DPIR-GSky: An Integrated Digital Twin Framework for Streaming Multi-Entity Intervention Prioritization in Vegetable Cold Storage

Triawan Adi Cahyanto[1], Taufik Djatna[2*], Annisa[3], Heru Sukoco[4], Hendra Rahmawan[5], Putro Ferro Ferdinant[6], Agung Prayudha Hidayat[7]

[1,3,4,5]*Doctoral Program in Computer Science, School of Data Science, Mathematics, and Informatics, IPB University, Bogor 16680, Indonesia*

[1]*Department of Informatics Engineering, Faculty of Engineering, Universitas Muhammadiyah Jember, Jember 68121, Indonesia*

[2]*Division of Agroindustrial Engineering and Technology, Faculty of Engineering and Technology, IPB University, Bogor 16680, Indonesia*

[6]*Department of Industrial Engineering, Faculty of Engineering, Universitas Sultan Ageng Tirtayasa, Serang 42435, Indonesia*

[7]*Industrial Management, Faculty of Vocational School, IPB University, Bogor 16128, Indonesia*

(Received: April 10, 2026; Revised: June 15, 2026; Accepted: September 10, 2026; Available online: October 29, 2026)

Abstract

When several vegetable cold storage units deviate from safe environmental conditions simultaneously, deciding which unit needs intervention first is difficult. Multi-criteria ranking methods generally require criterion weights fixed in advance, which does not fit well with continuously changing streaming conditions. This study developed and evaluated a digital twin framework built around DPIR-GSky, an incremental, dominance-power-based group skyline algorithm that prioritizes intervention across multiple storage entities without manually elicited criterion weights. The framework integrates streaming sensor acquisition, a virtual twin layer, and adaptive pruning with dominance-power scoring for sliding-window operation. Readings for temperature, relative humidity, ethylene, carbon dioxide, and illuminance were aggregated into six-dimensional deviation vectors, one combining temperature and ethylene into a shelf-life-based urgency index. The framework was evaluated using seven days of real sensor data from twenty storage racks for calibration and benchmarking, together with five simulated 72-hour disturbance scenarios generated from an autoregressive model calibrated on the real data, scored against independently defined ground-truth labels and compared with two uniformly weighted multi-criteria baselines. The virtual twin achieved a modest average coefficient of determination of 0.258, reflecting low signal variance under stable operating conditions, and is therefore better suited to directional trend monitoring than precise point prediction. DPIR-GSky reached an average F1-score of 0.685, trailing the weighted baselines by 0.042 points overall, yet it matched or exceeded both baselines under a heterogeneous, multi-variable disturbance scenario, detected an ethylene-related disturbance 65 minutes earlier than static-threshold monitoring, and processed each streaming window in 0.32 milliseconds at a throughput above 37,000 tuples per second. These findings indicate that the proposed framework is computationally and functionally feasible for real-time intervention prioritization without manually elicited criterion weights, with its advantage concentrated where several variables deviate together rather than one variable alone; validation at optimal 0–2°C storage conditions and across additional facilities remains necessary before operational deployment.

Keywords: Digital Twin, Group Skyline Query, Streaming Data, Multi-Entity Prioritization, Vegetable Cold Storage

1. Introduction

Postharvest shrinkage in vegetables directly affects their quality, distribution efficiency, and food security. As fresh commodities, vegetables are prone to quality deterioration when storage conditions are not optimal, for example, due to changes in temperature, relative humidity (RH), ethylene accumulation, increased CO_2 levels, or exposure to inadequately controlled environments [1], [2], [3]. Cold storage is used to slow down the ripening process through storage environment control [4]. In addition to temperature and humidity control, data-based early detection and computer vision are increasingly used to assess the quality of leaf commodities non-destructively [5]. Nevertheless, the success of cold storage is determined not only by the availability of refrigeration space but also by the system's ability to detect changes in conditions and respond rapidly. In practice, monitoring based on static thresholds tends to

***Corresponding author: Taufik Djatna(taufikdjatna@apps.ipb.ac.id)**
DOI: https://doi.org/10.47738/jads.v7i4.1507
This is an open access article under the CC-BY license (https://creativecommons.org/licenses/by/4.0/).
© Authors retain all copyrights

treat each sensor in isolation; therefore, the relationships between variables and the need to prioritize between zones are often missed when disturbances occur simultaneously [6], [7], [8]. The development of the Internet of Things (IoT) has enabled the continuous collection of sensor data from various storage points. However, large data streams require analytical approaches that can transform raw sensor readings into actionable information. Digital twins address this need as a virtual representation of cold storage synchronized with actual sensors, thus enabling the estimation of derivative conditions, such as degradation rate and remaining shelf life [9], [10], [11]. In line with this, studies on precision agriculture decision support systems emphasize the importance of real-time data and analytics for resource allocation and operational decision-making [12], whereas the implementation of IoT in precision livestock systems demonstrates the benefits of real-time sensors for the continuous monitoring of environmental conditions and biological status [13].

When several storage entities experience deviations simultaneously, the system must objectively determine the order of intervention. Aggregate-score and linear-weighting approaches still rely on criterion weights set at the outset, even though the importance of each variable can shift with the operating conditions. The Group Skyline Query (G-Skyline) offers an alternative: it identifies groups of entities that are not dominated by other groups without assigning explicit weights, a capability that has not yet been combined with digital twins for cold storage. Existing digital twin cold storage studies can monitor or predict the state of individual entities; however, they do not solve the online problem of selecting a fixed-capacity intervention group from several simultaneously deviating entities without manually elicited criterion weights. This study refers to the task of ranking several simultaneously deviating entities as multi-entity intervention prioritization, which is the central problem addressed here.

This study proposes an integrated digital twin framework incorporating a dominance power-based incremental refinement G-Skyline (DPIR-GSky) for monitoring vegetable cold storage. Its main contributions are as follows: (a) an integrated digital twin framework featuring DPIR-GSky for vegetable cold storage; (b) storage-entity representation as multidimensional objects encompassing environmental deviations and RSL-based quality impact; (c) an efficient incremental update mechanism with adaptive thresholds and dominance-power scoring for sliding-window streaming data; (d) soft real-time characterization (latency, jitter, throughput, WCET) with complexity and scalability analysis under increasing load; and (e) comparative evaluation against Naive-GSky, MDS, GPR, and SW-GSky algorithmic baselines for computational performance, and against uniformly weighted TOPSIS and PROMETHEE II for prioritization effectiveness across simulated operational scenarios.

2. Literature Review

2.1. Digital Twin in Agri-Food Systems

The digital twin concept, introduced by Grieves [14], is commonly defined as a virtual model that continuously reflects the condition of a physical entity, consisting of the physical entity, its virtual counterpart, and the data connection that synchronizes them. Kritzinger et al. [15] distinguished integration maturity by the directionality and automation of this data flow: a digital model exchanges data manually in both directions, a digital shadow provides automated, one-way sensing without automated actuation, and a digital twin closes the loop with an automated flow back to the physical entity. This distinction is later used to classify the maturity level of the proposed framework. Subsequent work extended this base model with data and service dimensions and broader functions such as monitoring, deterioration prediction, control optimization, and decision support [16], [17]. However, most agri-food implementations remain limited to one or two functions, and existing digital twins for smart farming and fruit supply chains [18], [19] do not accommodate streaming-based prioritization when multiple storage units deviate simultaneously, which is the gap this study addresses.

2.2. Streaming Data Processing in Food IoT Systems

Data streaming in IoT systems is characterized by high volume, velocity, variety, and veracity [20], forming a continuous heterogeneous flow that cannot be processed using conventional batch methods. It is commonly handled using a sliding-window mechanism that restricts the analysis to the most recent observations within a window of width W [21] while remaining robust to concept drift [22]. This paradigm is directly relevant to cold-storage monitoring, although its application in food surveillance remains limited. Xie et al. [23] applied sliding-window anomaly detection

to cold-chain logistics data without a digital twin or an entity-prioritization mechanism, and Sundmaeker et al. [24] identified persistent gaps between IoT infrastructure capability and the analytics needed for responsive food surveillance.

