Computers in Industry 172 (2025) 104347

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

A platform to support the fast development of digital twins for agricultural holdings

Jorge Laguna *, Mario E. Suaza-Medina, Rubén Béjar, Javier Lacasta, F. Javier Zarazaga-Soria

Aragon Institute of Engineering Research (I3A), Universidad de Zaragoza, Zaragoza, Spain

ARTICLE INFO

Keywords:
Digital twin
Open-field agriculture
Platform
Machine learning
Smart farming

ABSTRACT

Industry 4.0 has advanced in agriculture through Smart Agriculture initiatives, yet open-field farming lags in the adoption of digital twins. Although digital twins have transformed manufacturing since 2011, their application in open-field farming remains limited by environmental variability, data scarcity, and financial constraints. This paper addresses four gaps: the lack of affordable platforms for small farms that dominate European agriculture; the need to manage agricultural complexity through data-driven models rather than the physical modelling approaches prevalent in non-agricultural sectors; the absence of open sources solutions adapted to agriculture's slower innovation pace; the breach between technology and farmers. The platform features innovations in data workflow integration, open data incorporation, a cost-effective shared warehouse, and scalable data pipelines. To validate the proposed platform, a case study with two example digital twins mirroring two fields is conducted. This implementation ran efficiently on modest hardware (2 vCPUs, 4GB RAM). It achieved an average CPU usage of 60%, RAM usage of 2.5 GB, and a deployment time of around one minute. This helps lowering adoption barriers for small holdings and bridging the gap between basic monitoring and complex future systems.

1. Introduction

Agricultural holdings form the foundation of our food production systems. They are the operational units where land management, farming activities, cost management, and strategic decisions intersect. There is an important distinction between controlled-environment agriculture, e.g. the use of greenhouses, and open-field farming. Greenhouses provide stable conditions, making it easier to introduce technology and to collect data. In contrast, open-field farms face unique challenges when adopting digital technologies, due to unpredictable weather, variable soil conditions, and threats from pests and diseases. Within Europe, there are about 9.1 million farms, and nearly 64% of them are small-scale holdings of less than five hectares (European Union Eurostat, 2023). Several studies have shown slow rates of technology adoption, and hesitance among farmers regarding complex digital systems (Ruzzante et al., 2021; Mao et al., 2021; Gabriel and Gandorfer, 2023).

Current digital solutions do not meet all of the farmers' needs. These solutions include precision farming, the use of IoT devices, and farm management software. They are usually expensive and tailored mainly for larger farms. They typically focus on single aspects of farm management without offering the holistic view that would be better for effective decision-making (Sharma et al., 2022; Idoje et al., 2021). Besides this, some studies also point out other socio-economic issues with smart farming (De Alwis et al., 2022). Digital twins are a promising technology that can provide opportunities to address some of these issues by creating virtual replicas of farms. According to Grieves and Vickers (2017), digital twins can be categorised into several types: Digital Twin Prototypes (DTP) (prototypical physical assets), Digital Twin Instances (DTI) (specific physical assets) and Digital Twin Environments (DTE) (integrated spaces for operating several digital twins).

Digital twins can help to predict and control farming issues, reducing both risks and costs. They can also analyse and improve the efficiency of some interventions before carrying them out in the physical world. They have already proven themselves successful in other industries, where they have significantly improved efficiency, resource usage, and decision-making (Tao et al., 2022), with some early adopters already using them in 2011 (Pylianidis et al., 2021). In agriculture, the use of digital twins began to be explored later, around 2017. This delay is largely due to the complexity of modelling living ecosystems,

* Corresponding author.
E-mail address: jlagunar@unizar.es (J. Laguna).

https://doi.org/10.1016/j.compind.2025.104347
Received 29 October 2024; Received in revised form 28 June 2025; Accepted 27 July 2025
Available online 9 August 2025

unpredictable weather conditions, pests, and limited budgets in smaller farms. Current applications of digital twins in agriculture are focused on monitoring crops, but prediction and optimisation are increasingly being explored, although they face some serious challenges often related to their high costs. Potential solutions would include cheaper components, new business models like leasing, or the support by governments and non-profit organisations (Wang, 2024).

