

Article

Not peer-reviewed version

---

# Integrating Digital Twin and Optimization for First-Mile Milk Collection Systems: A Scenario-Based Evaluation Framework

---

Andrés Polo <sup>\*</sup>, Lina María Puentes Espejo, Fredy Cervera-Galindo, Marylin Beltran-Rodríguez

Posted Date: 1 May 2026

doi: 10.20944/preprints202605.0004.v1

Keywords: digital twin; first-mile logistics; milk collection systems; supply chain optimization; mixed-integer linear programming (MILP); scenario-based evaluation; adaptive decision-making; agro-industrial supply chains; network design; system resilience



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

# Integrating Digital Twin and Optimization for First-Mile Milk Collection Systems: A Scenario-Based Evaluation Framework

Andrés Polo \*, Lina María Puentes Espejo, Fredy Cervera-Galindo and Marilyn Beltran-Rodríguez

Department of Industrial Engineering, Universidad de Cundinamarca, Chía, 250008, Colombia

\* Correspondence: andrespolo@ucundinamarca.edu.co

## Abstract

This working paper develops a methodological approach for integrating mathematical optimization with a digital twin environment in the analysis of first-mile milk collection systems. The approach combines a mixed-integer linear programming (MILP) model for network design with a digital representation that enables the evaluation of system behavior under changing operating conditions. The optimization model determines the baseline configuration, including the location of collection points, capacity allocation, and producer assignments. This configuration is then embedded into the digital twin, where its performance is examined under a representative perturbation scenario involving a 20% reduction in milk supply. The analysis shows that the baseline configuration, while efficient under nominal conditions, is sensitive to variations in supply, leading to reduced utilization and higher unit costs. Allowing limited operational adjustments within the fixed network structure improves performance, although economic indicators do not fully return to baseline levels. The results also reveal uneven effects across performance dimensions, indicating the presence of trade-offs between economic, operational, and environmental outcomes. The contribution of this study lies in connecting optimization-based design with a digital evaluation environment that enables the assessment of network configurations beyond their initial formulation. The approach provides a structured way to examine how a given design responds to changing conditions without requiring immediate structural modifications. The analysis is illustrative and intended to demonstrate the integration mechanism. Future work will extend this approach through systematic scenario design, quantitative validation, and the incorporation of real-time data.

**Keywords:** digital twin; first-mile logistics; milk collection systems; supply chain optimization; mixed-integer linear programming (MILP); scenario-based evaluation; adaptive decision-making; agro-industrial supply chains; network design; system resilience

---

## 1. Introduction

Digital transformation in agri-food systems has gained significant momentum over the past decade, driven by the convergence of Internet of Things (IoT), artificial intelligence (AI), and, more recently, digital twin (DT) technologies [1]. Within agricultural contexts, digital twins are understood as dynamic representations of physical systems that integrate data streams, enable scenario exploration, and support decision-making under uncertainty [2]. Their potential extends beyond operational efficiency, including improved adaptability to environmental variability and increased resilience in production systems [3,4]. Despite this potential, their implementation in agricultural and livestock systems remains uneven and, in many cases, still exploration. In dairy systems, existing contributions have largely concentrated on monitoring and control at the process or animal level [5]. For example, digital twin architecture for dairy cows has enabled the integration of sensor data, positioning systems, and machine learning models to capture behavioral patterns and health

indicators, offering decision support for farm management [6]. These approaches demonstrate strong technical feasibility, particularly when combining multimodal data and learning algorithms. Their scope, however, is largely confined to perception and analytics, with limited extension toward broader operational or network-level decisions.

Additional contributions have emerged in dairy processing environments [7,8]. Recent studies integrating computer vision, IoT sensing, and machine learning have shown that key process variables—such as acidity during fermentation—can be estimated in near real time, reducing reliance on labor-intensive measurements and enabling more responsive control systems [8]. This reduces reliance on manual measurements and supports more responsive control strategies. Nevertheless, these developments are typically restricted to intra-facility operations and do not address the complexity of upstream logistics, where decisions involve multiple actors, distributed infrastructure, and significant uncertainty.

A closer examination of the literature reveals that most current agricultural implementations fall short of fully realized digital twins. Many systems operate as *digital shadows*, where data are collected and analyzed but rarely fed back into the physical system through automated or adaptive control mechanisms [9]. This partial integration reflects both technological and contextual challenges, including limited cyber-physical connectivity and the inherent variability of agricultural environments [10]. Consequently, DT applications remain fragmented, often addressing isolated components without achieving integrated decision-support capabilities [11]. At the same time, research on agri-food logistics, particularly first-mile systems for perishable products such as raw milk, has developed along a different trajectory. Optimization-based approaches have been widely used to design collection networks, focusing on strategic decisions such as facility location, producer allocation, and storage capacity [12]. These models provide valuable insights into cost efficiency, environmental performance, and system sustainability, frequently within multi-objective or robust optimization frameworks.

However, these approaches are typically formulated as static or scenario-based problems, with limited capacity to represent how the system behaves once a configuration has been selected. Although uncertainty is often incorporated through stochastic or robust formulations, the evaluation of system performance under evolving operational conditions remains limited [13]. Simplifications in the interaction between supply variability, transportation conditions, and infrastructure decisions further constrain the ability to capture real system dynamics [12]. These constraints hinder the evaluation of how designed systems perform when exposed to disruptions such as supply fluctuations, transportation delays, or infrastructure failures.