2.3. Skyline Query and G-Skyline

The skyline query, formalized by Börzsönyi et al. [25], identifies points in a multidimensional space that are not dominated by any other point, meaning that they are at least as good in every dimension and strictly better in at least one dimension, without requiring a single objective function. The Group Skyline Query (G-Skyline) extends this concept to groups of objects by selecting groups that are not dominated by other groups [26]. In this study, each rack was treated as a multidimensional object over deviations in temperature, RH, ethylene, CO_2, and light, plus the estimated remaining shelf life, so intervention priorities could be established directly from dominance relationships without assigning subjective criterion weights.

2.4. Cold Storage Management for Vegetables

The environmental storage requirements for vegetables vary by commodity. Leafy vegetables, such as lettuce and spinach, require a temperature of 0-2°C and RH of 95-100%, whereas tomatoes require 8-12°C and RH of 85-95% [27]. Deviation from optimal conditions affects quality through several mechanistic pathways: high temperature accelerates respiration and ethylene production, low relative humidity (RH) triggers water loss and wilting, and excessive CO_2 or ethylene induces undesirable changes in color and texture. Van Boekel [28] reviews kinetic models for predicting postharvest quality, including the Arrhenius and Q_{10} models. In the broader context of perishable systems, data-driven optimization models have integrated product deterioration functions into production and labor scheduling to minimize spoilage costs [29]. This study adopts the Q_{10} model (Equation 1), a practical equivalent of the Arrhenius formulation, which states that the reaction rate increases by a factor of Q_{10} for every $10°C$ rise in temperature.

$$k_T = k_{ref} \cdot Q_{10}^{(T-T_{ref})/10} \quad (1)$$

with a Q_{10} value of approximately 2–3 for most horticultural products.

Manolopoulou et al. [30] show that the storage quality of fresh-cut lettuce depends jointly on temperature and package atmosphere, which supports treating storage monitoring as a multivariate rather than single-variable problem.

2.5. Synthesis of Gaps and Research Position

Table 1 summarizes a comparison of previous relevant studies based on their dimensions. Previous studies indicate that digital twins, streaming processing, and multi-entity prioritization are still being developed independently. Because weighting-based multi-criteria decision-making (MCDM) methods, such as the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) and Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE), require explicit criterion weights [31], [32], this study positions G-Skyline as an alternative to Pareto prioritization for the digital twin architecture of vegetable cold storage.

Table 1. Comparison of dimensions in previous studies relevant to the proposed framework

Research	Digital Twin	Stream Processing	Multi-criteria	Cold Storage of Vegetables	Multi-entity Priority
[18]	✓	Partial	✗	✗	✗
[19]	✓	✗	✗	Partial	✗
[26]	✗	✗	✓ (G-Skyline)	✗	✗
This study	✓	✓	✓ (G-Skyline)	✓	✓

2.1. Comparison Methods for Multi-Criteria Evaluation

TOPSIS and PROMETHEE II were chosen as experimental baselines because they are widely used weight-dependent MCDM methods and both require explicit criterion weights, which are difficult to fix in advance when streaming conditions change. DPIR-GSky differs because it identifies a Pareto-optimal non-dominated subset rather than a total order, which suits settings where the intervention capacity is limited to a fixed group size. Both baselines use uniform

weights ($w_j = 1/d$ for all $d = 6$ dimensions), consistent with the absence of weight specifications from domain experts: TOPSIS ranks entities by their closeness coefficient to the positive and negative ideal solutions, and PROMETHEE II ranks entities by their net outranking flow based on pairwise preference. Top-k entities under each method are selected as the intervention priority group for comparison with the G-Skyline output.

3. Methodology

3.1. Research Design, Study Setting, and Variables

The research followed five phases, as summarized in figure 1: (1) problem identification and study setup; (2) acquisition of real streaming data and AR(1)-based scenario generation; (3) framework development, comprising sliding-window processing, a virtual twin with six-dimensional urgency representation, and the DPIR-GSky algorithm; (4) evaluation against an independently constructed ground truth; and (5) analysis and conclusion.

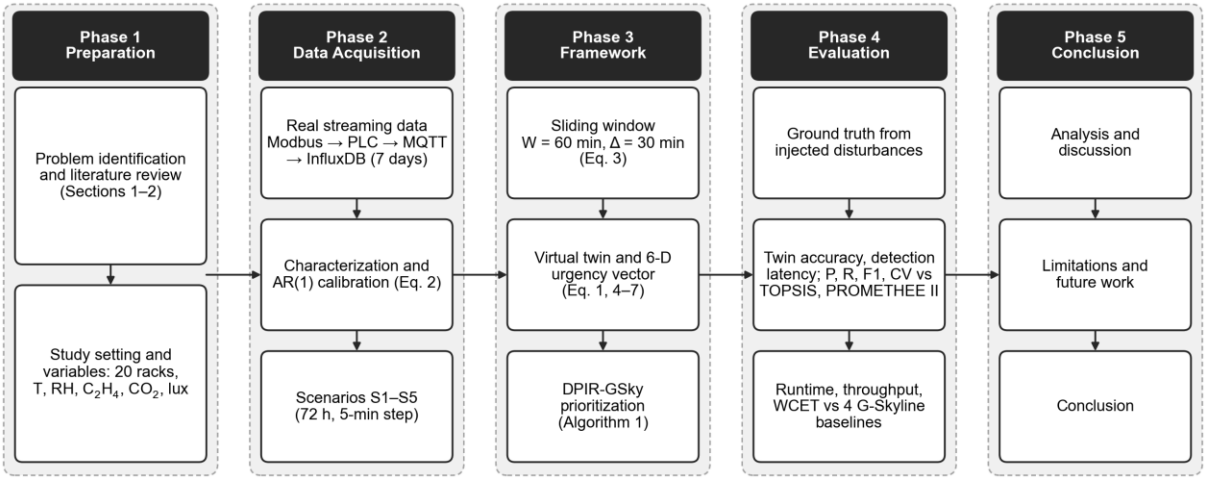


Figure 1. Research flowchart of the proposed DPIR-GSky digital twin framework.

This study adopted a systems engineering approach based on scenario simulation to design and validate an integrated framework, with functional validation of the acquisition pipeline using real sensor data from the ATP Packing House cold storage facility at the IPB University. The study object was a cold storage facility for curly lettuce (Lactuca sativa var. crispa), with a capacity of 10 tons and 20 racks in four thermal zones; curly lettuce was chosen for its high respiration rate and temperature sensitivity, its status as a highly perishable leafy vegetable, and the availability of deterioration kinetics reference data. The facility and its virtual representation are shown in figure 2.



Figure 2. Physical cold storage facility at the ATP Packing House and its virtual representation as part of the digital twin framework.

Each rack was monitored using five environmental variables: temperature ($PT100$, $\pm 0.2°C$, optimal $0 - 2°C$), RH (capacitive sensor, $\pm 1.5\%\ RH$, optimal $95 - 100\%$), ethylene (electrochemical, 0.05 ppm detection limit, optimal $<$ 0.1 ppm), CO_2 (NDIR, ± 100 ppm, optimal $400 - 1{,}000$ ppm), and illuminance (lux, optimal < 50 lx, a proxy for leaf-discoloration risk). Additional operational variables included door-opening frequency, energy consumption per zone, and an 8-hourly visual quality score.

3.2. Data Sources and Sensor Infrastructure

This study uses four data sources: (1) real pilot sensor data from a single rack ($\approx 11.5\ h$, February 2026) to validate the acquisition pipeline; (2) a real seven-day, 20-rack dataset at 1-minute resolution; (3) $AR(1)$-simulated scenario data calibrated from the seven-day dataset; and (4) the AR(1) generator itself, which reproduces periodic fluctuations from compressor cycles and door openings. Below-ambient CO_2 readings (mean 97.66 ppm) were attributed to active scrubbing by the facility's atmosphere-control system rather than a deviation trigger. Independent verification of the NDIR calibration and scrubbing system logs was not performed, which is a limitation of this study. The sensors were connected via Modbus RTU (RS-485) to a Haiwell PLC, aggregated on a SCADA HMI, and published to a HiveMQ MQTT broker backed by InfluxDB, with Haiwell Cloud as a backup channel. The periodic fluctuations in source (4) reflect the PLC's own closed-loop control, independent of the DPIR-GSky analytics layer: the Haiwell PLC operates each zone's compressor via threshold-based (hysteresis) setpoint logic, engaging cooling once the zone temperature exceeds its upper setpoint and disengaging once the lower setpoint is reached, with the resulting on/off cycling visible in the raw temperature series used to calibrate the AR(1) shocks described above.