The role of big data in agriculture is also growing in importance. Farms need to collect and process vast amounts of information from various sources to generate actionable insights. Lytos et al. (2020) emphasise the need to combine different types of data, open and proprietary, including descriptive data, vector cartography and satellite imagery. According to a recent analysis, open data has proven relevant in agriculture. However, around 83% of agriculture studies using open data mainly rely on satellite images, leaving other valuable data sources underused such as weather, pest data or economic indicators (Chamorro-Padial et al., 2024). Moreover, the agricultural sector remains in an early stage regarding the use of big data. It still faces challenges related with poor data quality, heterogeneity and sparsity of the available information, stakeholder distrust to share proprietary data, and the critical need for open platforms which can accelerate technological development and innovation in that sector (Wolfert et al., 2017).

Although the interest in digital twins for agriculture is rising, only a few existing products provide fully integrated digital twins. Inexperienced users find it challenging to use these systems (Tagarakis et al., 2024). Unlike other sectors with standardised processes and abundant data that enable plug-and-play solutions, open-field agriculture has an uncertain and difficult-to-model environment which presents some unique challenges. To help with these challenges, we propose a developer-focused, data-driven digital twin platform that provides a Digital Twin Environment where autonomous Digital Twin Instances coexist and communicate within a shared infrastructure. Instead of pursuing standardised plug-and-play solutions, which might not be feasible within the complex reality of open-field farming, our fully open-source platform intends to empower developers to be the link between technology and farmers. Our proposal leverages both proprietary and open data sources through a data warehouse that can be shared among several farms, in a way that makes digital twins achievable for farms of all sizes.

The rest of this paper is organised as follows: Section 2 reviews the state of the art in digital twin technology with a focus on agricultural applications. Section 3 presents the conceptual architecture of our digital twin platform for open-field farming. Section 4 provides a detailed breakdown of this architecture. Section 5 describes the technological implementation of the platform. Section 6 demonstrates the platform's functionality through a practical example. Finally, Section 7 summarises our main conclusions, and proposes some future work.

2. State of the art

Digital twin technology has evolved significantly over the past decade, transitioning from theoretical frameworks to practical implementations across multiple sectors. The natural evolution of digital twins was described by Matta and Lugaresi (2024): the "digital model" (offline physical object representation without autonomous data flows), the "digital shadow" (passive, unidirectional representations used for monitoring without controlling physical systems), and the "digital twin" (bi-directional and autonomous mirroring of the physical object). According to Tao and Zhang (2017), a digital twin can be conformed by five dimensions, with the virtual dimension being either model-based or model-free (data-driven). Hussain (2023) documented the industrial application of digital twins by major technology companies including Siemens, Microsoft, and General Electric. While digital twins originated in manufacturing to optimise processes and reduce costs, their application has expanded to aerospace, healthcare, and agriculture, among other sectors (Attaran and Celik, 2023). They fundamentally serve as virtual replicas of physical assets, processes, or systems that can be used for monitoring, analysis, and optimisation.

Iranshahi et al. (2025) conducted a cross-domain analysis of digital twin technology across seven engineering fields, rating overall technology maturity at TRL 4.8.[1] Their expert panel identified integration with digital threads as the most significant challenge (4.2 out of 9). They noted that agricultural digital twins lag behind other sectors due to the prevalence of small-to-medium enterprises, limited computational resources, and the complexity of modelling living systems. A review by Purcell and Neubauer (2023) examined 31 agricultural digital twin papers from 2016–2021. The review revealed that most focused on controlled environments, with only 37% implementing complete digital twins with bi-directional data flow. Key challenges included the need for domain-specific knowledge, substantial data requirements, and complex farm-level integration. Similarly, Tagarakis et al. (2024) analysed 34 papers on agricultural digital twins, finding that while all demonstrated simulation capabilities, only 56% included monitoring and 44% had user interfaces. Only 6 papers described true Digital Twins, with just two reaching prototype or deployment stages.

