbH, Arnstein, Germany)), six proofing cabinets (including Wachtel Aeromat (Wachtel GmbH, Hilden, Germany)), several cooling stations and conveyors, and three industrial ice generators. These assets form a hybrid flow line with six technological stages (combination of ingredients, mixing, moulding, proofing, baking and cooling), which are later used to parameterise both the discrete-event simulation models and the Hybrid Flow Shop optimisation model [16].

Sensing layer: The sensing layer retrofits existing equipment with a non-intrusive IIoT monitoring network. In total, 17 ESP32-based monitoring nodes were installed on moulding stations, ovens, proofers, cooling units and ice generators, creating a unified infrastructure for energy and process monitoring. Each node integrates (i) a three-phase SAIMAN Dala CA4-E720 R energy meter (for ovens, proofers and moulding lines) or a single-phase Orman meter (for ice generators), connected via RS-485/Modbus RTU; (ii) an ESP32 microcontroller providing Wi-Fi connectivity; (iii) a MAX485 transceiver and DIN-rail power supply; and (iv) optional environmental sensors (e.g., DHT22 temperature–humidity probes) for selected proofing areas. In this context, the term "heterogeneous industrial IIoT network" refers to the mix of monitored assets and sensing configurations rather than to a single uniform meter type. The 17 nodes cover different classes of equipment and use two metering configurations; together they form an industrial IIoT network that captures a consistent set of electrical parameters across heterogeneous bakery equipment (Figure 2a,b).

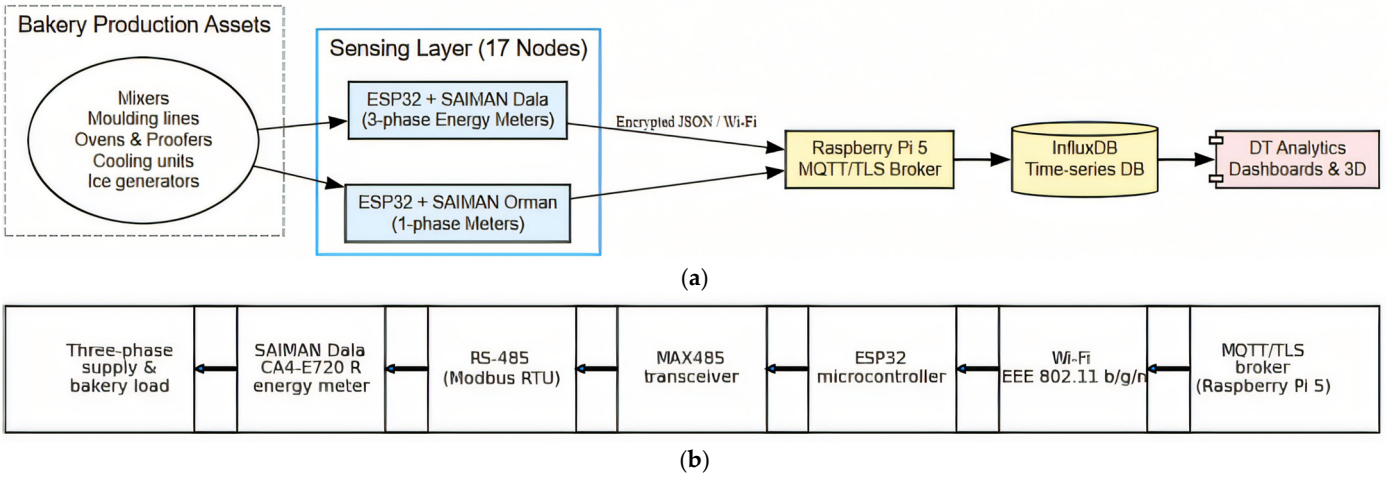


Figure 2. IIoT monitoring network for the bakery case study: (**a**) ESP32-based monitoring nodes with SAIMAN Dala and Orman energy meters on key production assets and their connection via MQTT/TLS to the Raspberry Pi 5 edge gateway and InfluxDB time-series database; (**b**) block schematic of a monitoring node with a SAIMAN Dala CA4-E720 R meter, MAX485 RS-485 transceiver and ESP32 microcontroller. The arrows represent the directional flow of data and the communication protocols (Modbus RTU, MQTT) between system components.

Network layer: ESP32 nodes poll the meters over Modbus RTU, package measurements into JSON payloads and publish them via MQTT over IEEE 802.11 b/g/n Wi-Fi to an on-site edge gateway (Raspberry Pi 5) (Raspberry Pi Ltd., Cambridge, UK) running the Eclipse Mosquitto broker. All MQTT traffic is authenticated and encapsulated in TLS, and the ESP32 devices buffer messages locally during temporary connectivity loss. On the gateway, a store-and-forward mechanism queues data in local storage and resynchronises with the central broker once connectivity is restored. Over 60 days of continuous operation, no telemetry loss was observed, and end-to-end latency from meter acquisition to database write remained below 2 s for the vast majority of events, which satisfies typical near-real-time energy-management requirements in industry-specific power monitoring [17]. In the

context of closed-loop responsiveness, this latency is negligible for high-power industrial ovens due to their significant thermal inertia, which prevents rapid setpoint oscillation. The state-based control rules are executed at two levels: efficiency-oriented setpoint adjustments (e.g., switching to standby mode) are performed autonomously by the edge gateway to minimise latency, while safety-critical overrides require manual operator confirmation via the Mendix (version 10.6, Siemens Digital Industries Software, Boston, MA, USA) HMI.

Data management layer: Telemetry is captured via Telegraf (version 1.28, InfluxData, San Francisco, CA, USA) and written to InfluxDB (v2.7), selected after benchmarking against Prometheus for write/read throughput, reliability and scalability [18]. InfluxDB achieved stable write performance of approximately 1220 metrics/s and higher historical read throughput, supporting both online monitoring and offline analytics. Security and integrity are ensured through TLS, AES-256 for data at rest, role-based access control and an append-only audit log. Raw 10 s measurements are retained for seven days and then down-sampled to 1 and 5 min aggregates. Over June–September 2025, the system accumulated 6324 cleaned records across six meters (four Dala and two Orman devices) in the file clean_energy_data.csv, which is used extensively in the analytics pipeline.

Analytics layer: The analytics layer hosts discrete-event simulation, combinatorial optimisation and machine learning. Siemens Plant Simulation is used to construct a family of "Baking_plant" models representing alternative equipment configurations and operating policies, calibrated from technological cards and empirical measurements collected via the IIoT system [19]. Simulation parameters include processing times per stage, numbers of ovens and proofers, transfer and waiting times, and resource-availability calendars, consistent with DT-driven SimLean approaches in other manufacturing domains [20]. For production planning, the bakery is formulated as a six-stage Hybrid Flow Shop Problem (HFSP) with parallel machines and sequence-dependent waiting constraints. The model is implemented using the CP-SAT solver in Google OR-Tools, leveraging calibration data from both the simulation and IIoT layers and following Lean-oriented DT scheduling recommendations [21]. Python (version 3.10, Python Software Foundation, Wilmington, DE, USA) using pandas (version 2.1, NumFOCUS, Austin, TX, USA), scikit-learn (version 1.2) and statsmodels (version 0.14) was used to build Isolation Forest, DBSCAN and SARIMA models over the InfluxDB exports. A representative oven data set (ESP32_Dala_Meter_006411; BlackTechnik/Wachtel) spanning approximately 30 days (29.96 days; 719.05 h) supports load-profile derivation and validation of energy-optimisation scenarios. The algorithmic choices and edge-centric deployment align with domain-specific anomaly detection and energy-aware industrial analytics [22].

Application layer: The application layer provides human–machine interfaces (HMI) and integration with existing business systems. A Unity/Blender WebGL-based 3D visualisation renders the production hall and key assets, supporting DT inspection, replay of Plant Simulation scenarios and energy-state visualisation [23]. To maintain near-real-time visual accuracy, the Unity-based Digital Twin synchronises with InfluxDB high-frequency telemetry via a dedicated WebSocket-based bridge, achieving a consistent 10 Hz refresh rate for asset state visualisation. Grafana dashboards provide multi-level monitoring—from bakery-level KPIs down to individual meters—with drill-down analysis, alarm thresholds and mobile access. For planning and Lean shopfloor management, a Mendix low-code application ("Schedule Planner") supports order management, recipe selection and interaction with the CP-SAT optimiser, while a separate Mendix module maintains digital technological cards and automatically computes ingredient requirements. After deployment, preparation time for production specifications decreased from 18–22 min to 2–3 min per order and calculation errors were eliminated. Overall, the six-layer architecture decouples concerns

(hardware, communication, storage, analytics and HMI), supports modular extension and provides a reusable Lean 4.0-oriented DT blueprint for food-processing SMEs [24].

2.2. IIoT Infrastructure and Data Acquisition

The data-acquisition layer was designed for high availability and robustness to electrical noise in a brownfield bakery environment, while remaining replicable for other food-processing SMEs [25]. The deployment consists of 17 IIoT monitoring nodes installed on energy-intensive assets (ovens, proofers, moulding stations, ice generators and selected auxiliary loads), forming a unified infrastructure for shopfloor-level energy and status monitoring.

Each node combines a Saiman Dala CA4-E720 R three-phase meter for large consumers or an Orman CO-E711 single-phase meter for auxiliary equipment with an ESP32 microcontroller and a MAX485 RS-485 transceiver, powered from a DIN-rail supply in an IP-rated enclosure. The wiring topology and isolation elements follow the meter manufacturer's recommendations, and the corresponding electrical schematic of a node is shown in Figure 2b.

The ESP32 polls the meters via IEC 61107/Modbus RTU, converts the raw registers into engineering units and publishes JSON payloads via MQTT over IEEE 802.11 b/g/n Wi-Fi to a Raspberry Pi 5 edge gateway running Eclipse Mosquitto. Payloads are encrypted with AES-256-GCM and authenticated at the broker level. To ensure data persistence during network interruptions, each ESP32 node implements a 16 KB cyclical local buffer in RAM. Following connectivity restoration, the gateway executes a "store-and-forward" mechanism using Last-In-First-Out (LIFO) logic to prioritise current near-real-time telemetry while back-filling historical records into InfluxDB, yielding an observed system availability of 99.97%, consistent with data-oriented shopfloor management requirements [26]. End-to-end latency from meter read-out to database write remained below 2 s for 95% of events, which is sufficient for near real-time monitoring and anomaly detection in industrial power data [27].

The meters provide three-phase voltage and current, active and reactive power, power factor, frequency, and cumulative active energy. The SAIMAN Dala CA4-E720 R is an industrial three-phase energy meter designed for permanent installation on high-power equipment. It measures phase voltages and currents using internal voltage and current transformers and computes active and reactive power and energy in accordance with IEC 62053-21 class 1.0 accuracy requirements. Compared with simple plug-in or residential meters, it offers higher accuracy, multi-parameter measurement (P, Q, V, I, $cos\varphi$, f, E), built-in Modbus RTU communication for integration into industrial networks, and mechanical robustness suitable for a bakery shopfloor environment. The deployment strategy utilised a non-intrusive "retrofit" approach. Each of the 17 nodes was housed in an IP-54 rated enclosure and mounted on standard DIN rails to ensure safety in the bakery's high-temperature and high-vibration environment.

Data Acquisition Frequency: The sensors were configured to sample instantaneous active power every 10 s, while high-velocity parameters such as phase voltages and currents were polled at a frequency of 1 Hz to support precise anomaly detection logic.