3.3. Acquisition and Characterization of Seven-Day Streaming Data

The streaming dataset was collected directly from 20 rack units over a full week at a one-minute resolution, resulting in 10,080 records per entity, using the temperature, RH, lux, ethylene, and CO_2 sensors of the MQTT/InfluxDB acquisition infrastructure described earlier. These data were used for descriptive statistics, trend and correlation analyses, multidimensional object formation, and scenario evaluation parameter calibration. Operational conditions during observation remained outside the optimal cold storage range, with an average temperature of approximately $14.6°C$ and ethylene concentration of approximately 15.7 ppm. These conditions saturated d_T and d_{Eth} at their maximum value of 1.0 and placed the RSL model outside its Q_{10} validation range. The streaming statistics and algorithm behavior can therefore be discussed as actual observed results, but absolute shelf-life estimates still require revalidation at the target $0 - 2°C$ range. Table 2 presents the descriptive statistics for the seven-day streaming data. Temperature and ethylene showed a mild but statistically significant upward trend over the week ($p < 0.05$), while illuminance remained flat. Temperature and ethylene were also strongly positively correlated ($r = 0.975$), whereas illuminance correlated weakly with the other variables ($|r| \leq 0.17$). All five sensor dimensions were nonetheless retained separately because they are causally distinct, and because G-Skyline operates on Pareto dominance rather than regression coefficients, multicollinearity does not affect it the way it would a regression-based method. A sensitivity check that excluded the ethylene dimension changed G-Skyline group membership by 8.3% on average, indicating partial rather than complete redundancy with temperature.

Table 2. Descriptive statistics of seven-day streaming data from the device

Variable	Unit	Mean	Std	Min	P25	Median	P75	Max
Temperature	°C	14.63	0.83	12.70	14.10	14.54	15.67	16.50
Relative humidity	%	89.86	1.82	83.55	88.90	89.97	90.96	94.46
Illuminance	lux	24.23	1.09	14.83	24.00	24.00	24.00	38.97
Ethylene	ppm	15.67	0.84	13.54	15.15	15.64	16.25	18.61
CO_2	ppm	97.66	6.24	62.46	91.96	97.85	102.68	120.75

3.4. AR (1) Stochastic Data Generation Model

Test scenarios were generated using a first-order autoregressive model, $AR(1)$, with parameters calibrated from the real seven-day dataset, given in Equation (2) to enable replication.

$$x_{i,d}(t) = \mu_d + \phi_d\left[x_{i,d}(t-1) - \mu_d\right] + \varepsilon_{i,d}(t), \qquad \varepsilon_{i,d}(t) \sim \mathcal{N}\left(0, \sigma_d^2\right) \tag{2}$$

$x_{i,d}$(t) is the value of variable d for entity i at time t, μ_d is its baseline level, set within the optimal range for scenario generation, φ_d is its autocorrelation coefficient, and $\varepsilon_{i,d}$(t) is Gaussian white noise representing shocks from compressor cycling and door openings, while the autoregressive term pulls the series back toward its baseline level. The calibrated autocorrelation coefficients (1-minute resolution) were 0.98 for temperature, RH, and ethylene, 0.90 for CO_2, and 0.58 for lux.

3.5. Stream Processing and Multidimensional Object Formation

In real-world implementation, sensor data flow through Modbus RTU → Haiwell SCADA → HiveMQ MQTT at one record per minute per entity, are stored in InfluxDB, and are aggregated by a window manager into a 60-minute sliding window (60 records/variable), from which a feature extractor computes a six-dimensional deviation vector per entity. From the seven-day dataset (10,080 records/entity) with a 30-minute evaluation step, 335 windows were obtained per entity, an example of which is shown in table 3.

A 60-min window was chosen to span multiple compressor duty cycles. This is consistent with the $AR(1)$ calibration: with $\varphi \approx 0.98/min$ for temperature, RH, and ethylene (Equation 2), the autocorrelation half-life is $\approx 34\,min$, so a 60-minute window spans just under two half-lives, long enough to avoid being dominated by a single transient event yet short enough to retain sensitivity to genuine drift.

Table 3. Example of sensor data transformation into a multidimensional object

Day	t	T	RH	Ethylene	CO_2	Lux	d_T	d_{RH}	d_{Eth}	d_{CO_2}	d_{Lux}	RSL	1/RSL
1	00:30	14.57	90.71	15.69	91.39	23.0	1.000	0.000	1.000	0.000	0.000	2.663	0.375
2	12:30	14.48	93.89	15.49	102.24	23.1	1.000	0.000	1.000	0.000	0.000	2.688	0.372

Feature aggregation in window h uses a mean operator. With $\Delta = 30\,minutes$ and $W = 60\,minutes$, 335 full windows are obtained, of which the first five were used for $AR(1)$ buffer warm-up, leaving 330 effective windows for comparative evaluation.

$$\tilde{x}_{j,q}(h) = \frac{1}{W}\sum_{\ell=0}^{W-1} x_{j,q}(t_h - \ell) \tag{3}$$

3.6. Five-Layer Architecture

The framework adopts a five-layer architecture adapted for vegetable cold storage. An overview of the architecture is illustrated in figure 3.

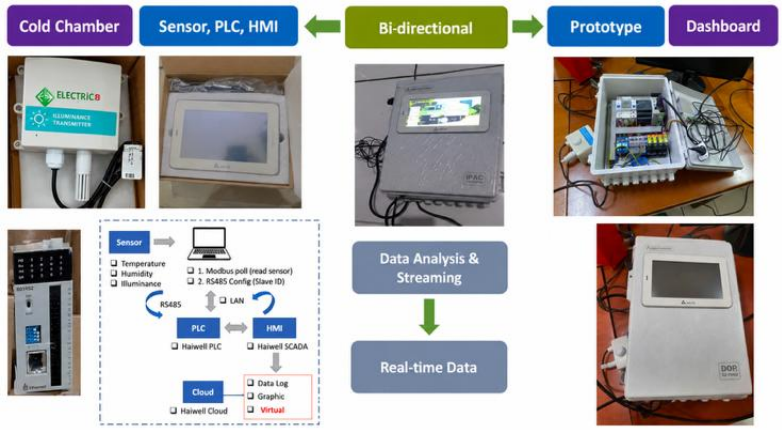


Figure 3. Five-layer digital twin architecture for multi-entity vegetable cold storage monitoring

The five layers span from physical sensing hardware to operator-facing decision outputs. Layer 1, Physical, comprises the cold storage infrastructure and an RS-485 sensor network (temperature, RH, lux, ethylene, and CO_2) read by a Haiwell PLC via Modbus RTU and visualized on the SCADA/HMI. In Layer 2, Communication, the HMI forwards the preprocessed data over LAN to a HiveMQ MQTT broker, with Haiwell Cloud as a backup. Layer 3, Virtual Twin Technology, stores the MQTT data in InfluxDB and exposes it via REST API, with each twin holding a state vector, a 24-hour history buffer, a degradation model, and an RSL predictor. Layer 4, Analytics and optimization, uses a stream ingestion engine, feature extractor, adaptive Z-score anomaly pre-filter, window manager, and dynamic optimization engine to convert the twin stream into a multidimensional G-Skyline representation. Layer 5, Decision, executes the G-Skyline on the active entity representations to generate non-dominated groups as intervention priorities, which are presented to the operator through a real-time dashboard, alarm notifications, and action recommendations.

Framed against Kritzinger et al.'s taxonomy [15], the framework supports automated bidirectional physical-control communication: layers 1–3 continuously update the virtual model from sensor readings, and the same PLC/Modbus RTU channel automatically issues control commands (e.g., the setpoint-based compressor cycling noted earlier). However, the multi-entity prioritization signal generated by DPIR-GSky in Layer 5 remains advisory and is presented to the operator as a ranked recommendation rather than being routed through the automated control channel. Therefore, the architecture is best described as a digital twin-oriented framework with automated infrastructure-level sensing and control combined with an advisory multi-entity decision layer for cross-entity prioritization.