This study presents a review of digital twin technology literature, sourced from Google Scholar. Our analysis examines digital twins in general, narrows to agricultural applications, and ultimately compares existing agricultural implementations with our proposed solution. We adopt the classification framework developed by Pylianidis et al. (2021), which distinguishes between *concepts* (theoretical system architectures) and *prototypes* (implemented systems for specific use cases). This classification provides a structured approach for evaluating the maturity and practical applicability of digital twin solutions in agriculture.

2.1. Digital twin approaches in non-agricultural sectors

The non-agricultural sector has witnessed a significant evolution of digital twin frameworks over the past decade. Zheng and Sivabalan (2020) advanced manufacturing applications with a Cyber-Physical System prototype architecture employing a tri-model approach that combined digital, computational, and graph-based models. However, such physical model-driven approaches prove challenging in agricultural contexts due to domain complexity. Masison et al. (2021) subsequently introduced a conceptual data-driven architecture using a global shared state for digital twin models. This enhanced modularity by eliminating direct dependencies between components but introduced potential data bottlenecks and security vulnerabilities. Building on these developments, Redeker et al. (2021) created a conceptual microservices platform utilising the Asset Administration Shell standard for Industry 4.0 interoperability. Despite standardisation benefits, this required significant infrastructure investment without clear AI service integration methodology.

As the field progressed, AboElHassan et al. (2023) developed a conceptual general-purpose digital twin platform emphasising digital shadows through distributed systems with publish–subscribe communications. Yet this approach lacked the bidirectional communication capabilities essential for active intervention in physical systems. That same year, Haghshenas et al. (2023) implemented a specialised prototype of digital twin for wind turbines incorporating sensors, 3D models, and augmented reality. However, it critically failed to account for external weather conditions and employed 3D modelling approaches impractical for agricultural applications. Ogunsoto et al. (2025) introduced a conceptual digital supply chain twin framework specifically

[1] Technology Readiness Level (TRL) is a scale from 1 to 9 that measures the maturity of technology development, where TRL 4.8 indicates transition from laboratory validation (TRL 4) to validation in relevant environments (TRL 5).

designed for resilience and recovery prediction during natural disasters but assumes single-vendor scenarios. Further complexity emerged with Lan et al. (2023)'s conceptual five-layered architecture integrating camera sensors and water quality sensors in a fish-type digital twin, where they use reinforcement learning to dispense food. In infrastructure monitoring, Fawad et al. (2025) developed an Immersive Bridge Digital Twin Platform (IBDTP) integrating Scan-to-BIM technology with Augmented Reality for structural health monitoring. The framework's heavy reliance on 3D scanning technology and AR hardware presents cost and complexity barriers for widespread agricultural adoption. Xu et al. (2024) presented a data-driven prototype for production line performance monitoring. Manufacturing machine states from Manufacturing Execution Systems (MES) were classified into fine-grained sub-states enabling comprehensive cycle time variability analysis. Still, the manufacturing-specific metrics may limit direct transferability to outdoor farming contexts.

More innovations include Robles et al. (2023)'s OpenTwins, a prototype open-source framework utilising Eclipse Hono, Eclipse Ditto, and Apache Kafka with Grafana dashboards. Despite its commendable open-source nature, this framework proved resource-intensive for agricultural applications with unnecessary 3D visualisation overhead and insufficient object storage capabilities. Most recently, Tao et al. (2024)'s makeTwin conceptual platform introduced replaceable modules and low-code tools aimed at democratising digital twin development. While promising, it remains in its early stages and needs significant maturation from an engineering software view. The platform lacks concrete implementation details, though its low-code approach with flexible customisation options could be particularly valuable for agricultural applications if achieved.

2.2. Digital twin approaches in agriculture and our proposed architecture

In the smart farming domain, Skobelev et al. (2018) proposed applying swarm concepts to agricultural contexts, where autonomous individuals collaborate to achieve collective intelligence. Their approach used unmanned aerial vehicles (UAVs) working together to complete tasks more efficiently. Building on this, Budaev et al. (2018) developed an autonomous multi-agent system functioning as a digital ecosystem of smart services. Their architecture employed ontology-driven knowledge bases with a virtual "round table" for coordinated decision-making among different services communicating through a service bus. However, ontology-driven systems encounter significant limitations when discovering hidden patterns within dynamic agricultural environments, where ontological representations may be insufficient for capturing the complexity of open-field conditions (Goldstein et al., 2021; Thomopoulos et al., 2013). Similarly, Kalyani and Collier (2023) constructed a multi-agent system incorporating a microservices layer that abstracts external services into standardised formats based on agreed semantic layer ontologies. While multi-agent and microservices architectures provide substantial benefits, they introduce considerable complexity overhead in communication protocols, reasoning processes, and coordination mechanisms that must be carefully balanced against their advantages (Blinowski et al., 2022; Luzolo et al., 2024).