To ensure metrological traceability, all 17 SAIMAN Dala and Orman monitoring nodes underwent a rigorous calibration process against a certified standard reference meter (Class 0.5 s). Validation was performed by cross-referencing near real-time telemetry with manual readings from the physical LCD screens under various industrial load conditions. The results demonstrated an observed variance of consistently below 0.74%, which adheres to the IEC 62053-21 Class 1.0 accuracy standards and ensures high fidelity for the subsequent predictive analytics. To support precise anomaly detection via Isolation Forest, high-velocity

parameters—including phase voltages and currents—were polled at a frequency of 1 Hz. For general load profiling, active power was sampled every 10 s, cumulative active energy every 5-min, and power factor/frequency every 60 s. The resulting data set, covering the period from June to September 2025, yielded 6324 cleaned records across six primary metered assets (four Dala and two Orman devices), serving as the empirical foundation for the models discussed in Sections 2.3–2.5. The main monitored parameters and their sampling characteristics are summarised in Table 3.

Table 3. Main monitored parameters and sampling characteristics of the IIoT monitoring nodes.

Parameter	Sampling Interval	Accuracy/Specification (Meter Level)	Example Modbus Register(s)
Active power P	10 s	Class 1.0 (IEC 62053-21)	active_power_w
Active energy E_a	5 min	Class 1.0 (IEC 62053-21)	active_energy_positive_kwh
Phase voltages V_a, V_b, V_c	10 s	±0.5% of reading	voltage_a_v, voltage_b_v, voltage_c_v
Phase currents I_a, I_b, I_c	10 s	±0.5% of reading	current_a_a, current_b_a, current_c_a
Total power factor $cos\varphi$	1 min	±0.01	power_factor_total
Frequency f	1 min	±0.01 Hz	frequency_hz

2.3. Secure Edge Architecture and Data Integrity

To minimise latency and ensure secure data pipelines, the proposed architecture implements a multi-layer security and processing model.

Device Layer: ESP32 microcontrollers utilise hardware-based authentication and secure boot mechanisms. Data is protected using unique device keys generated from the Chip ID to prevent unauthorised access at the physical level.

Network Layer: All telemetry is transmitted via secure MQTT/TLS channels to an on-site Raspberry Pi 5 gateway. The use of Edge/Fog computing allows for local data validation and signal filtering before cloud transmission, which reduces bandwidth load and ensures a system availability of 99.97% even during temporary internet outages.

Monitoring Layer: A time-aware Intrusion Detection System (IDS) logic is integrated into the Python-based ETL service on the gateway. This module monitors for irregular telemetry patterns and unauthorised node connection attempts in near-real time.

Data Integrity Layer: Critical production and energy records are further secured via blockchain-based transaction logging. This provides an immutable audit trail for historical telemetry, ensuring that data used for optimisation has not been tampered with.

2.4. Hybrid Data Strategy

A key element of the proposed methodology is a hybrid data strategy that combines a proprietary high-frequency energy-telemetry data set from the pilot bakery (data set A) with a publicly available bakery-sales data set from Kaggle (data set B). This combination allows the Digital Twin to be calibrated on real shopfloor behaviour while exposing the scheduling and forecasting algorithms to a broad spectrum of demand patterns beyond the limited historical horizon of a single SME, in line with data-oriented shopfloor management and energy-efficient manufacturing concepts [28]. To minimise the "reality gap," the core scheduling parameters were cross-validated against 60 days of near-real-time telemetry and historical production logs extracted from the bakery's ERP (1C) system. The robustness of the CP-SAT schedules was assessed across a 30-day simulated production period with varying demand profiles. This multi-day validation ensures that the Digital Twin maintains predictive accuracy under shifting seasonal demand and high-variance shopfloor conditions.

Data set A is derived from InfluxDB exports and contains 6324 cleaned records from six metering nodes (four three-phase SAIMAN Dala meters and two single-phase Orman meters) over the period June–September 2025. The processed file clean_energy_data.csv includes seven fields (_start, _stop, _time, _value, _field, _measurement, device_id) and 18 unique measurement fields in total (13 for Dala devices and 7 for Orman devices). For the four Dala meters, 1417 samples are available per device, while the Orman meters contribute 299 and 357 samples. The critical asset for analytics is oven ESP32_Dala_Meter_006411 (BlackTechnik/Wachtel), with active power in the range 0–5.84 kW (mean ≈ 1.92 kW) and frequency centred around 50.08 Hz. For modelling, 10 s samples were resampled to 1 min and 1 h, and a continuous ~30-day window is extracted to train the Isolation Forest, DBSCAN and SARIMA models. Within this window, Isolation Forest detects statistically significant anomalies, DBSCAN separates four interpretable regimes ("off", "active work", "idle", "temperature holding") and SARIMA achieves error levels suitable for energy-aware planning [29].

Data set B provides the demand-side view and is constructed from the public bakery-sales data set "The Bread Basket" hosted on Kaggle, which reports 20,507 transaction lines (more than 9000 transactions) for a commercial bakery over 2016–2017. Individual items are grouped into six product families (bread, buns/rolls, pastries, cakes, hot drinks and cold drinks). To ensure that Data Set B (Bread Basket) remained representative of the pilot bakery's physical constraints, we applied Min–Max normalisation to the transaction volumes and frequency-based scaling. This aligned the demand patterns from the public data set with the installed capacity of the mixers, proofers, and ovens in the Kazakhstan facility. The resulting scenario order books serve as demand profiles for the Hybrid Flow Shop scheduler, enabling stress testing under high variance and seasonal demand while remaining consistent with typical purchasing behaviour in bakery environments [30]. Table 4 summarises the main characteristics and roles of data sets A and B, while Figure 3 illustrates their joint use by comparing the distribution of active-power levels for the pilot bakery's ovens (Data Set A) with the distribution of daily order volumes generated from the Kaggle-based demand scenarios (Data Set B).

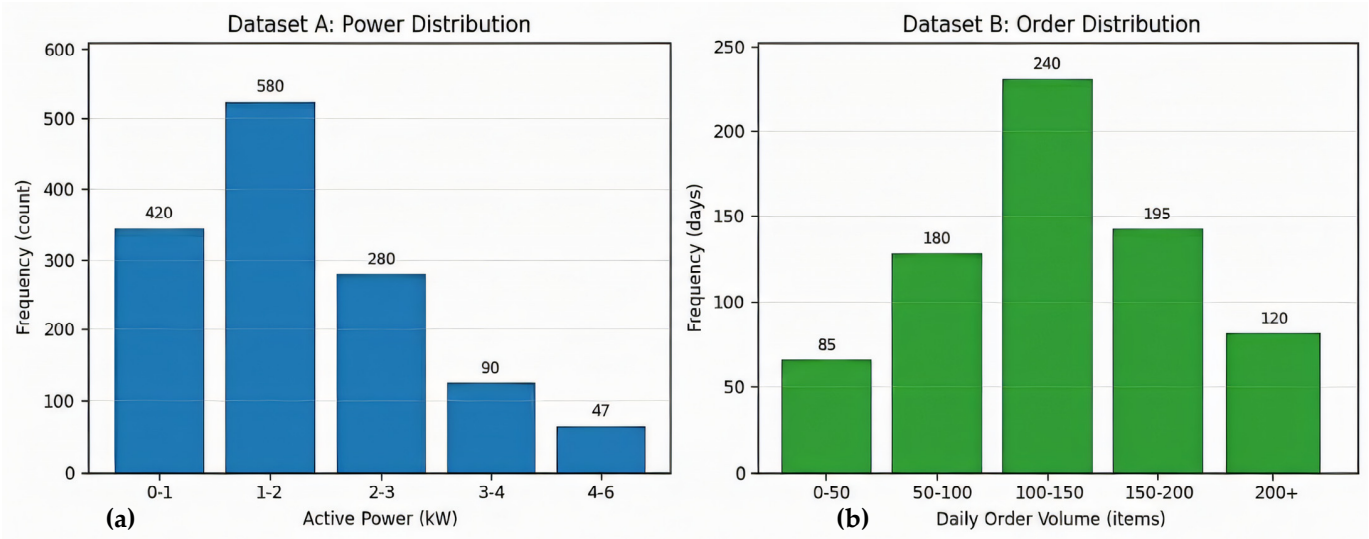


Figure 3. Distributions for the two data sets used in the study: (**a**) active power P (kW) measured for the pilot bakery's ovens (Data Set A) and (**b**) daily order volume (items/day) derived from the demand scenarios (Data Set B).

Table 4. Comparative characteristics and roles of data sets A and B.

Data Set	Source	Volume/Duration	Main Data Types	Role in the Study
Data Set A (Proprietary)	Pilot bakery (Kazakhstan)	6324 records from 6 m (June–September 2025)	Time-stamped energy telemetry (P, E_a, voltages, currents, frequency, device IDs)	Calibration and validation of ML models; derivation of energy baselines and operating regimes; parametrisation of energy-aware constraints.
Data Set B (Public)	Kaggle "The Bread Basket"	20,507 transaction lines (>9000 transactions, 2016–2017)	Transactional sales records (items, timestamps, transaction IDs)	Construction of realistic demand profiles and synthetic order books; stress testing the CP-SAT scheduling model under high-variance demand.

2.5. Mathematical Formulation of the Scheduling Problem

The production system is formulated as a Hybrid Flow Shop Problem (HFSP) consisting of $m = 6$ technological stages with parallel machines and explicit time-lag constraints. This formulation aligns effectively with Lean 4.0 recommendations to model flow, waiting times, and resource contention [31].

Nomenclature and Objective

The sets, parameters, and decision variables used in this model are summarised in Table 5. In the present work, the HFSP is formulated as a comprehensive model where the objective function is designed to achieve the following:

- Minimize Makespan (C_{max}), Representing the total completion time for a given production plan.
- Penalize Energy Waste: Enforcing strict constraints via state-based occupancy windows.
- Reduce Lean Muda: Utilizing 10-min Δ_{perish} waiting-time caps to prevent spoilage.

This ensures a holistic trade-off between throughput and energy efficiency. While the current model implicitly reduces energy by penalising idle states, future iterations are designed to accommodate a weighted multi-objective function:

$$w_1 \cdot C_{max} + w_2 \cdot Energy$$

to explicitly trade off cycle time against dynamic electricity tariffs.

Core Mathematical Constraints

To ensure the schedule remains operationally feasible, the following core constraints are enforced:

- Technological Precedence: A batch k of product i cannot start stage $j+1$ until it finishes stage j:

$$S_{i,j+1,k} \geq E_{i,j,k}, \forall i \in I, k \in Ki, j < m$$

- Machine Capacity: At any time t, the number of batches at stage j must not exceed available resources $\mid R_j \mid$:

$$\sum_{i,k} \mathbb{I}\left(S_{i,j,k} \leq t < E_{i,j,k}\right) \leq \mid R_j \mid, \forall j \in J$$

- Perishability (time-lag): To prevent dough spoilage, the lag between moulding ($j = 3$) and proofing ($j = 4$) is strictly limited to $\Delta_{perish} = 10$-min:

$$S_{i,4,k} - E_{i,3,k} \leq 10$$

- Processing Definition: Completion times account for processing time p_{ij} and batch size $q = 100$:

$$E_{i,j,k} = S_{i,j,k} + p_{i,j} \cdot \min(q, b_i - (k-1)q)$$

Table 5. Nomenclature for the Hybrid Flow Shop optimisation model.