3.7. Entity Representation and G-Skyline Dominance Definition

Each active storage entity j in window h is represented as a multidimensional point vector in a six-dimensional decision space (Equation (4)). All dimensions were expressed as urgency indices with the same orientation, meaning that the higher the value, the higher the intervention priority was.

$$o_{j,h} = (d_{j,T}(h), d_{j,RH}(h), d_{j,C_2H_4}(h), d_{j,CO_2}(h), d_{j,Lux}(h), u_{j,RSL}(h)\ ) \tag{4}$$

The first five components were temperature deviation, RH, ethylene, CO_2, and light intensity. For a variable q with the optimal range $[L_q, U_q]$, the bounded deviation sub-index is defined in Equation (5). The operational alert thresholds L_q and U_q applied in Equation (5) represent the practical monitoring bounds of the facility (e.g., $L_q(RH) = 85\%$), which may differ from the ideal postharvest recommendation ($95 - 100\%$). The deviation sub-index reaches 1.0 only when the window value lies at least one range width outside the operational range.

$$d_{j,q}(h) = min\left\{1, \left[\frac{max\{0, L_q - \tilde{x}_{j,q}(h), \tilde{x}_{j,q}(h) - U_q\}}{U_q - L_q}\right]^2\right\} \tag{5}$$

This formulation is equal to 0 when the variable is within the operational range and increases quadratically as the window value moves away from it. For variables with a single upper limit, such as ethylene, CO_2, and lux, the L_q value is taken as the physical lower bound or minimum operational limit and U_q as the recommended maximum limit. The sixth component is the urgency of shelf life, normalized from the remaining shelf life (RSL) (Equation 6).

$$u_{j,RSL}(h) = min\left\{1, \frac{RSL_{crit}}{RSL_j(h)}\right\} \tag{6}$$

In this study, $RSL_{crit} = 1\ day$ so that entities with remaining shelf life of less than one day are assigned maximum urgency. Table 3 presents the value of $1/RSL$ as a numerically equivalent form before the truncation at 1. RSL is modeled using a modified Q_{10} model that accounts for the effects of temperature and ethylene (Equation 7).

$$RSL_j(h) = \frac{RSL_0}{A_T(h)(1 + \alpha d_{j,T}(h))(1 + \beta d_{j,C_2H_4}(h))}, \qquad A_T(h) = Q_{10}^{(\tilde{x}_{j,T}(h) - T_{ref})/10} \tag{7}$$

with $RSL0$ as the reference shelf life under optimal conditions (21 days for lettuce at $1°C$ and 98% RH), $A_T(h)$ is the acceleration factor based on $Q_{10}(Q_{10} = 2.8; T_{ref} = 1°C)$, and $\alpha = 0.3$ and $\beta = 0.5$ are deviation contribution parameters. This RSL formula is bivariate as it includes only temperature (via A_T) and ethylene (via d_j, C_2H_4), the two variables with established kinetic effects in the postharvest literature [28], [30]. RH deviation, CO_2, and lux continued to contribute to prioritization through their G-Skyline vector dimensions, without being included in the shelf-life model. This restriction is an intentional scoping choice, not a complete multivariable Q_{10} model.

Extending this RSL model with humidity-deficit and cumulative light-exposure terms requires commodity-specific calibration against destructive postharvest measurements and is left for future work.

Definition of dominance. Since all dimensions of the urgency vector are expressed as a degree of deviation that becomes more critical as it increases, point p dominates point q (written $p \succ q$) if and only if p is not lower than q on any dimension and strictly higher on at least one dimension.

 ISSN 2723-6471

For a G-Skyline of size k, group G_1 g-dominates group G_2 if there exists a bijection between their members such that each member of G_1 is paired with exactly one member of G_2 that it dominates or equals, with at least one domination pair strict.

A G-Skyline of size k is defined as the set of all groups that are not g-dominated by other groups in the same window. A group size of $k = 3$ was chosen to represent the intervention capacity of the operator team (three racks handled simultaneously in one session), assumed for experimental purposes rather than derived from a documented facility staffing SOP.

3.8. DPIR-GSky algorithm

G-Skyline processing on streaming data follows an incremental workflow that uses the properties of the skyline layers; all members of a size-k G-Skyline group lie within the first k skyline layers [26]. Algorithm 1 presents the DPIR-GSky pseudocode. Its three main components are: (i) an adaptive pruning threshold, computed from the mean and standard deviation of the deviation scores of all active entities in the current window (Algorithm 1, lines 2–3), that eliminates unlikely candidates before the expensive dominance test; (ii) incremental group reuse, so that only entities whose deviation vectors change by more than a significant-change threshold trigger re-evaluation; and (iii) dominance-power scoring, which ranks returned G-Skyline groups by the average normalized urgency of their member entities, so groups with more urgent members rank higher. This scoring applies uniform implicit weighting to each dimension. The claim of "no weighting" refers to the lack of subjectively assigned criterion weights by the operator, not to the absence of implicit numeric weights.

An entity is retained as a pruning candidate if its deviation score $\rho_e(h)$, the mean of its six urgency components, is at least the adaptive threshold θ_h. Each candidate group was then ranked by its normalized dominance power score, which was computed as the mean urgency of its members. The dominance power score falls within $[0, 1]$ because all components are confined to the same normalized scale. Higher values indicate groups with greater urgency of need.

Algorithm 1. Dominance-Power-based Incremental Refinement G-Skyline (DPIR-GSky).

Input: Sequence S of deviation vectors $\{v_e(t) \mid e \in E, t \in [t_w - W, t_w]\}$; window size $W = 60\ min$; step $\Delta = 30\ min$; change threshold $\delta = 0.05$; minimum pruning threshold = 0.10
Output: G-Skyline group ranked as G^* along with dominance-power score $P(g, h)$

1: FOR each new window $[t_w - W, t_w]$ DO
2: Calculate μ_h and σ_h using all components $o_{e,h}$ in the active window
3: Set adaptive threshold $\theta_h = max(\theta_{min}, \mu_h - \sigma_h)$
4: FOR each entity $e \in E$ DO
5: IF $\|v_e(t_w) - v_e(t_w - \Delta)\|\infty > \delta$ THEN e marked as CHANGED; update v_e in buffer B
6: END FOR
7: $E_{changed} = \{e \in E: e\ marked\ as\ CHANGED\}$
8: FOR each $e \in E_{changed}$ DO
9: Delete e from all groups on G_{prev}
10: FOR each candidate group $g \in G_{prev} \cup \{\emptyset\}$ DO
11: IF $\forall e' \in g$: $\neg dominates(e', e)\ AND\ \rho_e(h) \geq \theta_h$ THEN
12: Add e to g; update $P(g, h) = \frac{1}{kd} \sum_{e \in g} \sum_{m=1}^{d} o_{e,m}(h)$
13: END FOR
14: END FOR
15: $G^* \leftarrow$ SortDescending($G_{prev}, key = P(g, h)$)
16: Output G^* with $P(g, h)$
17: END FOR

The use of a significant change threshold (δ) prevents unnecessary recomputation when marginal entity conditions change. Letting n denote the number of active entities in a window (distinct from the window length W), each update requires $O(n \cdot d)$ for deviation calculation and $O(n^2 \cdot k)$ for the pairwise g-dominance tests of incremental group refinement, where $d = 6$ and k is the target group size; exhaustive enumeration, as in Naive-GSky, instead faces a worst-case candidate space of $O(n^k)$.