These two research streams, digital twins in agricultural systems and optimization models for agri-food logistics, have developed in parallel, with minimal integration. DT research has advanced in representing and monitoring physical processes, yet its application to logistics network design and evaluation remains limited. Conversely, optimization models provide rigorous tools for determining efficient configurations but lack mechanisms to assess their performance dynamically or adaptively once implemented. This disconnect becomes particularly critical in first-mile dairy logistics. The collection of raw milk from dispersed producers to cooling or aggregation points involves tightly coupled decisions on location, allocation, and capacity, all within an environment characterized by supply variability, infrastructure constraints, and the perishability of the product. Although these systems play a pivotal role in ensuring product quality and supply chain continuity, existing approaches largely treat them as static design problems, without incorporating digital tools that enable continuous monitoring, simulation, and adaptive decision-making.

Addressing this gap requires moving beyond isolated modeling or monitoring approaches toward integrated frameworks that connect design and operation. In this context, digital twins offer a promising pathway, as they allow the representation of system states and the evaluation of alternative conditions over time. Building on this premise, this study proposes a methodological integration between optimization-based design of first-mile milk collection networks and a digital twin approach oriented toward decision support. The proposed framework enables the evaluation of

network configurations under changing conditions, incorporating economic, logistical, and environmental dimensions. By linking design decisions with system behavior, the study contributes to improving the assessment of logistics performance in agro-industrial supply chains.

## 2. Literature Review

The literature on digital twins in agro-industrial systems has grown rapidly, supported by advances in sensing, data processing, artificial intelligence, and cyber-physical integration [14]. In agriculture, digital twins are generally presented as digital representations capable of linking physical processes with data-driven models, simulation tools, and decision-support mechanisms [15,16]. Their value lies in the possibility of observing system states, anticipating changes, and evaluating alternative responses before interventions are implemented in the physical system. However, the practical maturity of these systems remains uneven. Many developments still emphasize data acquisition and visualization, while the transition toward feedback, control, and operational adaptation is less consolidated [10]. Dairy-related applications reflect this situation. One group of studies has focused on animal monitoring, where sensors, positioning systems, and machine learning techniques are used to identify behavioral patterns and support herd management. Zhang et al. [6] propose a digital twin architecture for dairy cows that integrates data from positioning sensors and inertial devices, combined with machine learning models to identify feeding and behavioral patterns. This type of work is valuable because it shows the technical feasibility of constructing animal-level digital representations. Its contribution, however, remains located mainly at the perception and monitoring level, without extending the digital twin toward logistics network decisions. A second group of studies has explored digital twins in dairy processing. Adeleke et al. [8] show how computer vision, IoT sensors, and machine learning can be combined to estimate variables such as pH and titratable acidity during fermentation. This line of research contributes to process monitoring and control, especially in operations where conventional measurements are time-consuming or intrusive. Still, the unit of analysis is the processing environment, not the upstream collection network. The coordination of producers, routes, tank capacities, and collection decisions remains outside the scope of these contributions.

Beyond these specific developments, recent reviews on digital twins in agriculture converge on a similar conclusion: most current applications remain fragmented [17]. Monitoring-oriented systems dominate where information flows from the physical environment to the digital model, but without closing the loop through feedback mechanisms that enable active intervention in the system [18]. This situation has been described as an incomplete transition from *digital shadows* toward fully functional digital twins, in which bidirectional interaction and control capabilities constitute essential elements. Consequently, digital twins remain primarily limited to decision-support roles for human operators, lacking the characteristics of integrated adaptive management systems.

The literature on first-mile agro-industrial logistics has followed a different path. In this field, mathematical optimization has been used to design collection networks, allocate producers, define capacities, and evaluate economic and environmental trade-offs. For the raw milk collection problem, Polo et al. [12] formulate a robust logistics model that considers cooling tank location, producer assignment, net present value, CO<sub>2</sub> emissions, and robustness evaluation through FePIA (Features-Perturbation-Impact-Analysis). This contribution is directly relevant because it addresses the structural decisions that define the first-mile dairy network.

Even so, optimization-based approaches face an important limitation when they are used as stand-alone design tools. They usually determine a configuration under baseline assumptions or predefined scenarios, but they provide limited support for evaluating how that configuration behaves after implementation. The network may be optimal under the modeled conditions and still lose performance when supply, transport conditions, or capacity availability change. Studies on robust and stochastic optimization recognize the importance of uncertainty, yet the dynamic assessment of already-designed systems remains underdeveloped. In particular, the literature highlights the need to incorporate robust metrics capable of capturing system behavior under real

disturbances, beyond the initial optimization stage [19]. On the other hand, the interaction among key variables—such as milk supply variability, transportation conditions, and storage capacity—is often simplified, which restricts the ability to accurately represent the operational dynamics of these systems [12].

This separation between digital twin research and logistics optimization defines the gap addressed in this working paper. Digital twin studies contribute representation, monitoring, and simulation capabilities, but rarely incorporate prescriptive models for first-mile dairy logistics. Optimization studies provide rigorous tools for network design, but they do not usually embed the resulting configuration into a digital environment where its performance can be re-evaluated under changing conditions. The missing link is an integrated approach in which the optimized design becomes the starting point for a digital twin evaluation process. Four specific gaps emerge from this reading of the literature. The first is the limited integration between digital twins and mathematical optimization for first-mile milk collection systems. The second is the lack of mechanisms for evaluating the post-design behavior of logistics configurations under perturbations. The third is the insufficient joint treatment of economic, operational, and environmental indicators within a digital evaluation environment. The fourth is the limited evidence of these approaches in real agro-industrial contexts characterized by dispersed producers, heterogeneous infrastructure, and high variability in supply.