In the agricultural digital twin domain, Attaran and Celik (2023) discuss the use of digital twins for tasks like farm management, weather modelling, and soil monitoring, but these systems are digital twin shadows. Similarly, Ariesen-Verschuur et al. (2022) and Howard et al. (2020) identify key variables for digital twins in greenhouse farming, like soil, water, and diseases. However, their approach does not integrate tools for autonomous decision-making and is limited to greenhouse-controlled environments. Other frameworks, such as Sung and Kim (2022), propose a high level architecture concept for an agricultural digital twin that is slightly generic for the agriculture domain. More advanced implementations include Mitsanis et al. (2024), who developed a 3D functional plant modelling framework integrating Functional Structural Plant Modelling (FSPM) with 3D plant phenotyping, though remaining largely conceptual. Lang et al. (2025) demonstrated the first complete digital twin greenhouse system for tomato harvesting optimisation using multi-view scanning and reinforcement learning, achieving 34.95% efficiency improvements but limited to controlled environments. A notable breakthrough in photorealistic simulation was achieved by Mirbod et al. (2025), who developed a digital twin of a commercial strawberry farm using procedural modelling techniques and NVIDIA's Omniverse platform. While their approach demonstrated impressive simulation-to-real transfer capabilities, their solution is focused on single-task optimisation rather than comprehensive farm management. Practical prototypes, such as the system developed by Gomes Alves et al. (2019), build a digital twin shadow using Orion, Draco, MySQL and Grafana, among others. This can be useful for a small experiment but does not seem to be scalable, at least in the visualisation level. Similarly, Chaux et al. (2021) propose a system for greenhouses that models energy and water use, but their approach is narrowly focused and is also less applicable to open-field farming.

We can infer that agricultural digital twins are less advanced compared to non-agricultural digital twins, encompassing mostly controlled environments and digital twin shadow models. Our proposed agricultural digital twin platform addresses key limitations identified in existing solutions. Unlike heavy-weighted frameworks such as those proposed by Robles et al. (2023) and Lan et al. (2023), our data-driven architecture handles lightweight technologies and prioritises a centralised shared data warehouse that eliminates redundancy while improving data governance and integration and removes the need of complex model-based approaches. Where Purcell and Neubauer (2023) found only 37% of agricultural solutions implemented true digital twins with bi-directional data flow, our platform enables full bidirectional communication between physical and virtual environments, supporting autonomous decision-making based on historical data analysis. We diverge from the complex ontology-driven approaches of Budaev et al. (2018) that struggle with incomplete agricultural knowledge, instead employing a data-driven strategy utilising tabular, time series, and imagery data for forecasting and monitoring. Microservices and agent-based architectures are designed for large-scale applications, introducing unnecessary complexity for open-field agriculture. These approaches may work well in industries with standardised processes, abundant data, and well-defined roles, but they might be less suitable to agriculture's uncertain, sparse, and difficult-to-model environment. This approach avoids the need for complex and resource-intensive real-time systems. Instead, it optimises for open-field agriculture's naturally slower pace using mainly batch processing capabilities and long-running tasks, with near real-time features reserved for scenarios such as weather alerts, SCADA monitoring, irrigation management, and equipment monitoring. While Tagarakis et al. (2024) found only two agricultural digital twins reaching prototype or deployment stages, our scalable open-source platform is designed to accommodate both small independent holdings and large agricultural operations.

3. Conceptual architecture

In this section we present the architecture of the proposed platform (see Fig. 1). It consists of a physical asset, from which the raw data is extracted; the *Digital Twin Environment*, where this data is processed; the consumption layer, where the valuable processed information is shown to the users in different ways.