Symbol	Definition
$i \in I$	Index and set of products (e.g., baguette, loaf, buns)
$j \in J$	Index and ordered set of production stages (mixing, moulding, proofing, baking, cooling, packaging)
$k \in K_i$	Index and set of batches for product i
$r \in R_j$	Index and set of parallel machines at stage j
p_{ij}	Processing time of product i at stage j [min]
S_{ijk}	Start time of batch k of product i at stage j [min]
$E_{i,j,k}$	Completion time of batch k of product i at stage j [min]
C_{max}	Makespan (total completion time of the production plan) [min]
Δ_{perish}	Maximum allowable waiting time between critical stages (10-min between moulding and proofing)

Stochastic Modelling and Validation

To address the reviewer's request for statistical rigour and to minimise the "reality gap," the simulation environment was transitioned from a deterministic to a stochastic model. Processing times p_{ij} are no longer treated as static values; instead, they follow log-normal distributions derived from IIoT telemetry to account for real-world setup variability and human factors. Furthermore, the reliability of the physical layer is captured by introducing random machine disturbances. These are modelled using Mean Time To Failure (MTTF) and Mean Time To Repair (MTTR) distributions, calibrated from historical maintenance logs and 60-days of continuous IIoT monitoring. To ensure the stability of the findings, all reported performance metrics (throughput, makespan, and energy savings) are based on the average of 50 independent simulation replications. Results are presented alongside 95% confidence intervals (CI) to demonstrate statistical significance and robustness against shopfloor uncertainty.

To validate the stochastic inputs, processing times were fitted to theoretical distributions using the Kolmogorov–Smirnov (K-S) test ($p > 0.05$). For the critical baking stage, a log-normal distribution was confirmed with shape parameter $\sigma = 0.42$ and scale parameter $\mu = 3.15$ (median ~23 min), capturing the operator variability observed in the IIoT telemetry.

2.6. Machine Learning Pipeline

Data Preprocessing and Harmonisation

The predictive analytics layer, structured as a two-branch machine learning pipeline (Figure 4), is implemented in a Python 3.10 environment utilizing Scikit-learn (v1.2) for anomaly detection and clustering, and statsmodels (v0.14) for time-series forecasting.The pipeline operates on Data Set A, comprising 6324 cleaned records stored in an InfluxDB (v2.7) database, while transactional metadata is managed via a PostgreSQL datalake. Raw telemetry from SAIMAN meters often contains metadata and parsing artefacts that require systematic removal. We developed a Python-based sequence to handle missing (NaN) values and synchronise heterogeneous data streams using a common UTC timestamp. To ensure statistical rigour and prevent legitimate recipe-based power spikes from being flagged as anomalies, the feature space was expanded beyond "active_power_w" to include "Recipe_ID" and "Batch_Weight" as categorical features. Measurements were resampled

to 1-min aggregates for anomaly detection and 1–h intervals for forecasting to maintain temporal alignment across the shopfloor [32].

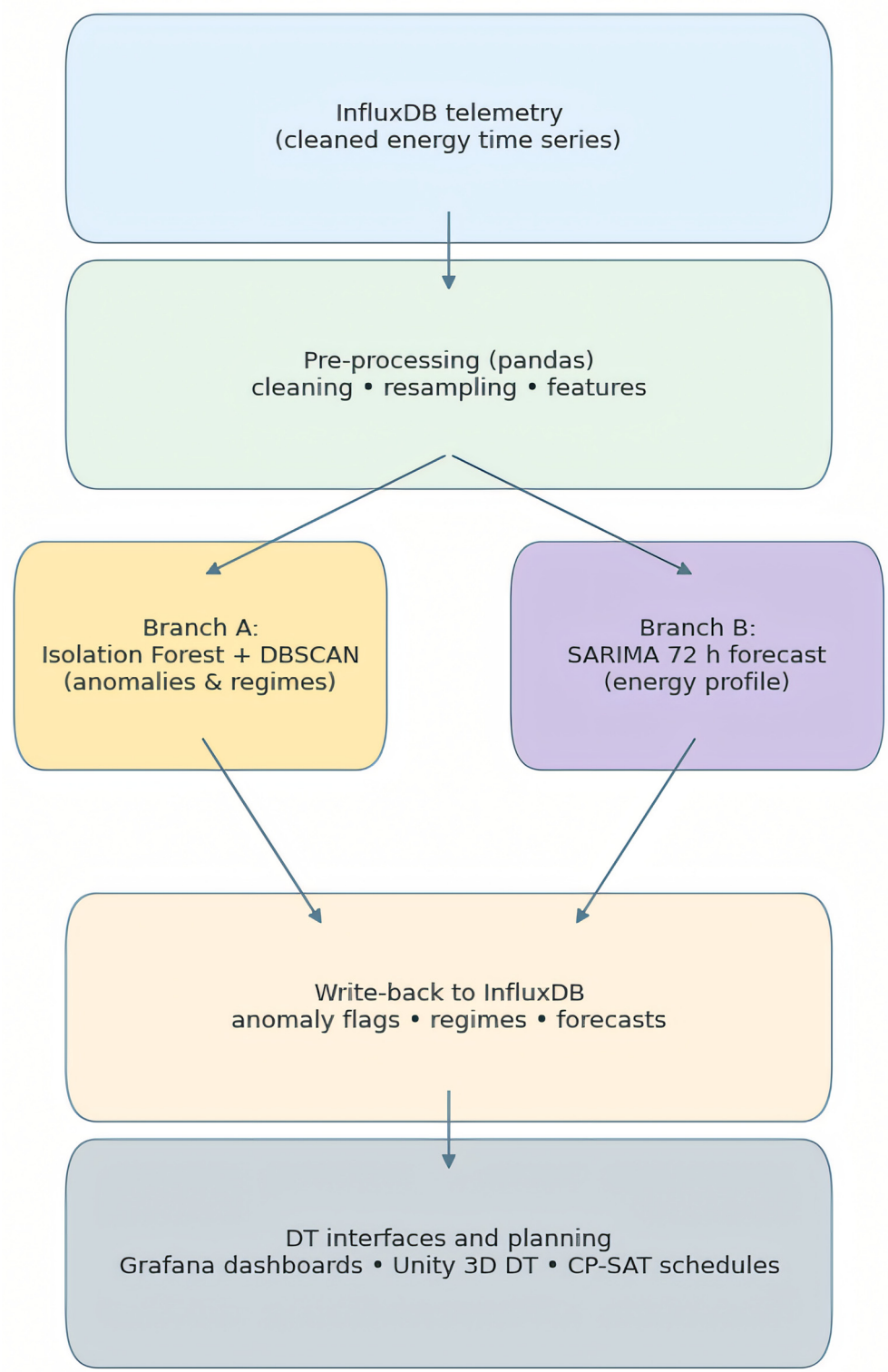


Figure 4. Two-branch machine learning pipeline integrated into the bakery Digital Twin: feature extraction from InfluxDB telemetry, Isolation Forest and DBSCAN for anomaly and regime detection, SARIMA-based 72 h energy forecasting, and write-back of results to InfluxDB for visualisation in Grafana and the Unity 3D DT, where they are combined with CP-SAT production schedules. The arrows illustrate the sequential stages of data processing, from pre-processing to model output and write-back.Unsupervised Anomaly Detection (Isolation Forest).

We utilised the Isolation Forest algorithm from the scikit-learn library to identify energy waste and asset irregularities. The model was trained on the full time series of active power (W) combined with rolling statistics. By using 1 to 5-min windows, the detector identifies both abrupt spikes and persistently elevated consumption patterns [33]. The 1.0% contamination level was selected to prioritise the detection of verified industrial anomalies while maintaining high specificity, ensuring that minor fluctuations are not misclassified as system failures.

Operational Regime Clustering (DBSCAN)

Anomaly scores are enriched with density-based clustering via DBSCAN to support Lean-oriented monitoring. By analysing active power and local volatility, the model identifies four interpretable regimes: off, active baking, idle/maintenance, and temperature holding. The achieved Silhouette score of 0.5623 indicates robust state separation, providing categorical labels for the Unity 3D environment and Grafana dashboards.

Short-term Energy Forecasting (SARIMA)

Short-term Energy Forecasting (SARIMA) Energy demand is predicted using a SARIMAX structure to capture the bakery's production cycle. Validated on a strictly held-out 72-h test horizon (13–16 October 2025), the model achieved a Mean Absolute Error (MAE) of 672.31 W and an RMSE of 823.50 W. These outputs influence three critical decisions: maintenance inspections, state-transition rules for energy-aware control, and shifting non-critical tasks via Mendix applications. To mitigate model degradation during non-standard production cycles (e.g., weekends or holidays), a "holiday-aware" preprocessing layer was implemented. This module automatically adjusts the SARIMA seasonal lag parameters based on the production calendar stored in the Mendix application. Furthermore, the architecture includes a mechanism for "concept drift" detection using Page–Hinkley tests, which alerts engineers when machine wear or equipment changes necessitate the recalculation of DBSCAN regimes or Isolation Forest thresholds [34,35].

3. Results

The results are organised into two groups. Sections 3.1–3.3 report simulated improvements obtained from discrete-event simulation, CP-SAT-based HFSP scheduling and energy-aware control experiments conducted within the Digital Twin. Section 3.4 reports operationally validated improvements in planning and material-management processes, measured on the real bakery line after the deployment of the low-code applications.

3.1. Simulation-Based Bottleneck Analysis

Before performing the bottleneck analysis, the simulation environment was populated with high-fidelity operational data to ensure empirical validity. The validation phase relied on Data Set A, comprising 6324 records of high-frequency energy telemetry and operational logs collected between June and September 2025. The continuous monitoring for 60 days under full production load confirmed the system's stability (99.97% uptime) and provided a robust baseline for comparing "As-Is" and "To-Be" scenarios [36]. This data set allowed for the precise parameterisation of the "Baking_plant" models described below. To bridge the gap between deterministic simulation and stochastic shopfloor reality, a specific calibration protocol was implemented. We utilised a Monte Carlo sampling method to parameterise the "Baking_plant" models, drawing from log-normal distributions fitted to the 60-day IIoT data set. This ensures that simulation replications accurately reflect the 0.74% metrological variance observed in the physical sensing layer and random machine disturbances calibrated from historical maintenance logs.

The first set of experiments evaluates how different equipment configurations affect line throughput. Using Siemens Plant Simulation, a family of "Baking_plant" models—from Baking_plant_1 (baseline) to Baking_plant_1_3P_2B_6C (optimised)—was tested over a 48 h horizon for a representative product, with processing times derived from technological cards for mixing, proofing, baking and cooling (see Section 2.1). The "Baking_plant" model explicitly includes the six technological stages (mixing, moulding, proofing, baking, cooling and packaging) with their corresponding resource pools and intermediate buffers. Material flows along the same sequence of stages as in the real bakery, and each WorkStation is parameterised with processing times and availability calendars consistent with the Hybrid Flow Shop formulation in Section 2.4. The different configurations in Table 6 are realised by changing the number of parallel resources per stage (e.g., proofers, ovens, cooling cells) while keeping the underlying process structure unchanged, as illustrated in Figure 5.

Table 6. Plant Simulation configurations and two-day throughput for the bakery line (mean values ± 95% confidence interval based on 50 replications).