This study responds to these gaps by proposing a digital twin–optimization framework for first-mile milk collection systems. The optimization model defines the baseline network configuration, while the digital twin provides the environment for evaluating that configuration under changing operating conditions. This integration shifts the role of the model from a one-time design instrument to a component of a broader decision-support process. The resulting framework links network design, scenario-based evaluation, and adaptive operational analysis within a single methodological structure.

### 3. Methodology

#### 3.1. General Architecture of the Proposed Approach

The proposed methodology is based on the integration of a mathematical optimization model with a digital twin oriented toward decision support for first-mile raw milk collection. This integration responds to a recurring limitation in the literature: digital twins applied to the dairy sector have advanced in animal monitoring, behavior recognition, process control, and traceability, but still present limited connection with prescriptive models capable of supporting logistics design and operational decision-making at the network level. In this work, the digital twin is not conceived solely as a visualization or monitoring platform, but as a computational environment capable of representing the physical network, evaluating disruption scenarios, and activating the optimization model as an analytical core to compare decision alternatives.

The architecture is structured into five interdependent layers. The first corresponds to the physical system, composed of milk producers, collection routes, transport modes, cooling tanks, aggregation points, and territorial conditions. The second layer corresponds to the digital representation of the network, where nodes, flows, capacities, costs, distances, slopes, emissions, and supply variables describing system operation are organized. The third layer contains the mathematical optimization model, responsible for defining viable logistics configurations through location decisions, producer–cluster allocation, capacity selection, and economic and environmental evaluation. The fourth layer corresponds to the simulation module, where operational parameters are modified to represent scenarios such as supply drops, route blockages, cost increases, or reductions in effective cooling capacity. The fifth layer corresponds to the decision module, where results are translated into indicators and recommendations for producers, associations, or logistics managers. This structure allows differentiating the proposed approach from current developments in agricultural digital twins. In several studies, digital architecture is limited to capturing data from

the physical system and projecting it onto an analytical interface, which is closer to a digital shadow than to a fully integrated digital twin. Zhang et al. [20], for example, develop an architecture for dairy cows based on sensors, positioning, and machine learning, useful for monitoring animal behavior, but focused on the perception and visualization level. Similarly, Adeleke et al. [8] integrate computer vision, IoT sensors, and machine learning models to monitor fermentation variables in near real time, although within an intraplant control problem. These contributions are valuable, but they do not address logistics decision-making related to location, capacity, allocation, and economic-environmental performance of a collection network.

In the proposed approach, the optimization model operates as the prescriptive engine of the digital twin. Based on network data, the model determines a base configuration: which tanks should be activated, what capacity should be installed, and how producers should be assigned. Subsequently, this configuration is embedded into the digital twin to evaluate its behavior under varying conditions. In this way, the mathematical model answers the design question—what configuration should be selected—while the digital twin addresses the operational question—how that configuration behaves when the system changes. This complementarity allows overcoming the traditional separation between static optimization models and digital monitoring platforms. The architecture also builds upon the first-mile model developed by Polo et al. [12], in which a multi-objective problem for raw milk collection was formulated, considering net present value, CO<sub>2</sub> emissions, tank location, producer allocation, and robustness evaluation through FePIA. This model provides a suitable foundation to be reinterpreted as the analytical core of the digital twin, as it integrates strategic and tactical decisions with economic, environmental, and social indicators. However, the methodological contribution of this work lies in shifting the use of the model from a one-time optimization logic toward a continuous digital evaluation approach, where the selected configuration can be subjected to operational and disruption scenarios.

In this way, the proposed architecture establishes a transition from a robust design model toward a digital twin-based decision-support system. This transition does not initially require the full implementation of IoT sensors or real-time control in the field. For the scope of this working paper, the digital twin is conceived as a computational representation fed by historical data, technical parameters, and simulated scenarios. This decision is consistent with the current state of digital twin adoption in agriculture, where many applications evolve progressively from digital models and digital shadows toward systems with greater cyber-physical integration and feedback capabilities.

### 3.2. Mathematical Model

The mathematical model aims to determine an efficient configuration of the first-mile milk collection network by simultaneously considering economic, logistical, and environmental criteria. The problem is formulated as a mixed-integer linear programming (MILP) model that integrates location, allocation, and capacity decisions. Let  $P$  denote the set of milk producers and  $A$  the set of candidate locations for the installation of collection points (cooling tanks). Each producer  $p \in P$  generates a quantity of milk  $q_p$ , while each location  $a \in A$  is associated with a maximum capacity  $Cap_a$ . The distance between producer  $p$  and location  $a$  is denoted by  $d_{pa}$ . The decision variables are defined as follows. A binary variable  $y_a$  takes value 1 if a collection point is activated at location  $a$ , and 0 otherwise. A binary variable  $x_{pa}$  indicates whether producer  $p$  is assigned to location  $a$ . The objective function seeks to maximize the overall system performance by combining economic and environmental components. It can be expressed as:

$$\max Z = NPV - C^{trans} - C^{inst} - \lambda E$$

where  $NPV$  represents the net present value of the system,  $C^{trans}$  corresponds to transportation costs,  $C^{inst}$  includes installation and operational costs of the collection points, and  $E$  denotes the emissions associated with transportation activities. The parameter  $\lambda$  controls the relative importance of environmental impact within the objective function.