ISSN 2723-6471

3.9. Simulation Parameters and Test Scenarios

The simulation was run for 72 hours at 5-minute intervals per sensor (864 points per variable per entity). Disturbance rates below are specified per minute and applied at each 5-minute step as $C(t + \Delta t) = C(t) + rate \cdot \Delta t$ with $\Delta t = 5\ min$. The main scenario evaluation uses 20 racks × 6 deviation dimensions, while the supporting benchmark code also tests scalability at 4, 8, 12, 16, and 20 entities as well as variations in group size $k = 2$ and $k = 3$. The AR(1) dynamics (φ_d and noise variance) were calibrated on the seven-day dataset (Table 2), while baseline levels were set within the optimal ranges so that S1 represents a stable optimal condition. Five operational scenarios were designed to represent increasingly complex disturbance conditions. The first scenario (S1, stable optimal condition) held all variables within range for $72\ h$ without any disturbance, verifying the absence of false positives. Building on this baseline, the second scenario (S2, localized temperature disturbance) warmed two racks at $+0.05°C/min$ from hour 12 until they were $> 3°C$ above the optimum, testing early detection before the absolute threshold is breached. The third scenario (S3, multi-zone RH fluctuation) extended the disturbance to three zones whose RH declined at different rates ($-0.1\ to\ -0.3\%/min$) from different start times, testing prioritization among entities that degrade concurrently but with different severity. Unlike these temperature and humidity disturbances, the fourth scenario (S4, external ethylene spike) introduced an ethylene surge of $+0.5$ ppm/min into one zone from hour 24, which then spread to adjacent zones, testing the detection of an external-source anomaly that does not immediately violate absolute thresholds. Finally, the fifth scenario (S5, simultaneous compound disturbances) combined concurrent deviations in temperature (R05), RH (R09), and ethylene (R14) with heterogeneous magnitudes and timings, thereby testing the core utility of G-Skyline when several entities require intervention with differing profiles.

3.10. Ground Truth Construction and Evaluation Protocol

Ground-truth urgency labels were constructed independently of DPIR-GSky, TOPSIS, and PROMETHEE II, using only each scenario's injected disturbance profile: an entity is labeled urgent once its injected variable crosses the same operational threshold used for the deviation sub-indices (e.g., R11 once its injected ethylene exceeds threshold in S4). Because this rule never references any method's output, all four methods are scored against the same independently defined label, avoiding circularity. DPIR-GSky's deviation dimensions (Equation 5) use the same operational thresholds, so any feature–label correlation affects all four methods alike and none is uniquely favored. This symmetry is not a substitute for validation against an independent physical quality outcome, which remains for future work.

3.11. Performance and Temporal Evaluation Metrics

Functional performance was evaluated using standard formulas for detection latency (the time difference between when degradation occurs and when the system identifies the entity as a priority candidate), virtual twin accuracy (R^2, the coefficient of determination between each sensor variable's $AR(1)$-predicted and observed value, averaged across variables), Precision, Recall, and F1-score for G-Skyline prioritization, and ranking stability via the coefficient of variation (CV) of an entity's position rank across successive stream updates.

End-to-end latency covers the interval from the sensor sample arrival to the priority group availability, and the response time covers the interval from a significant change trigger to the output update. Jitter is the standard deviation of the per-window computation time, and the throughput is the number of tuples processed per second. WCET is the maximum observed execution time, which is empirically estimated in this study.

The communication layer combines time-bounded deterministic Modbus RTU polling with the best-effort MQTT QoS 1 delivery. Functional validation showed a constant 60-second arrival interval with no dropout or backpressure across 10,080 clean records per entity.

The four baselines differ in their mechanisms. Naive-GSky tests group combinations exhaustively, MDS presorts by dominance power, GPR reuses groups from the previous window, and SW-GSky updates incrementally but without dominance-power ranking or adaptive pruning. DPIR-GSky's incremental group reuse (re-evaluating only changed entities) and adaptive pruning threshold (eliminating 69.7% of candidates on average) together make its runtime growth more gradual than Naive-GSky's empirically near-quadratic growth as entity count increases over the tested range (Figure 4), although complexity still grows polynomially, motivating spatial partitioning at a very large scale.

For larger facilities, zone-based spatial partitioning can further bound candidate-group formation by restricting comparisons to operationally meaningful, physically co-located zones: partitioning n entities into P zones of size n/P reduces the worst-case candidate space from $O(n^k)$ to $O(n^k/P^{(k-1)})$, a reduction by a factor of $P^{(k-1)}$ for a fixed group size k.

4. Results and Discussion

4.1. Digital Twin Representation Accuracy

Virtual twin accuracy averaged $R^2 = 0.258$ across scenarios ($0.239 - 0.275$), with the lowest R^2 for CO_2 (≈ 0.23) and the highest for temperature (≈ 0.27). RMSE values (temperature $0.71°C$, RH 1.54%, ethylene $0.71\ ppm$, CO_2 5.47 ppm, lux 0.98 lx) were of the same order as each signal's own standard deviation (Table 2). The low R^2 is partly structural: R^2 is computed on the raw sensor series, whose variability under stable conditions is small (Table 2), so most remaining variation is short-term fluctuation from compressor cycling, door openings, and sensor noise that a first-order autoregressive model cannot anticipate. The current twin configuration therefore supports directional trend monitoring but not accurate point-level prediction, a limitation for applications that need point forecasts.

Improving twin fidelity under stable conditions requires adaptive state-space filtering, such as an extended Kalman filter (EKF), an autoregressive LSTM/GRU network, or a kinetics-constrained physics-informed neural network (PINN), to separate measurement noise from the true degradation trend, which is a priority for future work.

4.2. Degradation Detection Performance

A comparison of the detection latencies of the proposed framework and static threshold system is presented in table 4.

Table 4. Detection latency of degradation per scenario (min).

Scenario	Static Threshold (min)	Proposed Framework (min)	Difference (min)
S2 – Temperature disturbance	0.0	0.0	0.0
S3 – RH Fluctuation	0.0	0.0	0.0
S4 – Ethylene surge	65.0	0.0	65.0
S5 – Compound disturbance	0.0	0.0	0.0
Average*	16.2	0.0	16.2

**The average was driven entirely by the S4 data. No latency advantage was observed in S2, S3, or S5.*

The detection advantage is specific to S4 (ethylene surge), where the proposed framework detected the disturbance 65.0 minutes earlier, before the absolute ethylene threshold was crossed. The static-threshold baseline checks every raw 5-minute sample, whereas DPIR-GSky is re-evaluated at each 30-minute step; a flag at the first DPIR-GSky step after onset is recorded as 0.0 min by convention, not literal instantaneity, so the 0.0-minute readings in S2, S3, and S5 mean that both methods flagged the disturbance within the first evaluation step.

The injected ethylene surge in S4 began at hour 24 and grew at a constant rate of 0.5 ppm per minute, and the static-threshold system only flags a rack once the raw ethylene reading crosses its fixed operational alert threshold, which, like the other operational bounds, lies well above the ideal postharvest limit. A 65-min lag (13 samples) at that rate means that the raw signal had to climb by 0.5 ppm/min × 65 min = 32.5 ppm above the pre-disturbance level before the static system triggered. DPIR-GSky, in contrast, flagged the affected rack at the first 30-min evaluation step after onset (recorded as 0.0 minutes in Table 4 under the convention explained above), because G-Skyline membership reflects each rack's urgency relative to the other entities rather than the crossing of a single absolute alarm; isolating which vector component drove this early flag requires a component-ablation analysis, which is left for future work.

4.3. Effectiveness of G-Skyline in Prioritization

The effectiveness of G-Skyline was evaluated by comparing the member entities of the G-Skyline group ($k = 3$) to the most urgent entities according to the scenario ground truth, and by comparison with TOPSIS and PROMETHEE II (table 5). For the main classification metrics, G-Skyline (F1 = 0.685) trails TOPSIS and PROMETHEE II (F1 = 0.727) by 0.042 points. Because both baselines also use uniform weights, this gap reflects the difference between compensatory aggregation into a single score and Pareto-based group selection, not an advantage of prior weight

knowledge. The deficit concentrates in single-dimension disturbances, such as S4 (G-Skyline F1 = 0.262 vs. TOPSIS 0.727), but disappears under the multi-dimension, heterogeneous disturbance S5, where G-Skyline scored higher (0.793) than both weighted baselines.