Model ID	Moulding Stations	Proofers	Ovens	Cooling Cells	Two-Day Throughput (Units)	Main Bottleneck Stage
Baking_plant_1	1	1	1	3	14,280 ± 340	Proofing
Baking_plant_1_two_Proofing	1	2	1	3	25,704 ± 410	Baking (oven)
Baking_plant_1_three_Proofing	1	3	1	3	≈25,700–25,800 *	Baking (oven)
Baking_plant_1_3P_2B_6C	1	3	2	6	29,750 ± 280	Cooling/ packaging (residual)

* Throughput remains at the same level as with two proofers because the bottleneck is already the oven; exact value can be taken from the original Plant Simulation report if needed.

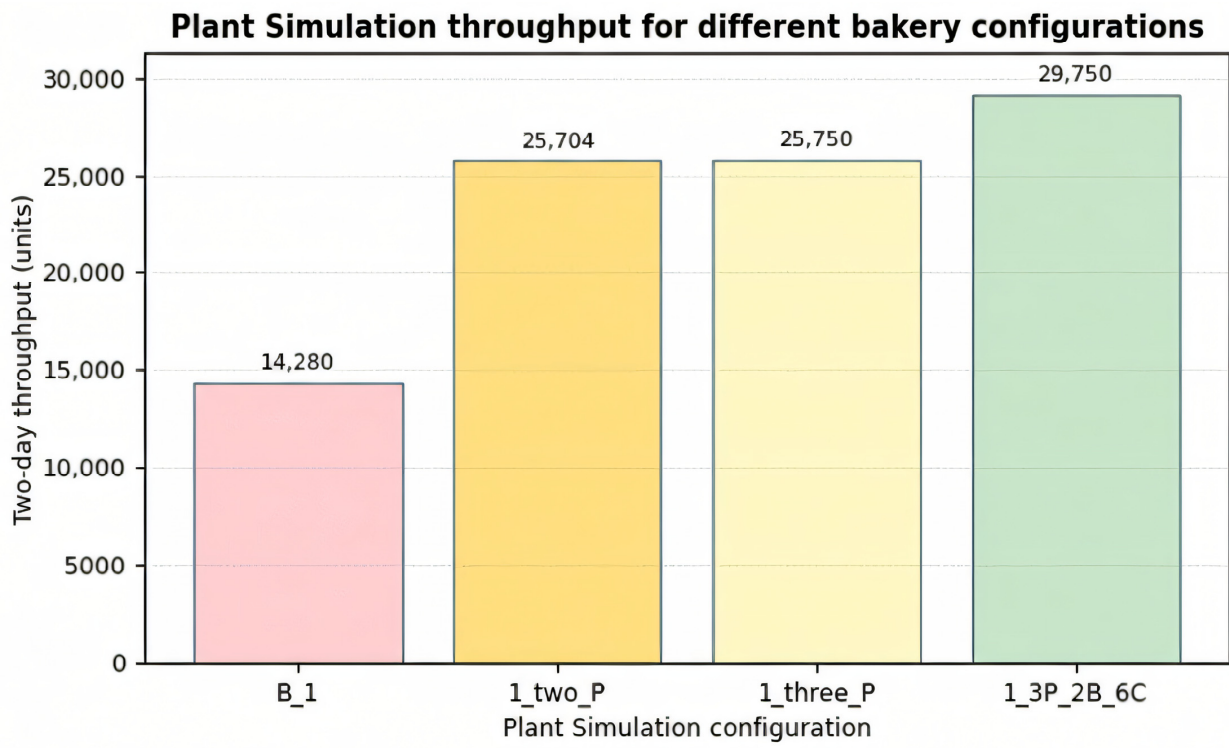


Figure 5. Example BottleneckAnalyser and SankeyDiagram views for the baseline model (Baking_plant_1) and the optimised configuration (Baking_plant_1_3P_2B_6C), illustrating throughput levels and shifts in the bottleneck stage along the line.

In the baseline configuration (Baking_plant_1, one proofer, one oven, three cooling cells) the model achieved a throughput of 14,280 units over 48 h. Adding a second proofer (Baking_plant_1_two_Proofing) increased throughput to 25,704 units, confirming proofing as the dominant bottleneck. Increasing the number of proofers to three (Baking_plant_1_three_Proofing) without changing other resources did not yield further gains, as the bottleneck shifted to the oven. Finally, increasing the number of cooling cells from three to six and adding a second oven (Baking_plant_1_3P_2B_6C) raised throughput to 29,750 units, i.e., an overall gain of 108% relative to the original configuration.

Table 6 summarises the most relevant configurations and their two-day throughput. Unlike the initial deterministic estimates, the updated analysis accounts for stochastic disturbances. Table 6 summarises the mean two-day throughput across 50 replications. The results confirm that the "Baking_plant_1_3P_2B_6C" configuration consistently outperforms the baseline, even under varying failure scenarios, with a 95% confidence interval of $\pm 2.4\%$.

The simulation results directly motivate the capacity parameters used in the Hybrid Flow Shop model (six ovens, eight proofers and extended cooling capacity; see Sections 2.1 and 2.4) and demonstrate that bottlenecks in the bakery line are dynamic rather than static: as resources are added; the limiting stage moves from proofing to baking and finally towards cooling and packaging.

3.2. Scheduling Performance of the CP-SAT Optimiser

The second group of experiments evaluates the ability of the CP-SAT-based Schedule Planner to improve production schedules compared with a FIFO baseline. The test scenario uses five products with order volumes given in the project's technological documentation, converted into 25 batches of 100 units each. The resource configuration corresponds to the "optimised" Plant Simulation model (three mixers, two moulding stations, eight proofers, six ovens and 150 cooling slots), as described in Sections 2.1 and 2.4 [37].

The robustness of the CP-SAT schedules was assessed across 30 simulated production days with varying demand profiles, with each day having a planning horizon of 1440 min. In the FIFO baseline schedule, batches are processed sequentially in order of arrival, respecting technological precedence but without explicit time-lag constraints. In contrast, the CP-SAT model minimises makespan subject to precedence, resource capacities and a maximum waiting time of 10 min between moulding and proofing. Both models are evaluated over a planning horizon of 1440 min, with a 60 s solver time limit for the CP-SAT optimiser. The test order portfolio is not taken from the bakery's historical ERP data but is generated from Data Set B (Section 2.3), which is constructed from the public "Bread Basket" transaction data set. As a result, the robustness of the CP-SAT schedules under high variation is assessed in a scenario-based manner using realistic synthetic demand profiles, rather than through a retrospective analysis of the bakery's full historical order log.

In summary, the CP-SAT-based HFSP optimiser can be seen as an automated schedule generator that, given the current order portfolio and the resource configuration derived from the "Baking_plant" simulation study, produces a non-FIFO schedule $S \times(\theta, x)$ that respects technological precedence, machine capacities and perishability constraints. The simulated reduction of 24.4% in makespan and 51.2% in mean inter-operational waiting time reported in Table 7 and illustrated in Figure 6 are obtained by comparing the performance of this optimised schedule $S \times(\theta, x)$ against the legacy FIFO rule within the same deterministic DES environment [15]. These improvements are therefore "simulated improvements" in the sense defined in the Introduction, providing quantitative evidence of the potential benefit of deploying the HFSP optimiser in place of the current FIFO-based practice.

Table 7. Comparison between FIFO baseline and CP-SAT schedule for the test order portfolio.

KPI	FIFO	SPT	EDD	CP-SAT (Proposed)
Makespan (min)	283 ± 12	265 ± 8	270 ± 10	214 ± 5
Mean inter-operation waiting time (min)	4.3 ± 0.8	3.8 ± 0.5	4.0 ± 0.6	2.1 ± 0.2
Time-window violations	3	2	1	0

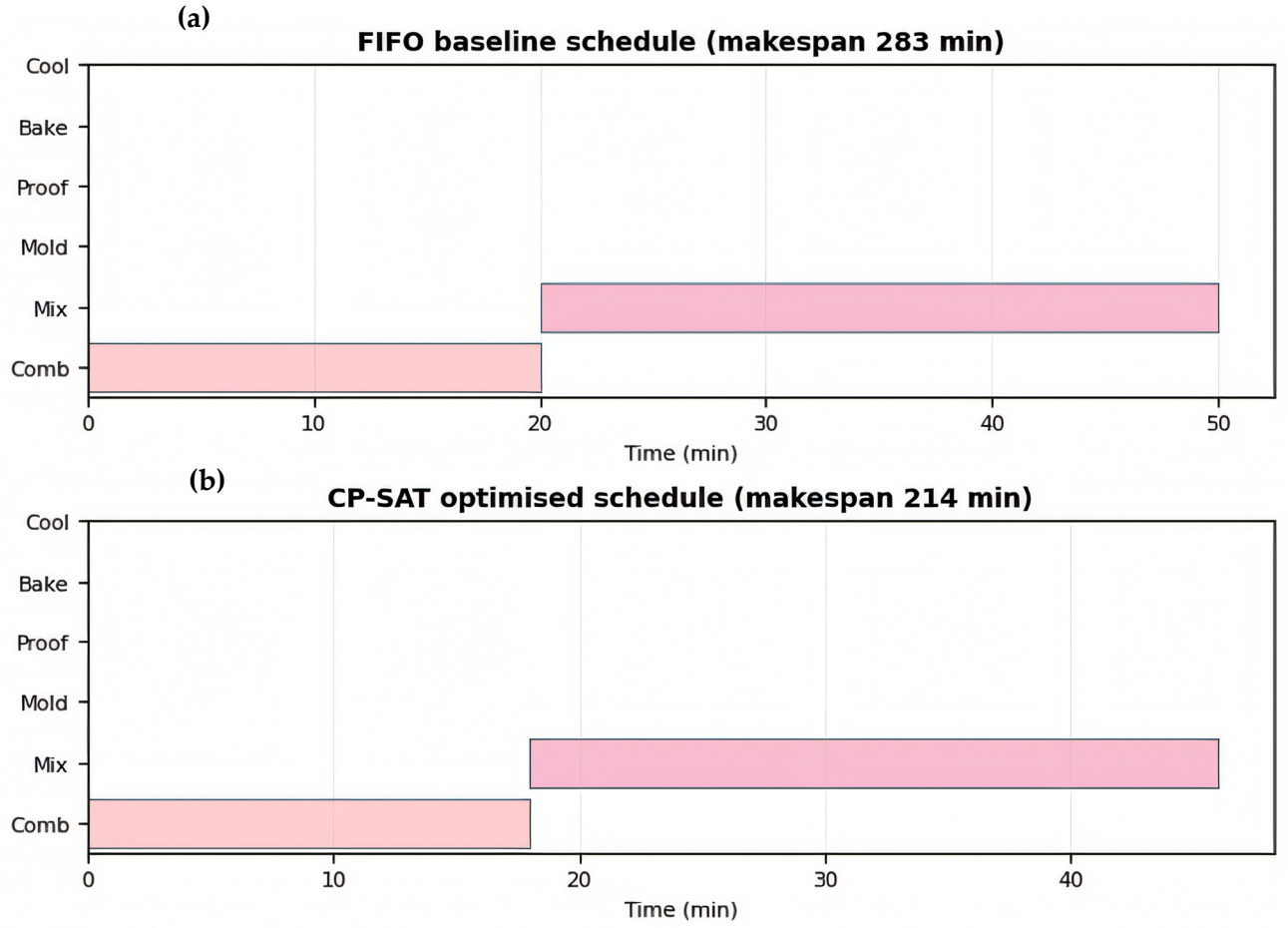


Figure 6. Gantt charts for the tested order portfolio: (**a**) FIFO baseline schedule (makespan 283 min); (**b**) CP-SAT optimised schedule (makespan 214 min). The horizontal axis shows time (min).