The model is subject to a set of constraints that ensure the feasibility of the network configuration. Each producer must be assigned to exactly one collection point:

$$\sum_{a \in A} x_{pa} = 1 \forall p \in P$$

The total amount of milk assigned to a collection point cannot exceed its installed capacity:

$$\sum_{p \in P} q_p x_{pa} \leq Cap_a y_a \forall a \in A$$

This constraint ensures that capacity is only available when a location is activated. In addition, assignments are restricted to active locations:

$$x_{pa} \leq y_a \forall p \in P, a \in A$$

The environmental impact of the system is modeled as a function of transportation distances and an emission factor  $\gamma$ :

$$E = \sum_{p \in P} \sum_{a \in A} \gamma d_{pa} q_p x_{pa}$$

Transportation costs are defined in a similar way, as a function of distance and unit transport cost:

$$C^{trans} = \sum_{p \in P} \sum_{a \in A} c_{pa} d_{pa} q_p x_{pa}$$

where  $c_{pa}$  represents the cost per unit distance. Installation costs are associated with the activation of collection points:

$$C^{inst} = \sum_{a \in A} f_a y_a$$

where  $f_a$  is the fixed cost of installing and operating a cooling tank at location  $a$ .

The model yields a baseline network configuration defined by the set of activated locations, their corresponding capacities, and the allocation of producers. This solution constitutes the structural design of the system. In contrast with traditional applications, the model is not used as a one-time decision tool. The resulting configuration is embedded into the digital twin and treated as a reference state for further evaluation. This reinterpretation allows the model to operate as part of a broader analytical process, where the behavior of the system is examined under changing conditions. In this sense, the mathematical model provides the structural foundation, while the digital twin enables the dynamic assessment of system performance.

### 3.3. Integration Between the Optimization Model and the Digital Twin

The integration between the optimization model and the digital twin establishes a unified analytical environment for evaluating first-mile milk collection systems under changing conditions. Optimization and simulation are not treated as independent components. A structured interaction links both elements, allowing the system to evolve from a static design problem into a continuous evaluation process. This integration reflects the presence of two distinct decision layers. Structural decisions determine the configuration of the network, including the selection of collection points, installed capacities, and producer–cluster assignments. These decisions are typically long-term and difficult to modify once implemented. In contrast, operational conditions—such as variations in milk supply, transport constraints, and cost fluctuations—change over time and directly affect system performance. The digital twin provides an environment where these variations can be explored without altering physical infrastructure.

The process begins with the solution of the optimization model under baseline conditions. Using spatial, operational, and economic data, the model identifies a network configuration that maximizes system performance according to the defined objective function. The resulting solution specifies active collection points, capacity levels, and producer assignments. This configuration is not

interpreted as the endpoint of the analysis. It serves as a reference state that is subsequently embedded into the digital twin. Once embedded, the configuration becomes part of a virtual representation of the system. This representation incorporates both the physical structure and the decisions obtained from the optimization model. The distinction is important, since the evaluation is not limited to observed system states but also includes configurations derived from the design phase.

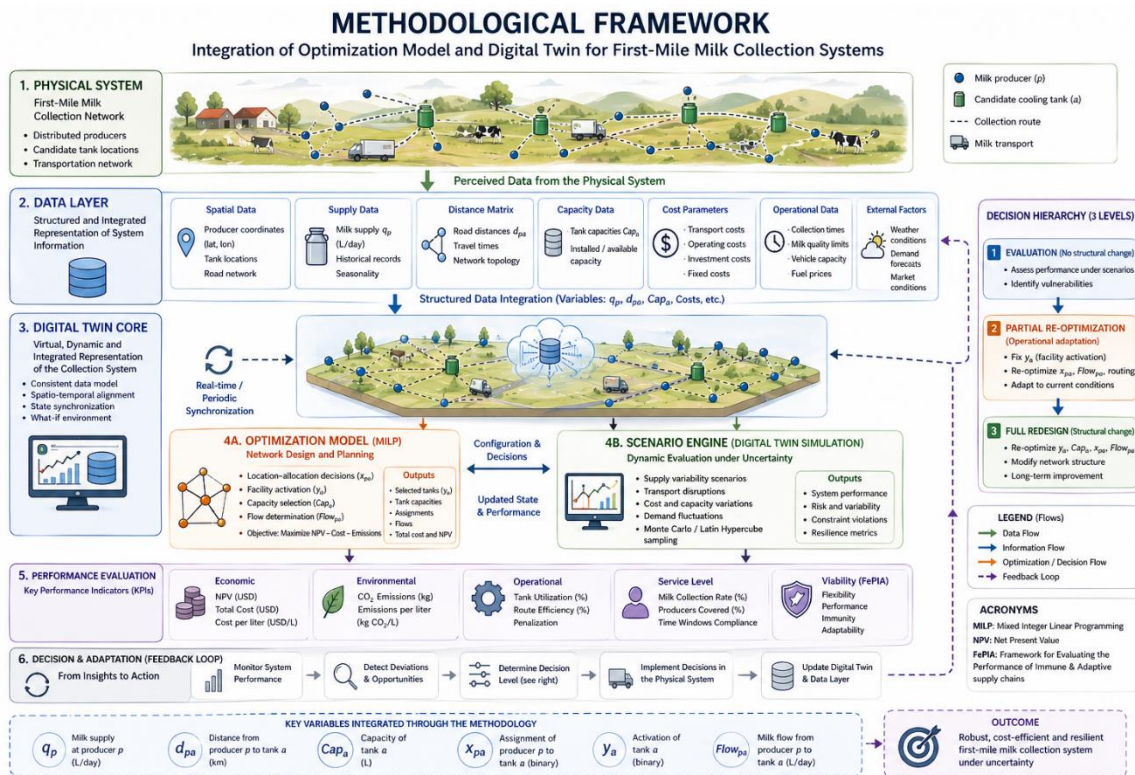
Within this environment, the system is subjected to controlled perturbations that reflect realistic operating conditions. These perturbations include changes in supply levels, disruptions in transport routes, variations in logistics costs, and reductions in effective capacity. Each perturbation defines a new system state, allowing the evaluation of how the configuration responds to different conditions.