Physical asset. This layer represents the external sources where the data that feeds each digital twin come from. Data can be open or proprietary data, structured, unstructured or semi-structured, and may be very heterogeneous.

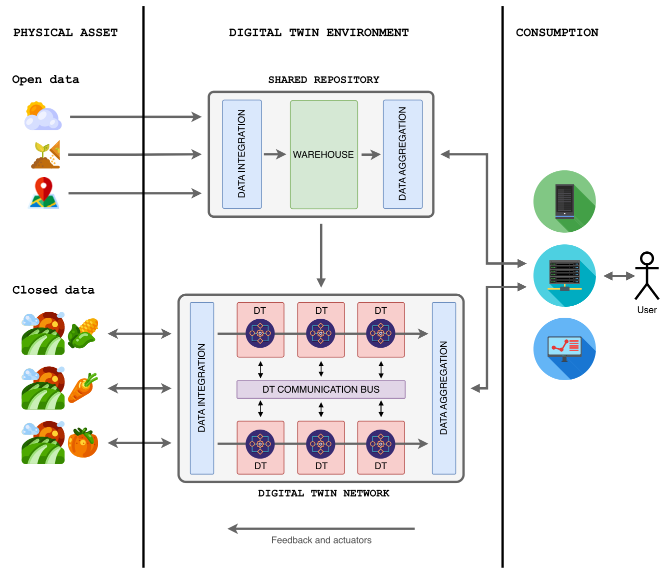

Fig. 1. Conceptual architecture.

Digital twin environment. The *Digital Twin Environment* represents the system infrastructure and it is the space were the different digital twins coexist and interact with themselves and with the outer world (Grieves and Vickers, 2017). The layer obtains data from the physical asset layer in order to feed the different digital twins. We have two main components, the *Shared Repository* and the *Digital Twin Network*. The *Shared Repository* is a global storage that contains all the open data that can be shared among the digital twins. This includes open data for agricultural tasks such as weather variables (e.g., temperature, humidity, wind), NVDI indexes, satellite imagery, economic data, topographic or even disease, pest or fertiliser data, which are stored inside the *Warehouse*. The *Digital Twin Network* is a collection of Digital Twin Instances that are configured for different or the same farm. Each instance gets the required data from the *Shared Repository* and combines it with the information provided by their local sensors and other suitable data. With it, they can build their AI models to make recommendations, forecasts and simulate various scenarios. To communicate with each other and share useful processed information, digital twins use the *DT Communication Bus*.

Consumption. The consumption layer gets data from the *Digital Twin Environment*. It translates the information produced by the digital twins and provides insights including real-time information, events, on-demand simulations and predictions based on AI models. Additionally, it can be used to send instructions to the physical asset to schedule activities like irrigation, giving an interface to communicate with the digital twin physical asset (*Feedback and Actuators* in the Figure).

4. Detailed design

Fig. 2 illustrates the essential components within the *Digital Twin Environment*, encompassing the *Shared Repository* and the network of digital twins. The architecture follows an event-driven architecture (Sivamani et al., 2014) for updating digital twin states based on events. It also uses a pipe-and-filter architecture (Wulf et al., 2016) to ingest and manage data flow and processing in both *Shared Repository* and Digital Twins through the *Data Pipelines* component. Also, the *Digital Twin Network* utilises a broker pattern (Tržec et al., 2022) through the *DT Communication Bus* for asynchronous communication between the digital twins, maintaining privacy and saving resources. As introduced in Sections 2 and 3, a platform such as the one we are proposing must fulfil the following characteristics in order to be fit-for-purpose according to Tao et al. (2022).