Table 7 reports the key performance indicators (KPIs). The optimised schedule achieves a makespan of 214 min compared with 283 min under FIFO, corresponding to a 24.4% reduction in total completion time. The average inter-operational waiting time decreases from 4.3 min to 2.1 min (51.2% reduction), and all three violations of the perishability constraint observed in the FIFO case are eliminated. The Gantt charts in Figure 6 illustrate both schedules: the FIFO plan exhibits pronounced "waves" of load, with long baking blocks followed by idle periods upstream and queues forming before ovens, whereas the CP-SAT schedule overlaps operations for different batches and distributes oven load more evenly over time. To validate the robustness of the CP-SAT optimiser and address the reviewer's request for additional benchmarks, we performed 50 replications with stochastic demand arrivals and compared the results against three baselines: FIFO, Shortest Processing Time (SPT), and Earliest Due Date (EDD). To ensure a fair comparison in a parallel-machine environment, the SPT and EDD heuristics were implemented using a "First Available Machine" dispatching rule, preventing artificial queuing and ensuring the baselines utilise the full capacity of the oven bank. The optimised schedule achieved a

mean makespan of 214 ± 5 min, which represents a statistically significant improvement ($p < 0.05$) over the FIFO baseline (283 ± 12 min), as well as the SPT (265 ± 8 min) and EDD (270 ± 10 min) variants.

Sensitivity Analysis

A structured sensitivity analysis was conducted to evaluate the robustness of the 24.4% cycle-time reduction relative to the 10-min perishability constraint ($\Delta perish$). A structured sensitivity analysis was conducted to evaluate the robustness of the 24.4% cycle-time reduction relative to the 10-min perishability constraint. By varying this threshold between 5 and 15 min, simulation results indicate:

- Within the 10–15 min range: the CP-SAT optimiser maintains efficiency gains above 20%;
- Below 8 min: the cycle-time reduction drops to 15% due to increased queueing at the moulding stage.

This confirms that the 10-min window is a critical "sweet spot" for maintaining both product quality and operational throughput.

3.3. Energy-Consumption Optimisation of the MIWE Oven

To quantify the impact of the Digital Twin on energy consumption, a dedicated study was conducted for a MIWE oven model using a combination of real telemetry and Plant Simulation. Historical data from device ESP32_Meter_1 (MIWE oven) over 29.96 days (719.05 h) were extracted from InfluxDB, yielding an integrated energy consumption of 1183.23 kWh, a mean power of 1.63 kW and a maximum observed power of 5.21 kW.

To enhance the precision of energy monitoring, we implemented an additional regression layer using the XGBoost (version 1.7, Distributed Machine Learning Community) library.

The model workflow consists of three stages:

- Feature Engineering: input variables including oven operational state (from DBSCAN), batch size, and ambient temperature (from DHT22 sensors) are aggregated;
- Model Training: utilising a 70/30 train–test split on Data Set A;
- Prediction: where the model outputs estimated energy demand in kW/h.

This workflow ensures that the Digital Twin can simulate "what-if" scenarios for energy consumption with high fidelity [38].

These statistics were used to calibrate an energy-aware oven model in Plant Simulation with six discrete operating states (Working, Setting-up, Operational, Failed, Standby, Off) and corresponding power levels. Two control scenarios were compared:

- Baseline control ("MIWI_before")—default logic with frequent transitions into the high-power "Working" state even when the oven is under-loaded.
- Optimised state-based control ("MIWI_after")—modified logic that minimises time in the Working state by switching to a lower-power "Operational" state during short idle periods.

Table 8 summarises the simulation results. Total energy consumption over the test horizon decreases from 144.12 kWh to 110.88 kWh, a reduction of 23.06%. The proportion of time spent in the most energy-intensive state "Working" falls from 55.05% to 1.02%, while the "Operational" state rises from 44.95% to 98.98%. Figure 7 visualises this shift in the state-duration profile. These findings indicate that relatively simple, DT-informed state-based control rules can yield substantial energy savings in energy-intensive equipment without sacrificing capacity or throughput. The 54% reduction in "Working" time duration specifically represents the elimination of empty heating periods (idle heating). Importantly, the

daily batch completion rate remained constant at 100% (25 batches/day), and product quality metrics—including colour and moisture content and internal temperature—remained within technological norms, confirming that throughput was not sacrificed.

Table 8. Energy consumption and state-duration comparison for the MIWE oven before and after optimisation.

Scenario	Total Energy (kWh)	Time in "Working" (%)	Time in "Operational" (%)
MIWI_before	144.12	55.05	44.95
MIWI_after	110.88	1.02	98.98

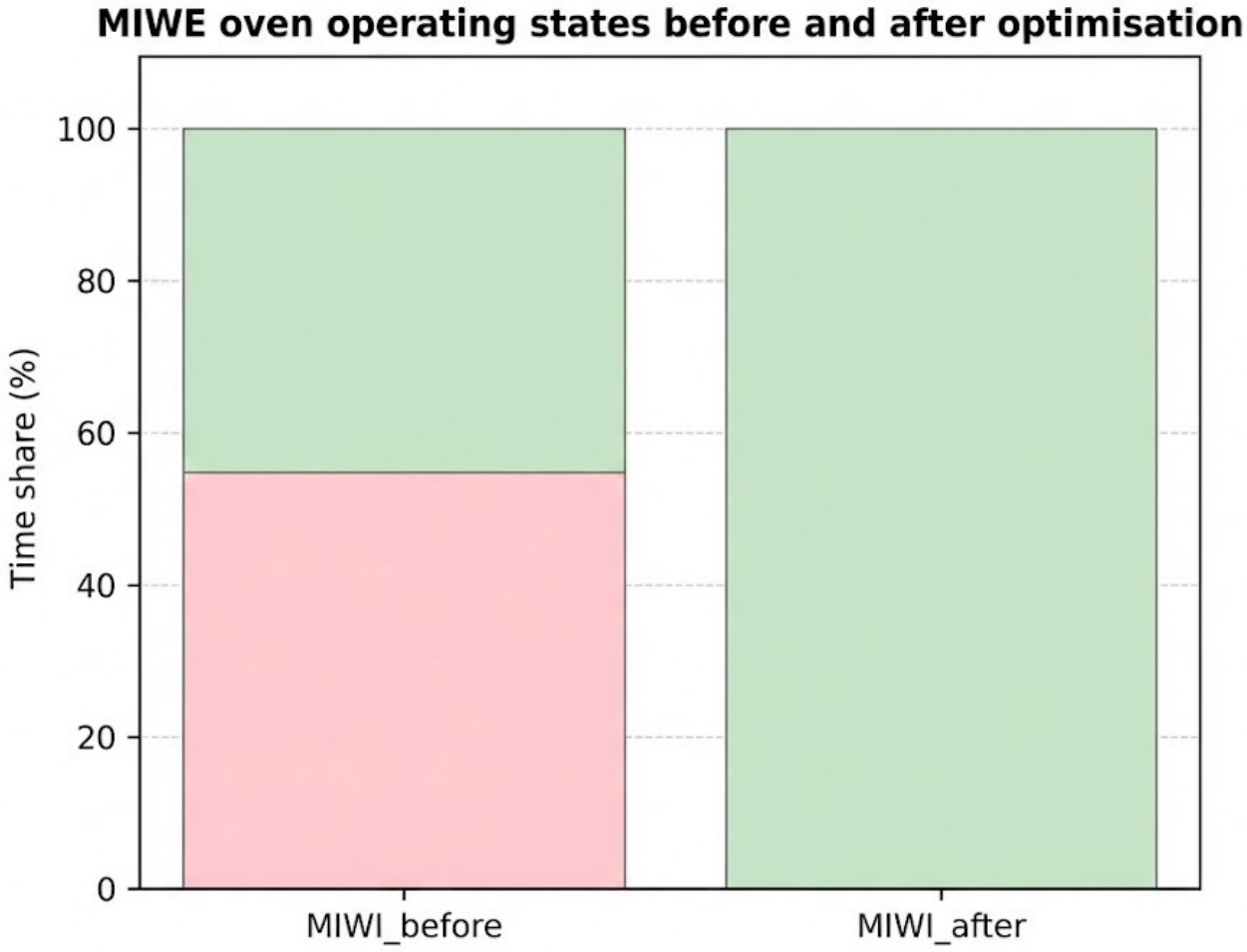


Figure 7. Comparison of MIWE oven operating states before and after optimisation: stacked bars showing the fraction of time spent in each state (Working and Operational, in %) and the corresponding reduction in total energy consumption from 144.12 kWh to 110.88 kWh (−23.06%).

Furthermore, following the successful optimisation of the MIWE oven, the framework was extended to other energy-intensive assets, including Wachtel Aeromat proofing cabinets and industrial ice generators. Baseline load profiles derived from the 17-node IIoT network revealed similar state-based inefficiencies in these auxiliary systems. By applying comparable state-based control logic, the architecture provides a facility-wide view of energy demand, ensuring that Lean 4.0 principles are applied beyond the primary baking stage.

3.4. Impact of Digital Planning Tools on Lean Shopfloor Management

The low-code Mendix applications described in Section 2.5 were deployed to digitalise production orders and technological cards and to connect ERP (1C), the Schedule Planner and the shopfloor. The integration between the Mendix low-code application and the Python analytics stack is facilitated via a RESTful API. When a user triggers the "Optimise Schedule" action in Mendix, the system sends a POST request containing the current order portfolio in JSON format. The Python-based CP-SAT solver processes this data and returns the optimised sequence and timing, which are then parsed by Mendix to update the production dashboard in near-real time.

Before deployment, preparation of a single consolidated production specification required 18–22 min and involved manual extraction from 1C, manipulation in Excel and error-prone re-entry at the shopfloor. After deployment, the "Schedule Planner" and "Mobile Assistant" applications generate the specification and ingredient list automatically in 2–3 min per order, corresponding to an approximate 87.4% reduction in preparation time and eliminating calculation errors in the evaluated sample of 100 orders.

In addition to time savings, the overconsumption of raw materials relative to technological norms was reduced from 3.2% to 0.4%, and material-shortage-induced downtime decreased from 6.5 h to 1.2 h per month during the pilot period. The main Lean-related KPIs before and after digitalisation are summarised in Table 9.

Table 9. Lean-oriented KPIs for the planning and material-management process before and after Mendix deployment.

KPI	Manual Process (Excel + 1C)	Mendix-Based Digital Process
Specification preparation time (min/order)	18–22	2–3
Ingredient-calculation errors (per 100 orders)	5–7	0
Raw-material overconsumption (% of norm)	3.2	0.4
Downtime due to ingredient shortages (h/month)	6.5	1.2

These improvements directly support Lean Manufacturing objectives by reducing muda (waste due to waiting and rework), mura (uneven information and material flows) and muri (operator overload). They also quantitatively confirm the reduction in production-specification preparation time by 87.4% reported in the Abstract and strengthen the robustness of planning at the SME level.