The framework incorporates a hierarchical mechanism for decision adaptation. In the first level, the system is evaluated under perturbed conditions without modifying the network structure. This stage reveals the extent to which performance deteriorates when external conditions change. In the second level, partial re-optimization is performed by fixing structural decisions and adjusting operational variables, such as the allocation of producers to collection points. This step captures the capacity of the system to adapt without requiring structural changes. A third level considers full re-optimization when persistent inefficiencies indicate that the initial configuration is no longer adequate. The interaction between the digital twin and the optimization model operates in both directions. Updated system states and scenario parameters are introduced into the model, while the resulting configurations and performance indicators are incorporated back into the digital representation. This feedback loop supports a continuous evaluation process in which system behavior is examined across multiple conditions, moving beyond a single set of assumptions.

From a methodological perspective, this integration extends the role of optimization models beyond static network design. The model becomes part of a broader analytical structure that supports scenario analysis, sensitivity assessment, and decision-making under variability. At the same time, the digital twin acts as an evaluation platform, where different configurations can be examined under realistic conditions. This approach addresses a limitation observed in both research streams. Digital twin applications in agriculture often focus on monitoring and visualization, with limited incorporation of prescriptive decision models. Optimization approaches, on the other hand, identify efficient configurations but provide limited support for evaluating their behavior after implementation. Integrating both components enables a more complete analysis in which design decisions are assessed according to their performance under changing conditions.

The proposed integration does not require real-time data in its initial implementation. The digital twin can be constructed using historical information, estimated parameters, and simulated scenarios. This allows the approach to be applied in contexts where sensing infrastructure is limited. At the same time, the structure is compatible with future extensions that incorporate real-time data streams and more advanced cyber-physical integration. Overall, the integration transforms the analysis of first-mile milk collection systems from a static design exercise into a dynamic evaluation process. The optimization model provides the structural basis of the system, while the digital twin enables the examination of its performance under varying conditions. Together, they provide a more robust foundation for decision-making in agro-industrial supply chains.

The integration between the optimization model and the digital twin transforms the analysis of first-mile milk collection systems from a static design problem into a dynamic evaluation process. The optimization model provides the structural backbone of the system, while the digital twin enables the exploration of its behavior under varying conditions. This combined approach offers a more realistic and robust basis for decision-making in agro-industrial supply chains. The overall methodological structure of the proposed approach, including the interaction between the physical system, the data layer, the digital twin core, and the optimization and simulation components, is summarized in Figure 1.

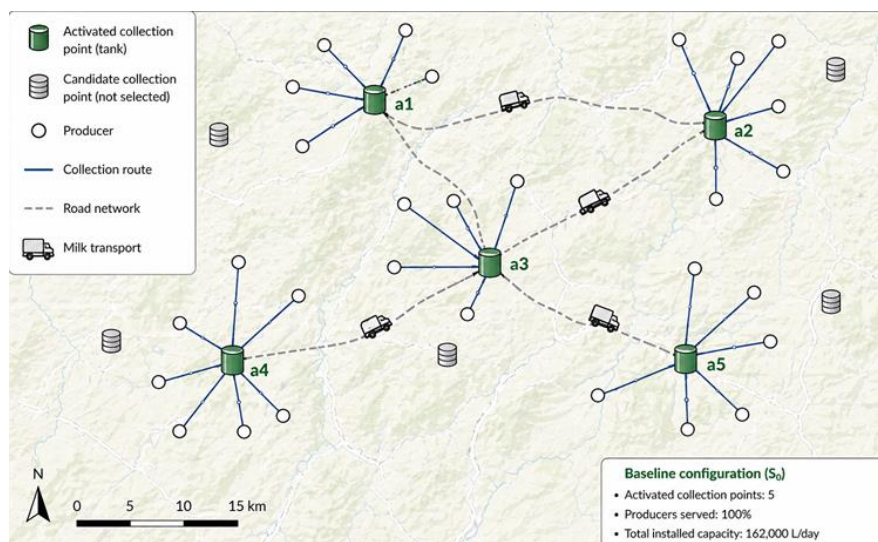


**Figure 1.** Integrated methodological framework for the digital twin–optimization approach in first-mile milk collection systems. The framework links the physical milk collection network with its digital representation through a structured data layer that integrates spatial, operational, and economic variables (e.g.,  $q_p$ ,  $d_{pa}$ ,  $Cap_a$ ). The digital twin core acts as a dynamic and synchronized environment where the optimization model (MILP) defines network configurations and the scenario engine evaluates system performance under uncertainty. A hierarchical decision is incorporated, distinguishing between evaluation, partial re-optimization, and full redesign levels. System performance is assessed using economic, operational, and viability indicators (FePIA), enabling continuous adaptation through a feedback loop. This integration transforms static network design into a dynamic decision-support process under varying operational conditions.