Data extraction and integration. The *Shared Repository* is the central hub for managing open data, including structured, semi-structured, and unstructured data stored in the Warehouse. Open data comes from Information Technology (IT) systems such as web services, databases, cloud platforms, or external Enterprise Resource Planning (ERP) systems. Examples include Normalised Difference Vegetation Index (NDVI),[2] weather information, and imagery. This data is ingested in batches through *Data Pipelines* using a pipe-and-filter architecture pattern, with the Orchestrator managing schedules and tracking workflows. Once ingested, raw data is stored, enriched, and transformed in the Warehouse before being distributed to the digital twins in the *Digital Twin Network*. Raw data is kept in its original format, while processed data is transformed into a common format, providing higher performance and compression. The *Digital Twin Network* provides a shared environment for hosting digital twins, which are virtual replicas of physical assets. Digital twins rely on two types of data: open data from the *Shared Repository* and proprietary data from the physical assets. Open data is processed in the *Shared Repository* and sent asynchronously to digital twins via an event-driven architecture. For example, if two twins require weather data from the same station, the first twin triggers its retrieval and caching in the *Shared Repository*, allowing the second twin to access the already processed data. This approach ensures optimised resource use, reduced latency, and consistent shared data. Proprietary data originates from IT systems or Operational Technology (OT) systems that supervise physical processes, devices, and infrastructures (e.g. SCADA[3] systems, PLCs[4] or sensory systems). It is collected in real-time from physical assets via the *Communication Bus*, which processes and delivers the data to the specific *DT Storage* of each twin. Both open and proprietary data are stored in the *DT Storage*, a high-speed solution that maintains historical records, analysis results, predictions, and simulation outputs.

Automation and optimisation. In order to extract data from the physical asset using the *Data Pipelines*, several adapters must be used in order to make the extraction and integration process automatic. Heterogeneity in the data sources is managed through the flexibility of workflows. Each workflow can create its own adapter to ingest data independently of its format or type. External systems must be able to communicate future changes to synchronise the system adapters as other systems evolve. The data flow should be completely automated, from the extraction to the storage in the *Shared Repository* and the *DT Storage*. The *Orchestrator* is responsible for this task, and it will use the *Data Pipelines* to control the data flow.

Modelling. Digital twins resemble a physical holding. Every one of them can store a number of finite models, as many as it takes to better mimic the holding lifecycle, and thus, predicting its behaviour more accurately. As we said, it is prepared for a data-driven approach, where models are created and trained based on historical open and proprietary data stored in *DT Storage*, and also using knowledge about the crop and its lifecycle. The platform embraces data-driven model approaches like statistical, rule-based, machine learning models or even reinforcement learning models. Adding a new model follows a workflow where developers first create and train models offline using their preferred tools and frameworks. These models are then registered in the platform's *Model Registry* along with their metadata and versioning information. Once registered, models become available throughout the platform for offline testing, on-demand predictions, and integration into Digital Twin *Data Pipelines* for retraining, forecasting and decision support. This way, it enables developers to incorporate their preferred modelling approaches while maintaining consistency.

[2] NDVI measures the amount and vigour of vegetation on the land surface https://ipad.fas.usda.gov/cropexplorer/Definitions/spotveg.htm

[3] SCADA, or Supervisory Control and Data Acquisition, is a system used to monitor and control industrial processes, such as power plants, manufacturing, and infrastructure, by collecting and analysing real-time data.

[4] PLCs, or Programmable Logic Controllers, are rugged computers used to automate electromechanical processes in industrial settings.