Table 5. Comparison of the effectiveness of prioritization methods (mean S2–S5)

Methods	Precision	Recall	F1-Score
Static Threshold	0.685	0.787	0.726
TOPSIS (uniform weights)	0.686	0.789	0.727
PROMETHEE II	0.686	0.789	0.727
G-Skyline (Proposed)	0.642	0.748	0.685

TOPSIS and PROMETHEE II report identical Precision, Recall, and F1 to three decimal places in Table 5. With only 20 entities, uniform criterion weights, and the same six normalized deviation dimensions, the two methods reached equivalent classification outcomes despite computing different intermediate scores (closeness coefficient vs. net outranking flow). Whether their underlying per-window rankings are also identical is left for future rank correlation checks. Ranking stability, measured as the coefficient of variation (CV) of entity positions across stream updates, averaged 0.498 for G-Skyline (S2–S5, highest stability in S5 at 0.461) versus 0.231 for TOPSIS/PROMETHEE II, indicating greater positional variation under the evaluated conditions. Isolating the source of this variation, whether a genuine condition change or single-variable noise, requires a dedicated perturbation-sensitivity experiment.

4.4. Soft Real-Time Characterization

The proposed system is classified as a soft real-time application: the operational deadline is the 30-minute evaluation interval (Δ), and missing it reduces the currency of priority information without causing catastrophic failure. Measured from the streaming benchmark (Table 6), DPIR-GSky operated well within this deadline: an average execution time of 0.32 ms/window, throughput of 37,734 tuples/s, an empirical worst-case execution time (WCET) below 1 ms, peak memory of 8.42 KB, and 93.0% window coverage, so the 30-minute deadline is roughly 5.6×10^6 times the average runtime. The communication results were consistent with this characterization: Modbus RTU bounded field-level acquisition, and MQTT was the main source of temporal uncertainty. The absence of dropout or backpressure indicates that the tested LAN was adequate for this data load under normal conditions, although jitter injection, packet loss, and broker failover under adverse conditions remain to be investigated.

4.5. Benchmarking Algorithms and Scalability

Table 6 presents the benchmark of DPIR-GSky against four G-Skyline baselines on a synthetic streaming workload calibrated from the seven-day real dataset. This algorithmic benchmark uses a synthetic stream (windows of 12 tuples, step 6) distinct from the 60-minute operational window, isolating the per-window computational cost under controlled, repeatable conditions. The quality score is the average normalized deviation between the dominance power of the returned G-Skyline groups and that of the exact (brute-force) baseline groups for the same window (lower is better). A lower value therefore indicates closer agreement with the exact solution, not lower urgency.

Here, a "tuple" is one per-entity deviation vector fed to the dominance-comparison engine, and the "window"/"step" counts in Table 6 (12 tuples, step 6) are synthetic workloads rather than the 60-minute/30-minute temporal window; Figure 4 uses the same setting while varying the entity count from 4 to 20. These figures therefore characterize algorithmic behavior under controlled load rather than the literal operational cadence, which is addressed as future work.

Table 6. G-Skyline algorithm benchmark on a synthetic streaming workload calibrated from the seven-day real dataset

Comparison Dimension	Naive-GSky	MDS	GPR	SW-GSky	DPIR-GSky (Proposed)
Streaming support	✗ (batch/window)	✗ (batch/window)	Partial (reuse)	✓ (incremental)	✓ (incremental)
Dominance-power ranking	✗	✓	✗	✗	✓ (with reuse)
Adaptive pruning	✗	✗	✗	✗	✓ (69.7%)
Average runtime (ms)	0.61	0.44	0.38	0.31	0.32
Throughput (tps)	19,565	26,998	31,562	38,408	37,734

ISSN 2723-6471

Memory (KB)	3.33	4.33	6.11	6.06	8.42
Quality Score (↓)	0.374	0.374	0.374	0.386	0.386
Window coverage (%)	100.0	100.0	100.0	93.0	93.0
Speedup vs Naive	1.00×	1.39×	1.61×	1.97×	1.91×

Relative to Naive-GSky, DPIR-GSky reduced the average runtime from 0.61 to 0.32 ms (1.91×) and raised throughput from 19,565 to 37,734 tps at 93.0% window coverage. Because SW-GSky, which lacks adaptive pruning, achieves a similar speedup (1.97×), most of this gain is attributable to incremental re-evaluation of changed entities rather than to pruning. The remaining 7.0% are windows skipped because no entity changed enough to warrant re-evaluation, not failed evaluations. The higher Quality Score of 0.386 vs. 0.374 for Naive-GSky (lower is better) is shared with SW-GSky and coincides with the 7.0% of skipped windows, a modest trade-off for the 1.91× speedup.

SW-GSky is marginally faster (1.97× vs. 1.91×) and lighter (6.06 vs. 8.42 KB), with the same Quality Score and coverage. What DPIR-GSky adds over SW-GSky is dominance-power ranking, which returns a ranked list of groups rather than an unordered skyline set, at the cost of extra per-group state.

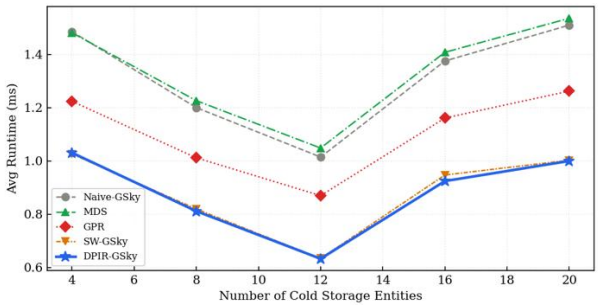


Figure 4. Average runtime (ms) versus the number of cold storage entities (window = 12, $k = 2$): DPIR-GSky increases more gradually than Naive-GSky.

In terms of scalability (figure 4), the observed runtime of Naive-GSky grows approximately quadratically with the entity count over the tested range, whereas DPIR-GSky increases more gradually through incremental reuse and adaptive pruning, supporting its viability for soft real-time operation within a 30-minute evaluation step.

4.6. G-Skyline Dominance Analysis on 20 Entities

In a representative 20-entity evaluation window, the G-Skyline computation at $k = 3$ consistently prioritized rack R02 across the returned intervention groups, reflecting its highest $1/RSL$ value (shortest remaining shelf life, 62.8 h) among the active entities.

4.7. Discussion

The results support three points that link the empirical findings to the research gap. On integration maturity, the sensing and actuation backbone (layers 1–3 and the PLC/Modbus channel) reaches the Digital Twin tier of Kritzinger et al.'s taxonomy [15], whereas the DPIR-GSky output reaches the operator as an advisory ranked recommendation; the framework as a whole is therefore digital twin-oriented, pairing automated sensing and control with a human-in-the-loop decision layer [33], [34]. A modest coefficient of determination ($R^2 = 0.258$) does not undermine the framework's contribution: G-Skyline prioritization runs on window descriptive statistics (Equation 3) rather than the autoregressive forecaster's output; therefore, weak twin fidelity does not carry through to the F1-score, ranking stability, or detection latency results. The low R^2 instead reflects the operating regime: under stable conditions, the raw signals carry little variance beyond short-term fluctuations that an autoregressive persistence model cannot anticipate, so R^2 mainly reflects how little explainable variance the stable regime offers rather than a failure to track the signals [35], [36]. The current twin therefore works as a trend-responsive monitor; accurate point forecasting is a separate problem for future work.

The comparison with the uniformly weighted TOPSIS and PROMETHEE II (table 5) is the most direct test of the framework's central claim. The F1 deficit in S4 is consistent with how the methods aggregate: in a compensatory score, a large deviation in one dimension directly raises an entity's overall rank [37], [38], [39], whereas DPIR-GSky scores

groups by mean urgency over six dimensions, so a single-dimension shock competes with entities that are moderately elevated across several dimensions. This disadvantage disappears under S5, where temperature, relative humidity, and ethylene deviate together with different timings and magnitudes: DPIR-GSky posts the highest F1 of any method (0.793) and its own most stable ranking across scenarios (CV = 0.461 against its S2–S5 average of 0.498). S5 also matches the condition for which the framework was designed (several variables moving at once, with no single weighting scheme trustworthy in advance), making it the result most relevant to the study's aim, although a single scenario cannot establish a general property of weight-free ranking. The resulting testable hypothesis is that DPIR-GSky's disadvantage in single-variable disturbances turns into an advantage once disturbances are compound and heterogeneous, since it never depends on the disturbed dimension being the one an analyst happens to weight highly [26]. Testing this claim against more heterogeneous scenarios, with a formal significance test and rank-specific metrics such as Precision@k or NDCG [40], remains for future work. Because all four methods draw on the same threshold-derived deviation dimensions used to build the ground-truth labels, none is uniquely favored by the labeling, although this symmetry is not validation against an independent measure of physical quality [28], [30]. Overall, these results do not show DPIR-GSky beating weighted MCDM baselines across the board; they show a computationally and functionally workable framework whose advantage concentrates where a fixed weighting scheme is hardest to defend in advance.