## 4. Results and Discussion

### 4.1. Baseline Network Configuration

The optimization model was first solved under baseline conditions to determine the structural configuration of the first-mile milk collection network. The resulting solution defines the set of activated collection points, the allocation of producers to clusters, and the installed capacity at each location. The model identifies a spatially coherent configuration in which producers are grouped into clusters based on proximity and logistical efficiency. The selected configuration minimizes transportation costs while maintaining feasible capacity levels, resulting in a balanced distribution of milk flows across the network. From an operational perspective, the solution reflects a trade-off between consolidation and accessibility, ensuring that collection points are sufficiently close to producers while avoiding excessive fragmentation of the network. As shown in Figure 2, the fixed configuration exhibits significant underutilization when no adjustments are allowed, whereas partial re-optimization improves the distribution of flows and increases system efficiency.



**Figure 2. Baseline network configuration ( $S_0$ ) of the first-mile milk collection system.** The figure illustrates the optimal configuration obtained from the MILP model, including the activated collection points (tanks), producer locations, and assignment routes. Producers are grouped into clusters based on spatial proximity and logistical efficiency, with milk transported through the road network to the selected collection points. Non-selected candidate locations are also shown for reference. The resulting configuration ensures full-service coverage while minimizing transportation costs and balancing installed capacity across the network.

The baseline results provide key performance indicators, including total system cost, net present value (NPV), and resource utilization levels. These indicators establish a reference point for subsequent evaluation within the digital twin environment. Importantly, the baseline solution is not interpreted as a final or definitive outcome, but rather as a designed configuration whose performance must be assessed under varying operational conditions.

#### 4.2. Embedding the Optimized Configuration into the Digital Twin

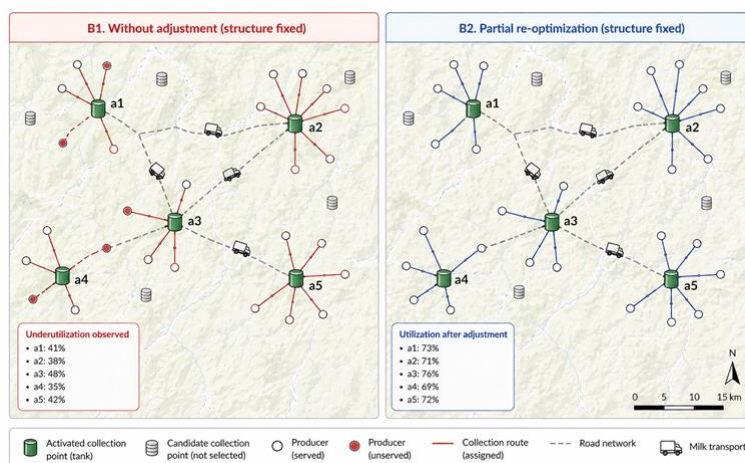
Once the baseline configuration is obtained, it is embedded into the digital twin, which acts as a virtual representation of the system. In this stage, the digital twin incorporates not only the physical characteristics of the network, but also the decision structure derived from the optimization model. The embedding process consists of fixing the structural variables—such as the activation of collection points and installed capacities—while allowing the system to be evaluated under modified operating conditions. This transformation is central to the proposed approach, as it enables the analysis of system behavior without altering the physical infrastructure.

Within the digital twin, the system is represented as a state defined by spatial, operational, and economic variables. This state serves as the baseline configuration ( $S_0$ ), from which alternative system conditions can be generated. Unlike traditional simulation approaches, the digital twin maintains consistency between the optimized design and the evaluation environment, ensuring that all analyses are grounded in a feasible and coherent network structure.

#### 4.3. Illustrative Scenario: Supply Reduction

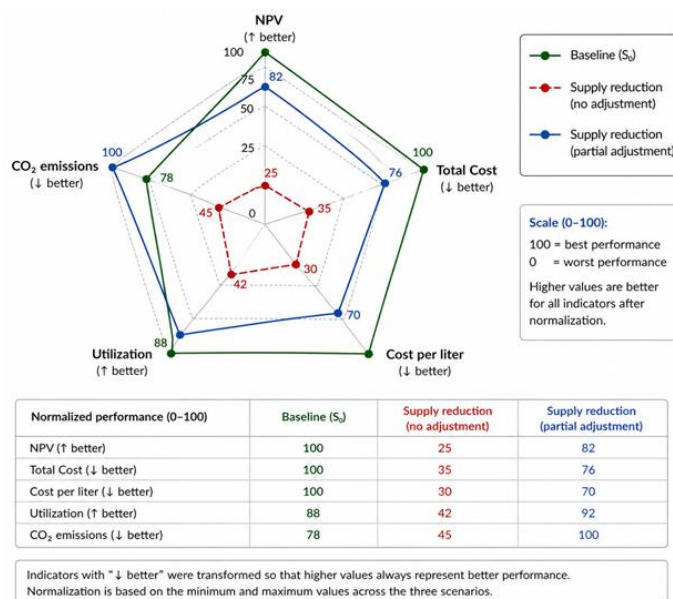
To demonstrate the functionality of the proposed framework, an illustrative scenario is considered in which milk supply at the producer level is reduced by 20%. This scenario reflects realistic conditions such as seasonal variability, production constraints, or external disruptions affecting farm output. The evaluation is conducted in two stages. First, the system performance is assessed without modifying the baseline configuration. In this case, the fixed structure reveals its sensitivity to supply changes, as reduced volumes affect the utilization of collection points and the overall efficiency of transportation routes. The results show a decrease in economic performance,

accompanied by underutilization of certain facilities and increased costs per unit collected. In the second stage, partial re-optimization is performed by maintaining the structural configuration while allowing adjustments in the allocation of producers to clusters. This represents an operational response within the digital twin, where the system adapts to new conditions without requiring structural changes. The adjusted solution improves the distribution of flows, partially mitigating the impact of the supply reduction.