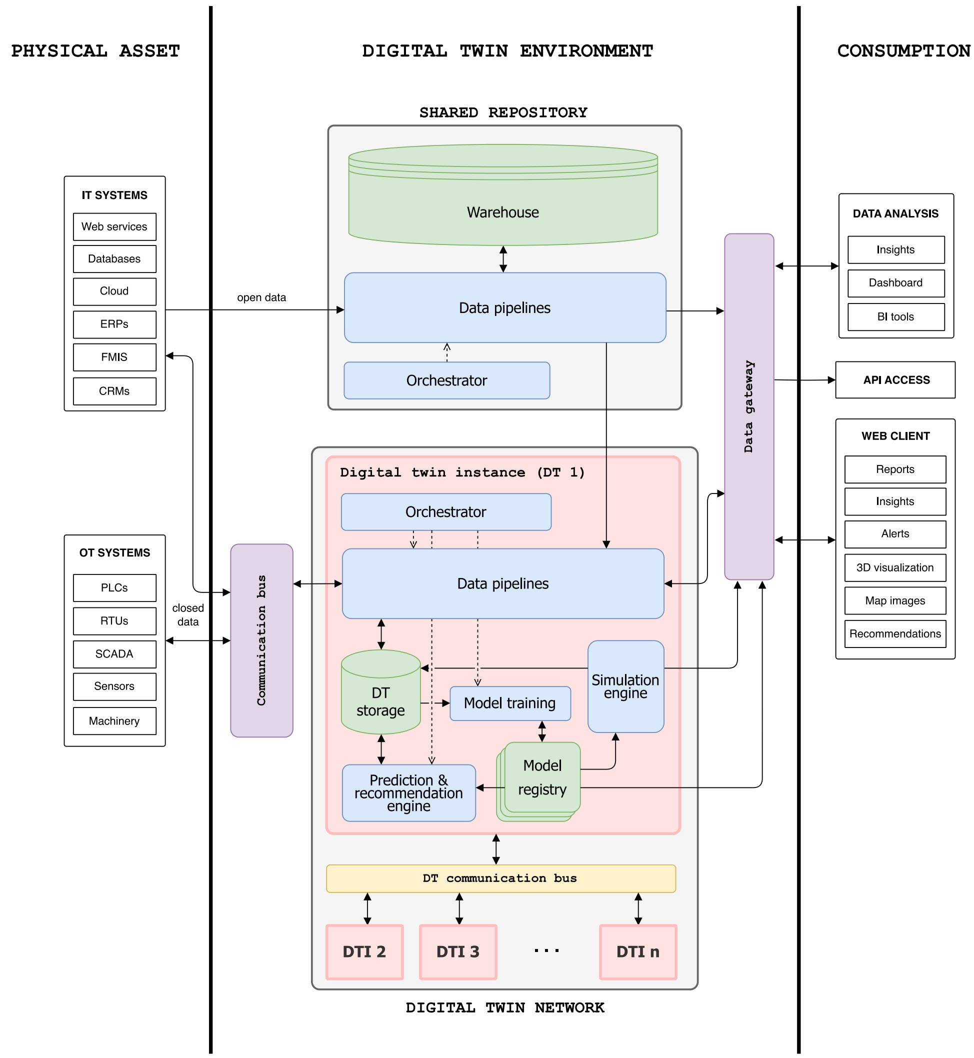

Fig. 2. Complete design.

Real-time monitoring and alerts. As we explained in Section 1, agricultural processes evolve slowly over time. Nevertheless, some external processes like natural disasters or plagues could require a faster response. This real-time data received from the holding sensors or other type of data is sent to the user and stored in the *DT Storage*. The *Prediction & Recommendation Engine* will use this data to alert the user if something goes wrong —in a synchronous way using the event-driven architecture—, taking into account the historical data and the models stored in the *Model Registry*.

Predictions and recommendations. The digital twin must be able to predict future crop yields, disease outbreaks, and even market trends. The *Prediction & Recommendation Engine* is responsible for this task, and it will use the appropriate model of the digital twin *Model Registry* to make the best decision. The results will be stored in the *DT Storage* and will be available for the user to consult.

Feedback and actuators. The digital twin must be able to send feedback to the physical asset in order to perform certain actions using events. The component capable of taking decisions autonomously is the *Prediction & Recommendation Engine*. This engine will use the appropriate model of the digital twin *Model Registry* in order to make the best decision, and send feedback to the actuators of the holding. For instance, if irrigation was programmed today by the digital twin and a storm suddenly starts, the digital twin should be able to stop the irrigation process by sending feedback to the irrigation system to stop it. Likewise, based on historical data, the digital twin could predict that a certain amount of water is needed to be irrigated today. It can then send feedback to the irrigation system to start the process.

Simulations. The digital twin and users must be able to simulate different scenarios based on the historical data and the models stored in the *Model Registry*. These simulations could be used to predict the crop yield, the best moment to harvest, the best moment to irrigate, or the best moment to apply a certain pesticide. They can even run a parametrised simulation to test extreme scenarios different from the typical scenarios. The *Simulation Engine* is in charge of this task. It will use the *Model Registry* to select the appropriate model to simulate the scenario. Results will be stored in the *DT Storage* and will be available for the user to consult.