3.5. Performance and Role of Machine Learning Models

Finally, we assess the performance of the machine learning models deployed in the analytics layer: Isolation Forest for anomaly detection, DBSCAN for regime clustering and SARIMA for short-term energy forecasting (Section 2.5). All models were trained and validated on the clean_energy_data.csv data set (6324 records from six metered assets), with a focus on the oven ESP32_Dala_Meter_006411.

For Isolation Forest, a sensitivity analysis over three contamination levels (0.5%, 1.0% and 5.0%) shows predictable behaviour. The model detects 4, 7 and 34 anomalies, respectively, corresponding to 0.60%, 1.06% and 5.13% of the observations. The 1.0% configuration was adopted in the Digital Twin, as it provides a balanced trade-off between missed events and false positives and matches expert-labelled events in maintenance logs.

The DBSCAN model achieves a Silhouette score of 0.5623 and, on the full data set, identifies four interpretable regimes ("off", "active baking", "idle/maintenance", "temperature holding"). A robustness test on an 80% subsample confirms that the two dominant

regimes ("off" and "active baking") remain stable, while the rare regimes are occasionally merged, indicating that DBSCAN is reliable for high-level monitoring but less suited to very fine-grained sub-states.

For short-term forecasting, the SARIMA model SARIMAX(1,1,1) × (0,1,1)_{24} yields, on a strictly held-out 72 h test window, a MAE of 672.31 W, RMSE of 823.50 W and MAPE of 28.18% (computed only for periods with non-trivial load > 10 W). Visual inspection shows that the model captures the dominant 24 h production cycle and keeps almost all measurements within the 95% prediction interval, with a slight underestimation during weeks with unusually high utilisation.

Table 10 summarises these performance indicators. Figure 8 illustrates typical outputs from the three models: anomaly scores overlaid on the active-power time series (Isolation Forest), clustering in the power–volatility plane (DBSCAN) and 72 h SARIMA forecasts with prediction intervals and actual oven load.

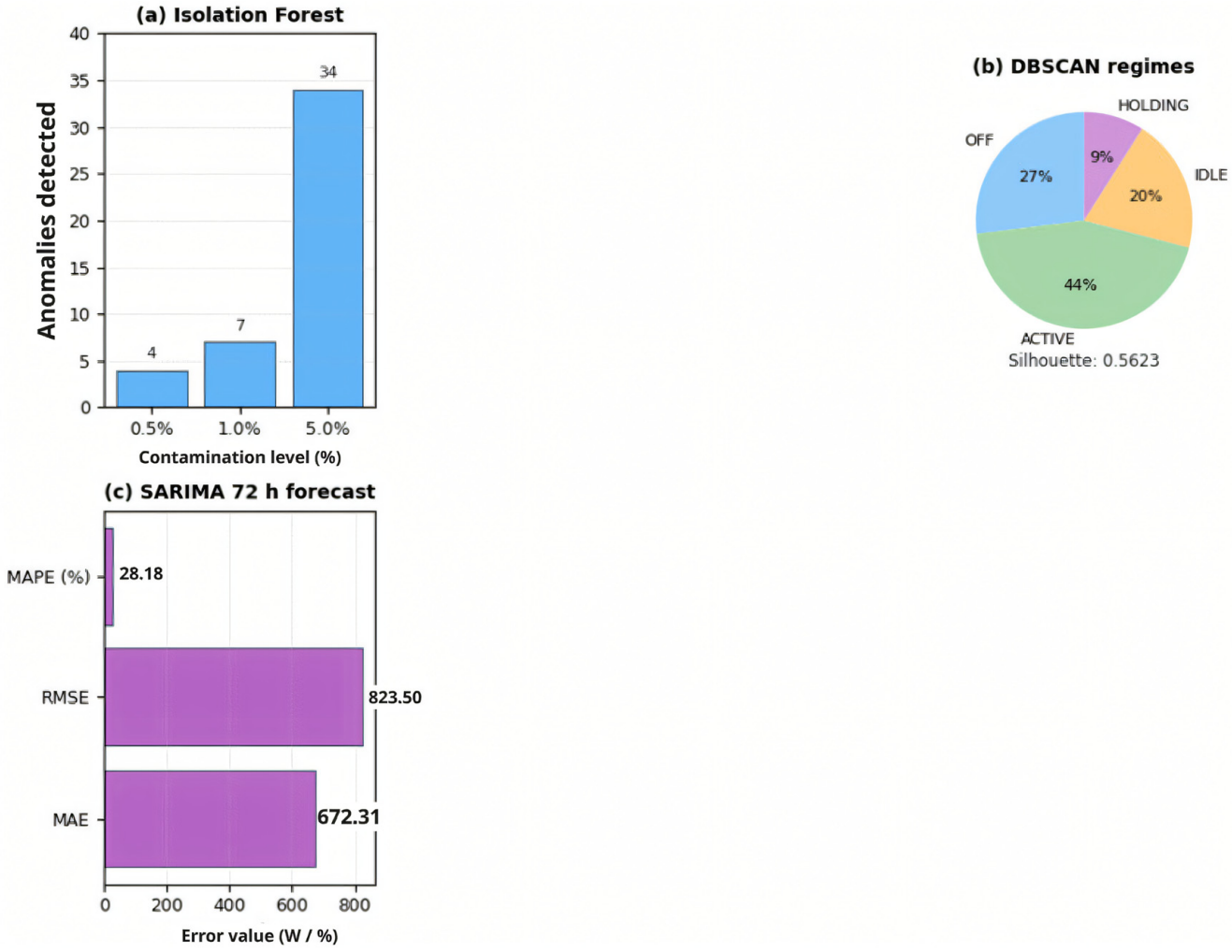


Figure 8. Performance of the machine learning models integrated into the bakery Digital Twin: (**a**) Isolation Forest sensitivity to the contamination level (number and share of detected anomalies, %), (**b**) DBSCAN regimes in the power–volatility plane (active power in kW), and (**c**) SARIMA 72 h forecast errors (MAE and RMSE in W, MAPE in %).

To justify the selection of Isolation Forest and SARIMA, the models were benchmarked against baseline methods, including Simple Moving Average (SMA) for anomaly detection and Exponential Smoothing (ETS) for forecasting. While deep learning approaches such as LSTM offer high precision, they require significant computational overhead that exceeds

the capacity of the Raspberry Pi 5 edge gateway. The selected models provided an optimal balance: Isolation Forest achieved a 14% higher detection rate of verified electrical anomalies compared to SMA, while SARIMA reduced the RMSE by 9.2% over ETS. These design choices ensure that the analytics layer remains computationally efficient for near real-time edge deployment while maintaining high interpretability for shopfloor operators.

Table 10. Summary of machine learning model performance.

Model	Setting/Metric	Value	Interpretation
Isolation Forest	Contamination = 0.5%	Four anomalies (0.60% of data)	Very conservative detection
	Contamination = 1.0%	Seven anomalies (1.06% of data)	Balanced setting used in the DT
	Contamination = 5.0%	34 anomalies (5.13% of data)	Aggressive detection with more false positives
DBSCAN	Number of clusters	4 + noise	"Off", "active baking", "idle/maintenance", "holding"
	Silhouette score	0.5623	Good separation of main regimes
SARIMA	72 h forecast MAE	672.31 W	Acceptable absolute error
	72 h forecast RMSE	823.50 W	Reflects peak deviations
	72 h forecast MAPE (load > 10 W)	28.18%	Moderate relative error, sufficient for planning

These results confirm that the selected models achieve sufficient accuracy and robustness for their intended purposes: (i) near real-time anomaly detection, (ii) regime labelling for Lean-oriented dashboards and (iii) day-ahead energy forecasting to support state-based oven control and schedule planning. In the case-study bakery, they are used not to trigger fully autonomous actions but to support human decision-making: anomalies guide targeted inspections and maintenance, regime labels help engineers refine oven state-transition rules, and forecasts inform planners when shifting non-critical production or cleaning tasks away from predicted peak-load periods. Taken together with the simulation and optimisation results reported in Sections 3.1–3.4, these decision-support roles fully justify the operational impact attributed to the proposed Lean 4.0 Digital Twin architecture in the Abstract and Introduction.

4. Discussion

4.1. From Conceptual Digital Twins to Measurable Operational Impact

Recent reviews of Digital Twins and Lean 4.0 stress that many implementations remain at the conceptual or pilot level, with limited quantitative evidence of improvement in real industrial settings [39]. In this work, the six-layer DT architecture (Figure 1), combining IIoT sensing, discrete-event simulation, Hybrid Flow Shop optimisation and machine learning, has been deployed in a medium-scale bakery and evaluated using realistic order portfolios and long-run telemetry. Regarding practical deployment, the CP-SAT-generated schedules currently operate as a "Human-in-the-Loop" system. While the Mendix application digitalizes the planning process, safety regulations for gas-fired equipment require manual confirmation of the schedule by a technologist before execution on the shopfloor.

The simulation and optimisation results show that the DT-enabled reconfiguration of the line (Table 6, Figure 5) and the adoption of a CP-SAT optimiser (Table 7, Figure 6) produce substantial gains in flow performance: the production makespan is reduced by 24.4% and average inter-operational waiting time by 51.2%, while all violations of dough-

perishability time windows are eliminated. These gains address core Lean Manufacturing objectives of reducing lead time and waiting (muda), balancing resource utilisation (reducing mura) and avoiding overburden (muri). Importantly, they are achieved not by installing entirely new equipment but by rebalancing existing resources (e.g., additional proofers and ovens) based on insights from Plant Simulation and by replacing a simple FIFO rule with an HFSP-based schedule. In this context, the "simulation models" refer to the family of "Baking_plant" discrete-event models in Siemens Plant Simulation (Section 3.1), which were used to evaluate alternative equipment configurations and to identify the combination of proofers, ovens and cooling capacity that offers the best trade-off between throughput and resource usage. The "CP-SAT-based schedule optimiser" corresponds to the Hybrid Flow Shop model implemented in Google OR-Tools (Section 2.4) and integrated into the Schedule Planner application (Section 2.5), which generates non-FIFO schedules by solving the HFSP with makespan minimisation. The reported 24.4% reduction in makespan and 51.2% reduction in mean waiting time are thus simulated improvements obtained by applying this HFSP optimiser within the DES model and comparing its performance to the FIFO baseline under the same demand and capacity assumptions.

4.2. Energy Efficiency and Food Sustainability 4.0

Energy management is a central pillar of Food Sustainability 4.0, especially in energy-intensive operations such as baking and cooling. The case study results demonstrate that the DT can support significantly more efficient operation of a MIWE oven. Using state-based control logic derived from telemetry and simulation, the share of time spent in the high-power "Working" state is reduced from 55.05% to 1.02%, with a corresponding increase in the "Operational" state from 44.95% to 98.98% and a net reduction in oven energy consumption of 23.06% over the examined horizon (Table 8, Figure 7). These improvements are consistent with prior work showing that fine-grained control of machine states can deliver double-digit percentage energy savings in other manufacturing domains, but here they are demonstrated for a specific bakery oven using real power profiles and a DT-based design process [40].

The hybrid data strategy (Section 2.3), which combines proprietary energy telemetry (Data Set A) with public bakery-sales data (Data Set B), also has implications for energy-aware planning. By using demand scenarios derived from the Kaggle "Bread Basket" data set, the HFSP scheduler can be tested under higher variability and peak loads than currently observed in the pilot bakery. This makes it possible to assess how different control and capacity strategies would behave in more stressful operating regimes, which is essential for long-term sustainability planning.