**Figure 3. Illustrative scenario under 20% supply reduction.** The figure compares system performance under a supply reduction scenario with and without operational adjustment. Panel B1 shows the baseline configuration evaluated without modifying producer assignments, revealing underutilization of collection points and reduced efficiency. Panel B2 presents the result of partial re-optimization, where assignments are adjusted while maintaining the original network structure. The comparison highlights the ability of the digital twin to support adaptive responses that improve utilization and mitigate performance degradation without requiring structural redesign.

The multi-criteria performance comparison presented in Figure 4 provides a synthesized view of system behavior across scenarios using normalized indicators. The baseline configuration achieves the highest economic performance; however, the supply reduction scenario without adjustment leads to a substantial and generalized degradation across all dimensions. When partial re-optimization is introduced within the digital twin, the system exhibits a heterogeneous recovery pattern. Notably, operational and environmental indicators such as utilization and CO<sub>2</sub> emissions show significant improvement, in some cases surpassing baseline performance due to more efficient resource allocation under constrained conditions. In contrast, economic indicators such as NPV and cost-related metrics, while improved relative to the unadjusted scenario, do not fully recover to baseline levels. This asymmetric response highlights the presence of trade-offs between efficiency, cost, and environmental impact, reinforcing the importance of adaptive decision-making within the proposed digital twin framework.



**Figure 4. Performance comparison across scenarios using normalized indicators.** The radar chart shows that partial re-optimization mitigates the deterioration caused by supply reduction, although its effect is heterogeneous across indicators. The adjusted scenario improves utilization and CO<sub>2</sub> performance relative to the baseline, while economic indicators remain below the nominal configuration, revealing operational–economic trade-offs within the digital twin environment.

#### 4.4. Discussion

The results obtained from the baseline configuration and the illustrative scenario provide insight into how first-mile milk collection networks respond to changes in operating conditions. The optimization model identifies a configuration that is efficient under nominal assumptions; however, its performance deteriorates when milk supply decreases. This effect is reflected in lower utilization levels at collection points and higher costs per unit collected, indicating that the network structure is sensitive to variations in supply. Embedding the configuration into the digital twin allows this behavior to be observed explicitly. Under reduced supply, some collection points operate below their intended capacity, while transport routes become less efficient due to reduced consolidation of flows. These effects are not captured during the optimization stage, where the model assumes stable input conditions. The digital twin complements this limitation by exposing how the system performs once the configuration is subjected to external changes.

The introduction of partial re-optimization reveals an important aspect of system behavior. Adjusting producer–cluster assignments, while maintaining the existing network structure, improves flow distribution and increases the utilization of active collection points. This adjustment does not fully restore economic performance, but it reduces the impact of the perturbation without requiring new infrastructure. In practical terms, this suggests that part of the system’s resilience can be achieved through operational decisions, such as reallocating producers or modifying collection routes. The comparison of performance indicators confirms that system response is not uniform across dimensions. Economic indicators such as net present value and total cost remain below baseline levels after the perturbation, even with operational adjustments. In contrast, utilization and environmental performance may improve relative to the unadjusted scenario and, in some cases, approach or exceed baseline values. These differences indicate the presence of trade-offs between efficiency, cost, and environmental impact. Evaluating the system from a single objective would not capture these effects.

From a methodological perspective, the proposed integration connects two lines of research that are often treated separately. Optimization models provide structured solutions for network design, while digital twins enable the observation of system states under changing conditions. Combining

both allows the evaluation of a configuration beyond its initial formulation, focusing on how it behaves when inputs change. This shift is particularly relevant for agro-industrial systems, where variability in supply and operating conditions is unavoidable. The analysis presented here is intentionally limited to a single perturbation scenario to illustrate the integration mechanism. A broader evaluation, including multiple scenarios and statistical validation, is left for future work. Extending the framework in this direction would allow a more systematic assessment of system performance and provide stronger evidence of its applicability in real-world settings.

## 5. Conclusions

This working paper develops an approach that integrates mathematical optimization with a digital twin environment for the analysis of first-mile milk collection systems. The model defines a baseline network configuration in terms of collection point location, capacity allocation, and producer assignment. This configuration is then used as a reference state within the digital twin, where its performance is evaluated under changing operating conditions. The results show that configurations obtained under nominal assumptions are sensitive to variations in milk supply. A reduction in supply leads to lower utilization of collection points and higher unit costs, revealing structural dependencies within the network. Embedding the configuration into the digital twin makes these effects observable, allowing the evaluation of system behavior beyond the initial design stage.

Allowing partial re-optimization improves system performance without modifying the network structure. Adjustments in producer allocation increase utilization and reduce inefficiencies, although economic indicators do not fully return to baseline levels. This indicates that operational flexibility can mitigate part of the impact of disruptions, while structural limitations remain. The analysis also shows that system performance evolves differently across dimensions. Economic, operational, and environmental indicators respond in distinct ways to the same perturbation, highlighting the presence of trade-offs that cannot be captured through a single objective function.

From a methodological perspective, the study connects optimization-based network design with a digital evaluation environment. The resulting approach enables the assessment of a given configuration under varying conditions, extending the role of the model beyond one-time design. This is particularly relevant for first-mile milk collection systems, where supply variability and infrastructure constraints affect daily operations. The scope of the analysis is intentionally limited to an illustrative scenario. Future work will expand the framework through systematic scenario design, quantitative validation, and the incorporation of real-time data. These extensions will allow a more comprehensive evaluation of system performance and support the development of adaptive decision-support tools for agro-industrial supply chains.