4.8. Limitations and Future Work

The seven-day streaming dataset used for calibration and benchmarking was collected at a mean temperature of approximately 14.6°C, well outside the 0–2°C range recommended for leafy vegetables. This saturated the temperature and ethylene deviation sub-indices and pushed ranking discrimination onto relative humidity, carbon dioxide, and the shelf-life index instead. The RSL model itself remains bivariate, covering only temperature and ethylene, and a preliminary check found that dropping the ethylene dimension altered G-Skyline group membership by 8.3% on average, indicating only partial redundancy with the other dimensions. Twin accuracy and the F1-score gap against weighted baselines were each measured at a single operating point, not swept across window sizes, pruning thresholds, or group sizes. Validation was limited to one facility, one commodity (curly lettuce), and simulated rather than field-observed disturbances, and the autoregressive scenario generator does not preserve the strong temperature–ethylene correlation ($r = 0.975$) seen in the real data.

Four items are immediate priorities for future work. First, empirical validation at the target 0–2°C storage range: without it, the deviation sub-indices and RSL estimates are model-derived urgency indicators, not validated shelf-life measures. Second, an independent physical ground truth, using destructive quality measurements such as weight loss, chlorophyll content, and texture, to check the prioritization output against actual deterioration rather than threshold-crossing alone. Third, additional heterogeneous disturbance scenarios, a formal significance test, and rank-specific metrics to confirm or rule out the S5 advantage described above. Fourth, an operational-cadence benchmark that replays the literal 60-minute/30-minute window configuration, to check whether the runtime and throughput figures in Table 6 hold beyond the entity-count micro-benchmark used here. Secondary extensions include a multivariate RSL model with uncertainty propagation, adaptive state-space filtering for twin point-accuracy, sensitivity sweeps over the pruning threshold, window length, and k, a correlation-preserving scenario generator, network stress testing, cross-commodity validation, formal WCET analysis on an RTOS, and an interventional study linking algorithmic gains to measured spoilage reduction. Toward closed-loop operation, the next step is routing DPIR-GSky's output through the same PLC channel for a narrow set of reversible actions, with safety interlocks, authorization logic, and fallback behavior yet to be designed and tested.

5. Conclusion

Existing digital twin implementations for cold storage can monitor or forecast individual entities, but none resolves the online problem this study set out to address: ranking several simultaneously deviating entities for intervention without criterion weights fixed in advance. DPIR-GSky addresses this gap by pairing a streaming digital twin architecture with an incremental, dominance-power-based group skyline engine that requires no manually elicited criterion weights. The results show that the framework is computationally efficient, at 0.32 ms per window, with an empirical worst-case

execution time below 1 ms and throughput above 37,000 tuples per second, and functionally capable of soft real-time multi-entity prioritization.

Classification performance against uniformly weighted TOPSIS and PROMETHEE II averages lower for DPIR-GSky (F1 = 0.685 vs. 0.727), yet it matches or exceeds both baselines under the heterogeneous, multi-variable disturbance scenario (S5) the framework targets, and detects an ethylene-related disturbance 65 minutes earlier than static-threshold monitoring. Weighted methods therefore remain preferable when criterion importance is stable and known; DPIR-GSky is most useful where a fixed, defensible weighting scheme is hardest to specify in advance.

This is a feasibility result rather than a deployment-ready one: virtual twin accuracy was modest ($R^2 = 0.258$) under stable operating conditions, the prioritization ground truth was built from the same thresholds used to construct the deviation vectors, and validation was confined to one facility, one commodity, and simulated disturbances. Operational deployment still requires validation at the target 0–2°C storage range, an independent physical ground truth, and multi-facility field trials.

6. Declarations

6.1. Author Contributions

Conceptualization: T.A.C. and T.D.; Methodology: T.A.C., T.D., and A.; Software: T.A.C.; Validation: T.A.C., T.D., A. and H.S.; Formal analysis: T.A.C. and P.F.F.; Investigation: T.A.C. and H.R.; Resources: T.D. and A.P.H.; Data curation: T.A.C.; Writing – original draft: T.A.C.; Writing – review and editing: T.D., A., H.S., and H.R.; Visualization: T.A.C. and P.F.F. All authors have read and approved the final version of this manuscript.

6.2. Data Availability Statement

The data presented in this study are available upon request from the corresponding author.

6.3. Funding

This research was supported by a Dissertation Grant from Universitas Muhammadiyah Jember based on the Rector's Decree No. 768/11.3.AU/REKTORAT/A/2021.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

[1] Food and Agriculture Organization of the United Nations, *The state of food and agriculture 2019: Moving forward on food loss and waste reduction*. Rome, Italy: Food and Agriculture Organization of the United Nations, vol. 2019, no. 1, pp. 1–182, 2019, doi: 10.18356/32f21f8c-en.

[2] P. M. A. Toivonen and D. A. Brummell, "Biochemical bases of appearance and texture changes in fresh-cut fruit and vegetables," *Postharvest Biol. Technol.*, vol. 48, no. 1, pp. 1–14, Apr. 2008, doi: 10.1016/j.postharvbio.2007.09.004.

[3] T. G. Hoffmann, C. K. de Souza, A. D. Sonawane, U. Praeger, F. Büchele, D. A. Neuwald, R. Jedermann, M. Linke, B. Sturm, and P. V. Mahajan, "Challenges and future perspectives in postharvest cold storage: Technological review on sustainable and efficient cold storage of fresh produce," *Food Res. Int.*, vol. 221, no. 1, pp. 1–15, Dec. 2025, doi: 10.1016/j.foodres.2025.117406.

[4] R. D. Jawanjal, P. G. Shivarkar, R. R. Ghatale, and A. O. Sonone, "Cold storage and management system," *Int. Res. J. Adv. Eng. Manage.*, vol. 3, no. 4, pp. 1261–1267, Apr. 2025, doi: 10.47392/IRJAEM.2025.0206.

[5] D. Akhiyar, "Development of color segmentation and texture analysis algorithms for early detection of green vegetable deterioration in retail environments," *J. Appl. Data Sci.*, vol. 7, no. 2, pp. 882–896, May 2026, doi: 10.47738/jads.v7i2.1094.

[6] M. Wei, J. Wang, Z. Ma, and M. Zhu, "An introduction to normal-operating-zones for process monitoring," *IEEE Int. Conf. Ind. Technol.*, vol. 2025, no. Mar., pp. 1–6, Mar. 2025, doi: 10.1109/ICIT63637.2025.10965223.

[7] W. Li, Y. Xia, and W. Ren, "A novel robust anomaly detection for enhancing reliability of industrial equipment under variable operating conditions via multi-sensor information fusion," *Eng. Res. Express*, vol. 8, no. 2, pp. 1–10, Jan. 2026, doi: 10.1088/2631-8695/ae37cf.

[8] H. Darvishi, D. Ciuonzo, E. R. Eide, and P. S. Rossi, "Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture," *IEEE Sens. J.*, vol. 21, no. 4, pp. 4827–4838, Feb. 2021, doi: 10.1109/JSEN.2020.3029459.

[9] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary Perspectives on Complex Systems*, F.-J. Kahlen, S. Flumerfelt, and A. Alves, Eds. Cham, Switzerland: Springer, vol. 2017, no. 1, pp. 85–113, 2017, doi: 10.1007/978-3-319-38756-7_4.

[10] A. Stefanone, M. Sala, R. Zanetti, M. Pirotta, M. Rossoni, and G. Colombo, "Real-time monitoring platform for commercial refrigerators using digital twin aggregates," *ASME Int. Mech. Eng. Congr. Expo.*, vol. 2025, no. Nov., pp. 1–10, Nov. 2025, doi: 10.1115/IMECE2025-166442.