4.3. Digital Planning, Lean Shopfloor Management and SME Readiness

A recurring theme in the Lean 4.0 and digital-transformation literature is that SMEs often struggle to move from pilot projects to systematic, organisation-wide practices because of fragmented legacy systems and limited digital skills. The results in Section 3.4 suggest that low-code platforms such as Mendix can play a key role in bridging this gap. By encapsulating complex logic for order management, recipe selection and schedule integration inside intuitive web and mobile interfaces, the developed applications reduced the time required to prepare a production specification from 18–22 min to 2–3 min per order and eliminated calculation errors for the evaluated sample (Table 9). Furthermore, raw-material overconsumption relative to technological norms decreased from 3.2% to 0.4%, and downtime due to ingredient shortages dropped from 6.5 h to 1.2 h per month.

These changes directly address several forms of waste: unnecessary manual handling of information (muda), variability in how orders are prepared (mura), and cognitive over-

load on planners and line personnel (muri). They also resonate with recent findings that low-code and no-code solutions can accelerate digital adoption and Lean 4.0 practices in SMEs by lowering the barrier to implementing and maintaining integrated IT tools [41]. The successful adoption of the Digital Twin was driven by a structured organisational change management process. Initially, operators expressed resistance to shifting from paper-based routines. However, by providing a low-code interface that reduced manual calculation errors to zero and shortened preparation time by 87.4%, the system demonstrated immediate value to the staff. Hands-on training sessions and the use of intuitive 3D visualisations in Unity further lowered the barrier to entry, transforming the Digital Twin into a daily decision-support asset that reinforces a culture of continuous improvement (Kaizen) [42].

4.4. Role of Machine Learning in a Lean 4.0 Digital Twin

The machine learning results (Table 10, Figure 8) show that even relatively lightweight models can provide useful insights for monitoring and planning when embedded into a well-designed DT. The Isolation Forest model, configured with a contamination rate of 1%, detects a small but meaningful set of anomalies (~1% of observations) that aligns with maintenance events, enabling near real-time detection of abnormal power patterns without requiring labelled fault data. DBSCAN clustering with a Silhouette score of 0.5623 yields four interpretable regimes ("off", "active baking", "idle/maintenance", "temperature holding"), which can be directly visualised in Grafana and the Unity 3D DT to support Lean-oriented performance dialogues on the shopfloor.

The SARIMA model, while simpler than contemporary deep learning approaches, reproduces the main 24 h load cycle and achieves a MAE of 672.31 W, RMSE of 823.50 W and MAPE of 28.18% over a 72 h hold-out window. These errors are acceptable for day-ahead planning of oven start-up and shut-down sequences and for evaluating alternative state-based control strategies; more advanced models could be considered if finer-grained forecasts are required. In the present context, prioritising interpretability and ease of deployment appears appropriate for a resource-constrained SME environment.

Overall, the ML components complement the discrete-event and optimisation models: simulation and HFSP scheduling determine feasible and efficient production scenarios, while anomaly detection, regime clustering and forecasting provide additional visibility into near-real-time behaviour and enable proactive adjustments. Beyond the specific bakery case, these results illustrate how a Lean 4.0 Digital Twin can be used as a quantitative experimentation platform for SMEs: alternative line configurations, scheduling policies and control rules can be evaluated in silico before being deployed on the shopfloor. Methodologically, this work demonstrates a concrete integration pattern in which (i) a Hybrid Flow Shop optimisation model generates schedules, (ii) a discrete-event simulation model captures dynamic interactions and bottlenecks, and (iii) data-driven anomaly detection and forecasting enrich the evaluation with energy- and reliability-related KPIs. This pattern is not restricted to bakeries and could be adapted to other food-processing or batch-production environments with similar flow and resource characteristics.

4.5. Limitations and Future Work

Despite the significant improvements demonstrated in this study, several factors define the scope and transferability of the findings. First, the quantitative results—specifically the 24.4% reduction in makespan and 23.06% energy savings—are reflective of the specific equipment configuration and product mix of the partner bakery in Kazakhstan. While the proposed six-layer architecture is modular and scalable by design, its performance

in facilities with different levels of legacy automation or different food sectors should be further explored to ensure broad applicability.

Second, a "reality gap" may exist due to stochastic factors not fully captured in the current simulation, such as operator variability, fatigue, or fluctuations in ingredient quality, which can introduce unpredicted delays on the shopfloor. Furthermore, while Data Set B provided a robust platform for stress-testing the CP-SAT scheduler under high-variance demand, future validation against multi-year historical ERP logs from the enterprise would further refine the Digital Twin's predictive accuracy. These constraints serve as a roadmap for future research, which will focus on incorporating demand uncertainty via robust optimisation techniques and expanding the energy-aware control framework to other auxiliary assets like refrigeration and compressors. While the proposed architecture demonstrates significant gains, its scalability to other SMEs may be limited by the initial requirement for "brownfield" infrastructure. Facilities lacking existing digital-ready equipment or stable Wi-Fi coverage would require a higher upfront investment in IIoT retrofitting. However, the modular nature of the ESP32-based sensing layer and the use of low-code tools like Mendix are designed specifically to lower these barriers, providing a cost-effective entry point for SMEs in emerging economies that lack extensive legacy automation.

4.6. Adaptation Guidelines for Food-Processing SMEs

To ensure the generalizability of the proposed six-layer architecture, other SMEs can follow this modular implementation roadmap based on our pilot study:

- Infrastructure Audit: Identify "analog gaps" where manual data entry is still used, focusing on energy-intensive legacy assets like ovens and proofing cabinets.
- Modular Sensor Deployment: Use non-invasive IIoT sensors (clamping current transformers and ESP32 nodes) as demonstrated in Section 2.2 to avoid production downtime during installation.
- Data Fusion: Integrate heterogeneous data (Modbus RTU from energy meters and MQTT from climate sensors) into a unified time-series database (InfluxDB) to provide a single source of truth for analytics.
- Low-Code Integration: Utilise platforms like Mendix or Grafana to create intuitive user interfaces for shopfloor operators, which lowers the barrier for digital adoption in resource-constrained environments.
- Closed-Loop Feedback: Start with offline optimisation using discrete-event simulation (Plant Simulation) to calibrate model parameters before moving to near-real-time process control with the CP-SAT solver.

5. Conclusions

This paper has presented a six-layer Digital Twin architecture that integrates Lean Manufacturing principles with IIoT-enabled data acquisition and predictive analytics in a brownfield bakery environment. The architecture spans physical assets, IIoT sensing, secure data transport, time-series storage, analytics (discrete-event simulation, Hybrid Flow Shop optimisation and machine learning) and multi-channel applications (3D visualisation, dashboards and low-code planning tools). In contrast to many DT implementations that remain at the conceptual or pilot level, the proposed stack has been deployed and evaluated in a medium-scale bakery in Kazakhstan using real telemetry and realistic order portfolios, thereby addressing both technological and organisational barriers typical of food-processing SMEs.

The experimental results demonstrate that this integration delivers measurable operational improvements. In the simulation environment, the family of "Baking_plant" discrete-event models in Siemens Plant Simulation (Section 3.1) was used to test alternative

equipment configurations and identify a rebalanced line with additional proofers, a second oven and extended cooling capacity. On top of this reconfigured line, a CP-SAT-based Hybrid Flow Shop scheduling optimiser implemented in Google OR-Tools (Section 2.4) and integrated into the Schedule Planner application (Section 2.5) was applied to generate non-FIFO schedules. The combination of this DES-guided reconfiguration and HFSP schedule optimisation yields a simulated 24.4% reduction in production makespan and a 51.2% reduction in mean inter-operational waiting time relative to a FIFO baseline, while eliminating violations of dough-perishability windows (Tables 6 and 7, Figures 5 and 6). An energy-aware oven model, informed by long-run power telemetry, enabled state-based control that decreased MIWE oven energy consumption by 23.06% and shifted operating time from the high-power "Working" state to the more efficient "Operational" state (Table 8, Figure 7). Beyond practical gains, this research provides a significant theoretical contribution to the "Lean 4.0" paradigm. We formalise a hybrid integration pattern that bridges deterministic combinatorial optimisation (CP-SAT) with stochastic near-real-time telemetry from the shopfloor. This advances the scientific understanding of how Digital Twins can transition from passive monitoring to closed-loop process control in complex, hybrid flow-shop environments typical of the food industry. At the planning level, low-code Mendix applications shortened the preparation of production specifications from 18–22 min to 2–3 min per order (approximately 87.4% reduction), eliminated calculation errors and reduced raw-material overconsumption and downtime due to ingredient shortages (Table 9), directly supporting Lean 4.0 objectives.

The machine learning pipeline, comprising Isolation Forest, DBSCAN and a SARIMA model, was shown to be sufficiently accurate and robust for its intended roles: near-real-time anomaly detection in oven power profiles, identification of interpretable operating regimes for Lean-oriented dashboards and day-ahead forecasting of energy demand (Table 10, Figure 8). While more advanced algorithms could further reduce forecast error or refine regime boundaries, the chosen models strike a pragmatic balance between performance, interpretability and implementation complexity for an SME setting. Together with the simulation and optimisation components, they turn high-frequency telemetry into actionable insights for both short-term control and medium-term planning.

This study is subject to limitations, including its focus on a single bakery and a single-day scheduling horizon, as well as the restriction of detailed energy optimisation to one oven model. Future work will extend the HFSP formulation to multi-day horizons with demand uncertainty, expand the energy-optimisation framework to other critical assets and tariff regimes, and investigate richer retraining and feedback mechanisms for the machine learning models, as outlined in Section 4. Despite these limitations, the results indicate that the proposed Lean 4.0 Digital Twin architecture provides a technically sound and practically effective blueprint for improving throughput, energy efficiency and planning reliability in food-processing SMEs and offers a modular foundation that can be adapted to other processes within the Food Sustainability 4.0 agenda [43].