## References

1. Tyagi M, Tyagi K. Industry 5.0 readiness factors for implementing digital twins in sustainable supply chains: A pathway to circularity. *J Clean Prod* 2025;529. <https://doi.org/10.1016/j.jclepro.2025.146795>.
2. Zhang Y, Zou Y, Liu L, Manzardo A, Wang X, Wu J. A digital twin-driven dynamic shelf life approach for mitigating food loss across fruit supply chains. *J Stored Prod Res* 2026;117. <https://doi.org/10.1016/j.jspr.2026.103016>.
3. Zohdi TI. A machine-learning enabled digital-twin framework for next generation precision agriculture and forestry. *Comput Methods Appl Mech Eng* 2024;431. <https://doi.org/10.1016/j.cma.2024.117250>.
4. Cesco S, Sambo P, Borin M, Basso B, Orzes G, Mazzetto F. Smart agriculture and digital twins: Applications and challenges in a vision of sustainability. *European Journal of Agronomy* 2023;146. <https://doi.org/10.1016/j.eja.2023.126809>.
5. Feng C, Reed KF, Giordano JO, You F. Precision decarbonization for clean dairy farming with digital twin and edge intelligence 2025;2. <https://doi.org/10.1016/j.ynexs.2025.100105>.

6. Zhang Y, Zhang Y, Gao M, Wang X, Dai B, Shen W. Multimodal behavior recognition for dairy cow digital twin construction under incomplete modalities: A modality mapping completion network approach. *Artificial Intelligence in Agriculture* 2025;15:459–69. <https://doi.org/10.1016/j.aiaa.2025.04.005>.
7. Polo A. A Mathematical Modeling Approach for Analyzing Dairy Supply Chains Under Heterogeneous Conditions 2026. <https://doi.org/10.20944/preprints202603.1710.v1>.
8. Adeleke I, Nwulu N, Adebo OA. A digital twin approach for real-time monitoring of amasi acidity using non-invasive computer vision, IoT, and machine learning. *Food Control* 2026;182. <https://doi.org/10.1016/j.foodcont.2025.111809>.
9. Purcell W, Neubauer T. Digital Twins in Agriculture: A State-of-the-art review. *Smart Agricultural Technology* 2023;3. <https://doi.org/10.1016/j.atech.2022.100094>.
10. Purcell W, Neubauer T, Mallinger K. Digital Twins in agriculture: challenges and opportunities for environmental sustainability. *Curr Opin Environ Sustain* 2023;61. <https://doi.org/10.1016/j.cosust.2022.101252>.
11. Gabellini M, Regattieri A, Bortolini M, Ronchi M. Conceptualization and validation of an intelligent digital twin design framework for supply chain risk management. *International Journal of Information Management Data Insights* 2025;5. <https://doi.org/10.1016/j.jjime.2025.100365>.
12. Polo A, Escobar JW, Montoya MP. Robust design of a logistics system using FePIA procedure and analysis of trade-offs between CO<sub>2</sub> emissions and net present value. *Heliyon* 2023;9. <https://doi.org/10.1016/j.heliyon.2023.e18444>.
13. Hosseini Dehshiri SJ, Amiri M. Considering the circular economy for designing closed-loop supply chain under hybrid uncertainty: A robust scenario-based possibilistic-stochastic programming. *Expert Syst Appl* 2024;238. <https://doi.org/10.1016/j.eswa.2023.121745>.
14. Amer Y, Soufali A, Zaghwan A. A digital twin-based framework for predictive quality assurance and supply chain resilience in the automotive industry. *Advanced Engineering Informatics* 2026;69. <https://doi.org/10.1016/j.aei.2025.103969>.
15. Pylianidis C, Osinga S, Athanasiadis IN. Introducing digital twins to agriculture. *Comput Electron Agric* 2021;184. <https://doi.org/10.1016/j.compag.2020.105942>.
16. Mitsanis C, Hurst W, Tekinerdogan B. A 3D functional plant modelling framework for agricultural digital twins. *Comput Electron Agric* 2024;218. <https://doi.org/10.1016/j.compag.2024.108733>.
17. Zhang Y, Zhang Y, Gao M, Dai B, Kou S, Wang X, et al. Digital twin perception and modeling method for feeding behavior of dairy cows. *Comput Electron Agric* 2023;214. <https://doi.org/10.1016/j.compag.2023.108181>.
18. Hurtado E, Suryadevara C, Walker J, Tayag G, Ryait H. Digital twins as decision-support tools for automation in agriculture: A case study on robotic vaccination. *Smart Agricultural Technology* 2026;13. <https://doi.org/10.1016/j.atech.2025.101675>.
19. González JP, Sanchez-Londoño D, Barbieri G. A Monitoring Digital Twin for Services of Controlled Environment Agriculture. *IFAC-PapersOnLine*, vol. 55, Elsevier B.V.; 2022, p. 85–90. <https://doi.org/10.1016/j.ifacol.2022.09.188>.
20. Hosseini Dehshiri SJ, Amiri M, Olfat L, Pishvae MS. A robust fuzzy stochastic multi-objective model for stone paper closed-loop supply chain design considering the flexibility of soft constraints based on Me measure. *Appl Soft Comput* 2023;134. <https://doi.org/10.1016/j.asoc.2022.109944>.

**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.