[11] H. Nozari, "Digital twin-based predictive maintenance in cold chain logistics," *RAIRO Oper. Res.*, vol. 60, no. 1, pp. 481–499, Jan. 2026, doi: 10.1051/ro/2025167.

[12] D. Yi, "Transforming agriculture: An insight into decision support systems in precision farming," *J. Appl. Data Sci.*, vol. 5, no. 4, pp. 1614–1624, Dec. 2024, doi: 10.47738/jads.v5i4.274.

[13] S. Efendi, "An IoT-enabled smart system utilizing linear regression for sheep growth and health monitoring," *J. Appl. Data Sci.*, vol. 6, no. 3, pp. 2056–2071, Sep. 2025, doi: 10.47738/jads.v6i3.901.

[14] M. W. Grieves, "Product lifecycle management: The new paradigm for enterprises," *Int. J. Product Develop.*, vol. 2, no. 1/2, pp. 71–84, 2005, doi: 10.1504/IJPD.2005.006669.

[15] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018, doi: 10.1016/j.ifacol.2018.08.474.

[16] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *Int. J. Adv. Manuf. Technol.*, vol. 94, no. 9–12, pp. 3563–3576, Feb. 2018, doi: 10.1007/s00170-017-0233-1.

[17] C. Verdouw, B. Tekinerdogan, A. Beulens, and S. Wolfert, "Digital twins in smart farming," *Agric. Syst.*, vol. 189, no. 1, pp. 1–12, Apr. 2021, doi: 10.1016/j.agsy.2020.103046.

[18] R. G. Alves, G. Souza, R. F. Maia, A. L. H. Tran, C. Kamienski, J.-P. Soininen, P. T. Aquino, and F. Lima, "A digital twin for smart farming," *IEEE Global Humanitarian Technol. Conf.*, vol. 2019, no. Oct., pp. 1–4, Oct. 2019, doi: 10.1109/GHTC46095.2019.9033075.

[19] Y. Zou, J. Wu, X. Meng, X. Wang, and A. Manzardo, "Digital twin integration for dynamic quality loss control in fruit supply chains," *J. Food Eng.*, vol. 397, no. 1, pp. 1–12, Sep. 2025, doi: 10.1016/j.jfoodeng.2025.112577.

[20] T. Elsaleh, S. Enshaeifar, R. Rezvani, S. T. Acton, V. Janeiko, and M. Bermudez-Edo, "IoT-Stream: A lightweight ontology for Internet of Things data streams and its use with data analytics and event detection services," *Sensors*, vol. 20, no. 4, pp. 1–18, Feb. 2020, doi: 10.3390/s20040953.

[21] M. Stonebraker, U. Çetintemel, and S. Zdonik, "The 8 requirements of real-time stream processing," *ACM SIGMOD Rec.*, vol. 34, no. 4, pp. 42–47, Dec. 2005, doi: 10.1145/1107499.1107504.

[22] B. F. Arkana, "An adaptive random forest for data stream sentiment classification under concept drift," *J. Appl. Data Sci.*, vol. 7, no. 2, pp. 926–939, May 2026, doi: 10.47738/jads.v7i2.1153.

[23] Z. Xie, H. Long, C. Ling, Y. Zhou, and Y. Luo, "An anomaly detection scheme for data stream in cold chain logistics," *PLoS One*, vol. 20, no. 3, pp. 1–15, Mar. 2025, doi: 10.1371/journal.pone.0315322.

[24] H. Sundmaeker, C. Verdouw, S. Wolfert, and L. P. Freire, "Internet of Food and Farm 2020," *Digitising the Industry Internet of Things Connecting the Physical, Digital and Virtual Worlds*, vol. 2022, no. 1, pp. 129–151, Jan. 2022, doi: 10.1201/9781003337966-4.

[25] S. Borzsonyi, D. Kossmann, and K. Stocker, "The skyline operator," *Proc. Int. Conf. Data Eng.*, vol. 2001, no. Apr., pp. 421–430, Apr. 2001, doi: 10.1109/ICDE.2001.914855.

[26] J. Liu, L. Xiong, J. Pei, J. Luo, and H. Zhang, "Finding Pareto optimal groups: Group-based skyline," *Proc. VLDB Endow.*, vol. 8, no. 13, pp. 2086–2097, Sep. 2015, doi: 10.14778/2831360.2831363.

[27] R. E. Paull, "Effect of temperature and relative humidity on fresh commodity quality," *Postharvest Biol. Technol.*, vol. 15, no. 3, pp. 263–277, Mar. 1999, doi: 10.1016/S0925-5214(98)00090-8.

[28] M. A. J. S. van Boekel, "Kinetic modeling of food quality: A critical review," *Compr. Rev. Food Sci. Food Saf.*, vol. 7, no. 1, pp. 144–158, Jan. 2008, doi: 10.1111/j.1541-4337.2007.00036.x.

[29] M. S. S. Putri, "A data-driven mixed integer nonlinear programming model for cost-optimal scheduling of perishable production and workforce," *J. Appl. Data Sci.*, vol. 6, no. 4, pp. 2449–2460, Dec. 2025, doi: 10.47738/jads.v6i4.1019.

[30] H. Manolopoulou, G. R. Lambrinos, E. Chatzis, G. Xanthopoulos, and E. Aravantinos, "Effect of temperature and modified atmosphere packaging on storage quality of fresh-cut romaine lettuce," *J. Food Qual.*, vol. 33, no. S1, pp. 317–336, Sep. 2010, doi: 10.1111/j.1745-4557.2010.00321.x.

[31] G. Haseli, R. Sheikh, J. Wang, H. Tomaskova, and E. B. Tirkolaee, "A novel approach for group decision making based on the best–worst method (G-BWM): Application to supply chain management," *Mathematics*, vol. 9, no. 16, pp. 1–17, Aug. 2021, doi: 10.3390/math9161881.

[32] D. Sharma, S. Sridhar, and D. Claudio, "Comparison of AHP-TOPSIS and AHP-AHP methods in multi-criteria decision-making problems," *Int. J. Ind. Syst. Eng.*, vol. 34, no. 2, pp. 203–223, 2020, doi: 10.1504/IJISE.2020.105291.

[33] T. Defraeye, C. Shrivastava, T. Berry, P. Verboven, D. Onwude, S. Schudel, A. Bühlmann, P. Cronje, and R. M. Rossi, "Digital twins are coming: Will we need them in supply chains of fresh horticultural produce?" *Trends Food Sci. Technol.*, vol. 109, no. 1, pp. 245–258, Mar. 2021, doi: 10.1016/j.tifs.2021.01.025.

[34] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP J. Manuf. Sci. Technol.*, vol. 29, no. 1, pp. 36–52, May 2020, doi: 10.1016/j.cirpj.2020.02.002.

[35] D. L. J. Alexander, A. Tropsha, and D. A. Winkler, "Beware of R^2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models," *J. Chem. Inf. Model.*, vol. 55, no. 7, pp. 1316–1322, Jul. 2015, doi: 10.1021/acs.jcim.5b00206.

[36] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/j.ijforecast.2006.03.001.

[37] E. Triantaphyllou and A. Sánchez, "A sensitivity analysis approach for some deterministic multi-criteria decision-making methods," *Decision Sci.*, vol. 28, no. 1, pp. 151–194, Jan. 1997, doi: 10.1111/j.1540-5915.1997.tb01306.x.

[38] M. Behzadian, S. K. Otaghsara, M. Yazdani, and J. Ignatius, "A state-of-the-art survey of TOPSIS applications," *Expert Syst. Appl.*, vol. 39, no. 17, pp. 13051–13069, Dec. 2012, doi: 10.1016/j.eswa.2012.05.056.

[39] J. P. Brans and Ph. Vincke, "Note—A preference ranking organisation method: The PROMETHEE method for multiple criteria decision-making," *Manage. Sci.*, vol. 31, no. 6, pp. 647–656, Jun. 1985, doi: 10.1287/mnsc.31.6.647.

[40] K. Järvelin and J. Kekäläinen, "Cumulated gain-based evaluation of IR techniques," *ACM Trans. Inf. Syst.*, vol. 20, no. 4, pp. 422–446, Oct. 2002, doi: 10.1145/582415.582418.