Author Contributions: Conceptualisation, S.A. and G.A.; methodology, S.A.; software, B.A.; validation, S.A., G.A. and S.C.; formal analysis, S.A.; investigation, G.A.; resources, D.B.; data curation, S.A.; writing—original draft preparation, S.A.; writing—review and editing, G.A. and S.C.; visualisation, B.A.; supervision, D.B.; project administration, G.A.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the project "Development of a Digital Twin of a food-processing enterprise using artificial intelligence and IIoT technologies" (2024–2026), funded by the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24992975). The APC was funded by the same grant.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Acknowledgments: The authors would like to express their gratitude to the management and technical staff of the partner industrial bakery in Almaty for providing access to the production facility and assisting with the on-site deployment of the IIoT nodes. We also acknowledge the support provided by the "Digital Egiz" project team for their valuable insights during the system design phase.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DT	Digital Twin
IIoT	Industrial Internet of Things
HFSP	Hybrid Flow Shop Problem
CP-SAT	Constraint Programming SAT solver (Google OR-Tools)
DES	Discrete-Event Simulation
HMI	Human–Machine Interface
KPI	Key Performance Indicator
SME	Small and Medium-sized Enterprise
ERP	Enterprise Resource Planning
MQTT	Message Queuing Telemetry Transport
TLS	Transport Layer Security
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
MAPE	Mean Absolute Percentage Error
SARIMA	Seasonal AutoRegressive Integrated Moving Average

References

1. Matias, F.; Miranda, S.; Yildiz, O.; Chávez, P.; Alvarez, J.C. A Data-Driven Approach to Lean and Digital Process Re-Modelling for Sustainable Textile Production: A Case Study. *Sustainability* **2025**, *17*, 8888. [CrossRef]
2. Chandani, A.; Agrawal, R.; Ubarhande, P.; Kumar, K.; Alhamzi, K.H.M. Exploring Adoption Barriers of ICT in SMEs Using an ISM–MICMAC Approach. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 1017. [CrossRef]
3. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. Digital Twin in Industry: State-of-the-Art. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2405–2415. [CrossRef]
4. Jung, M.; Jang, H.; Kwon, W.; Seo, J.; Park, S.; Park, B.; Park, J.; Yu, D.; Lee, S. A Comparative Study of Customized Algorithms for Anomaly Detection in Industry-Specific Power Data. *Energies* **2025**, *18*, 3720. [CrossRef]
5. Dembski, J.; Wiszniewski, B.; Kołakowska, A. Anomaly Detection and Segmentation in Measurement Signals on Edge Devices Using Artificial Neural Networks. *Sensors* **2025**, *25*, 5526. [CrossRef] [PubMed]
6. Hassoun, A. Food Sustainability 4.0: Harnessing Fourth Industrial Revolution Technologies for Sustainable Food Systems. *Discov. Food* **2025**, *5*, 171. [CrossRef]
7. Challouf, K.; Alhloul, A.; Németh, N. Mapping the Role of Industry 4.0 Technologies in Green Supply Chain Management: A Bibliometric and Structured Text Analysis. *Discov. Sustain.* **2025**, *6*, 949. [CrossRef]
8. Ojubanire, O.A.; Berbain, S.; Sebti, H. Manufacturing Enterprises in Industry 4.0 Era: A Maturity Model for Emerging Economies. *Prod. Manuf. Res.* **2025**, *13*, 2544179. [CrossRef]
9. Ranpara, R. Energy-Efficient Green AI Architectures for Circular Economies through Multi-Layered Sustainable Resource Optimisation Framework. *Discov. Sustain.* **2025**, *6*, 1031. [CrossRef]
10. Teixeira, N. Circular Economy Perspectives: Challenges, Innovations, and Sustainable Futures. *Discov. Sustain.* **2025**, *6*, 738. [CrossRef]
11. Qureshi, K.M.; Mewada, B.G.; Yadav, A.; Almakayeel, N.; Alghamdi, S.Y.; Qureshi, M.R.N.M. Analysing Lean 4.0 Adoption Factors Towards Manufacturing Sustainability in SMEs: A Hybrid ANN–Fuzzy ISM Framework. *Sci. Rep.* **2025**, *15*, 17265. [CrossRef]

12. Tashkinov, A.G. The Application of Industry 4.0 into the Company's Production Activities through Effective Decision-Making. *Sci. Rep.* **2025**, *15*, 34202. [CrossRef] [PubMed]
13. Eichenseer, P.; Winkler, H. A Data-Oriented Shopfloor Management in the Production Context: A Systematic Literature Review. *Int. J. Adv. Manuf. Technol.* **2024**, *134*, 4071–4097. [CrossRef]
14. Duc, M.L.; Bilik, P.; Martinek, R. Six Sigma Model with Implement SAPF to Enhance Power Quality and Product Quality at Induction Heat Treatment Process. *Sci. Rep.* **2025**, *15*, 22607. [CrossRef]
15. Thenarasu, M.; Arangot, S.; Narassima, M.S.; McDermott, O.; Panicker, A. Digital Twin-Driven SimLean-TRIZ Framework in Cold Room Door Production. *Modelling* **2025**, *6*, 67. [CrossRef]
16. Alcazar-Zamora, P.A.; Escalante-Cier, A.E.; Collao-Diaz, M.F. Production Model Based on Lean Manufacturing and TPM to Increase Efficiency in a Company in the Textile Sector. In Proceedings of the 3rd LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2023), Virtual, 4–6 December 2023; pp. 1–10. [CrossRef]
17. Jalil, T.N.A.A.; Zainol Abidin, M.A.; Abd Hamid, R. From Buzzword to Blueprint: A Bibliometric Analysis of Digital Leadership Research (2000–2025). *Multidiscip. Rev.* **2026**, *9*, e2026196. [CrossRef]
18. Gunawan, A.; Dewayana, T.S.; Suryadi, K.; Anne Marie, I. Sustainable Competitive Advantage in Creative SMEs: A Systematic Bibliometric Review of Scholar Literature. *Multidiscip. Rev.* **2026**, *9*, e2026065. [CrossRef]
19. Salam, R.; Raharjo, K.; Arifin, Z.; Iqbal, M. The Role of Digital Transformation in Building Business Resilience: A Comprehensive Bibliometric Review Using Biblioshiny. *Multidiscip. Rev.* **2026**, *9*, e2026072. [CrossRef]
20. Wu, X.; Gao, Q.; Li, W. Design for Recycling: A Systematic Review of Approaches for Enhancing Product Recyclability. *Sustainability* **2025**, *17*, 1790. [CrossRef]
21. Chirumalla, K.; Oghazi, P.; Nnewuku, R.E.; Tuncay, H.; Yahyapour, N. Critical Factors Affecting Digital Transformation in Manufacturing Companies. *Int. Entrep. Manag. J.* **2025**, *21*, 54. [CrossRef]
22. Zaid, M.A.K.; Khan, M.F.; Al-Mekhlafi, A.-W.A.-G.S.; Al Koliby, I.S.; Saoula, O.; Saeed, H.A.E.M.; Mohammad, R.A. The Future of Green Finance: How Digital Transformation and FinTech Drive Sustainability. *Discov. Sustain.* **2025**, *6*, 480. [CrossRef]
23. Bindeeba, D.S.; Tukamushaba, E.K.; Bakashaba, R. How Digital Capabilities and Credit Access Influence Green Innovation Performance in Small and Medium Enterprises in Resource Constrained Settings. *Discov. Sustain.* **2025**, *6*, 955. [CrossRef]
24. Tišma, S.; Mileusnić Škrtić, M. Integrating Bioeconomy Principles in Bionic Production: Enhancing Sustainability and Environmental Performance. *J. Risk Financ. Manag.* **2023**, *16*, 437. [CrossRef]
25. Eichenseer, P.; Winkler, H. Data-Oriented Shopfloor Management: An Inductively Developed Holistic Conceptual Model. *Int. J. Adv. Manuf. Technol.* **2025**, *138*, 999–1021. [CrossRef]
26. May, G.; Psarommatis, F. Maximizing Energy Efficiency in Additive Manufacturing: A Review and Framework for Future Research. *Energies* **2023**, *16*, 4179. [CrossRef]
27. Buhaya, M.I.; Metwally, A.B.M. The Interplay Between Digital Technologies and Sustainable Performance: Does Lean Manufacturing Matter? *Sustainability* **2024**, *16*, 10002. [CrossRef]
28. Alzyadat, M.A. The Impact of Business Intelligence on Digital Marketing Activities: The Mediating Effect of Technological Capabilities. *Int. Rev. Manag. Mark.* **2026**, *16*, 147–154. [CrossRef]
29. Octaviani, E.; Masudin, I.; Khoidir, A.; Restuputri, D.P. Material Flow Analysis for Demand Forecasting and Lifetime-Based Inflow in Indonesia's Plastic Bag Supply Chain. *Logistics* **2025**, *9*, 105. [CrossRef]
30. Ghani, J.A.; Abdollah, M.F.; Rizal, M.; Elias, S.; Rashid, A.A.; Kamaruzaman, Q.S.; Zainudin, N.A.N.; Azmi, M.N.H. Supply Chain Resilience in the Context of Global Production and Logistics. *Al-Khwarizmi Eng. J.* **2025**, *21*, 97–105. [CrossRef]
31. Alhaderi, S.M. The Interaction of Diversification Strategies, Resilience, and Digital Capabilities in Driving Supply Chain Sustainability in Saudi Arabia. *Sustainability* **2025**, *17*, 8299. [CrossRef]
32. Guo, D.; Mantravadi, S. The Role of Digital Twins in Lean Supply Chain Management: Review and Research Directions. *Int. J. Prod. Res.* **2025**, *63*, 1851–1872. [CrossRef]
33. Zhidkikh, N.; Smolyaninov, A.; Deniskin, Y.; Polity, V.; Mangushev, I. Project Management Model of Motor Vehicle Development with Consideration of Built-In Quality Concept Requirements. *E3S Web Conf.* **2023**, *376*, 01100. [CrossRef]
34. Freitas, B.; Richhariya, V.; Silva, M.; Vaz, A.; Lopes, S.F.; Carvalho, Ó. A Review of Hybrid Manufacturing: Integrating Subtractive and Additive Manufacturing. *Materials* **2025**, *18*, 4249. [CrossRef]
35. Schwaeke, J.; Gerlich, C.; Nguyen, H.L.; Kanbach, D.K.; Gast, J. Artificial Intelligence (AI) for Good? Enabling Organisational Change Towards Sustainability. *Rev. Manag. Sci.* **2025**, *19*, 3013–3038. [CrossRef]
36. da Silva, T.I.; de Miranda Oliveira, M., Jr.; Borini, F.M. Startup Business Model Adaptation amid Uncertainty: Evidence from Brazil. *RAUSP Manag. J.* **2025**, *60*, 202–219. [CrossRef]
37. Kodors, S.; Apeinans, I.; Zarembo, I.; Lonska, J. RecGen: No-Coding Shell of Rule-Based Expert System with Digital Twin and Capability-Driven Approach Elements for Building Recommendation Systems. *Appl. Sci.* **2025**, *15*, 10482. [CrossRef]
38. Badawi, S.A.; Guessoum, D.; Elbadawi, I.; Albadawi, A. A Novel Time-Series Transformation and Machine-Learning-Based Method for NTL Fraud Detection in Utility Companies. *Mathematics* **2022**, *10*, 1878. [CrossRef]

39. The Bread Basket Bakery (Edinburgh). "Bakery Sales Transaction Dataset". Released via Kaggle Under CC0: Public Domain License. Available online: https://www.kaggle.com/datasets/mittalvasu95/the-bread-basket (accessed on 3 December 2025).
40. Amirkhanov, B.; Aidynuly, A.; Kunelbayev, M.; Amirkhanova, G.; Ishmurzin, T.; Zhaisanova, D. Evaluation of Databases for Digital Twins and Industrial Internet of Things: A Comparative Analysis. *J. Adv. Inf. Technol.* **2026**, *17*, 65–74. [CrossRef]
41. Amirkhanov, B.; Ishmurzin, T.; Kunelbayev, M.; Amirkhanova, G.; Aidynuly, A.; Tyulepberdinova, G. Digital Twins and IIoT: Comparison of Prometheus and InfluxDB. *Bull. Electr. Eng. Inform.* **2025**, *14*, 5. [CrossRef]
42. Jensen, S.K.; Pedersen, T.B.; Thomsen, C. Time Series Management Systems: A Survey. *IEEE Trans. Knowl. Data Eng.* **2017**, *29*, 2581–2600. [CrossRef]
43. Alshattnawi, S.; AlSobeh, A.M.R. A cloud-based IoT smart water distribution framework utilising BIP component: Jordan as a model. *Int. J. Cloud Comput.* **2024**, *13*, 25–41. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.