Interoperability. The digital twin must be able to communicate with other digital twins inside the *Digital Twin Network* in order to share information of events that could affect its normal functioning. For instance, if a digital twin is located near a digital twin that is going to apply a certain pesticide, the first one must be able to know it and take the necessary precautions if some unexpected leak occurs that could affect near digital twins. Other examples include flooding, outbreaks of diseases or pests, fires, irrigation problems, drainage issues, and more. The *DT Communication Bus* handles this task. It will be able to send and receive information from other digital twins using the broker pattern. This pattern enables efficient asynchronous and anonymous

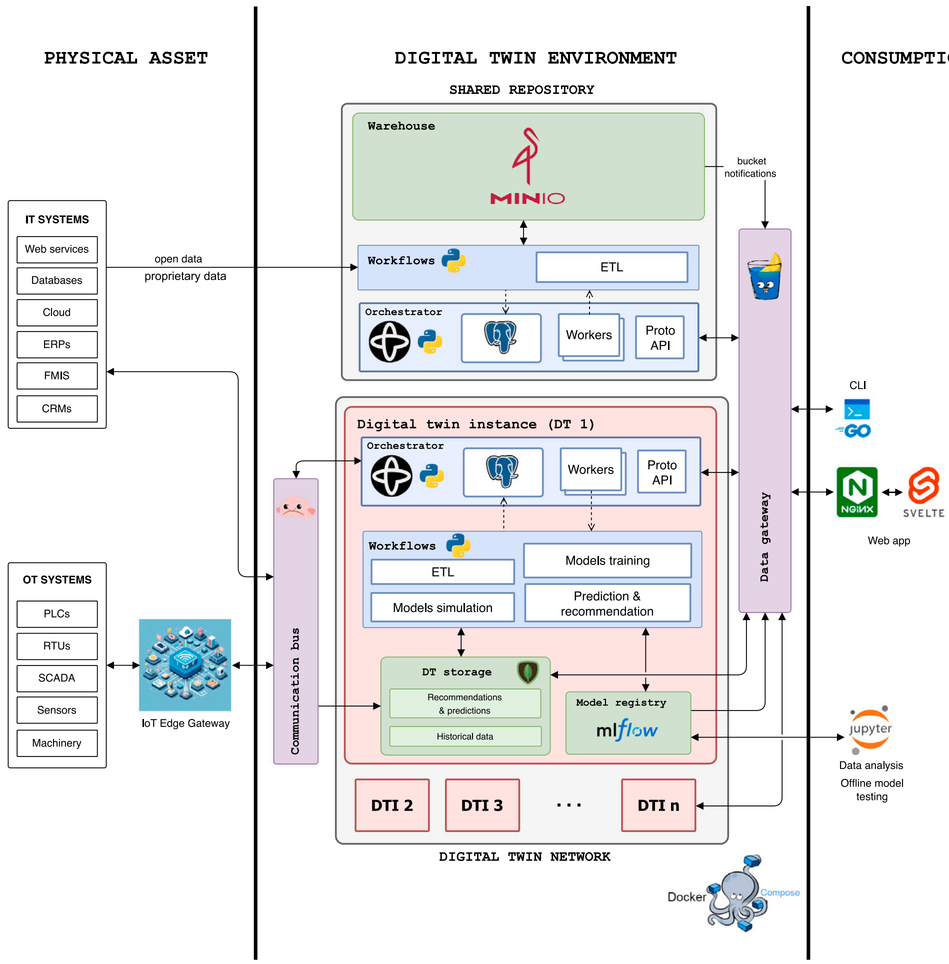

Fig. 3. System implementation.

communication between digital twins. This is the most efficient way to do it, as the data is not duplicated in each digital twin and proprietary information is kept private in the digital twin that owns it. Processing is done only once, in every digital twin, and then it is shared with the rest of the digital twins that need it.

5. System implementation

This section builds upon the design presented and justified in Section 4 by grounding it in a real technological framework. This section is essential not only to enable the experimental validation of our proposal, which is conducted in Section 6, but also to ensure the reproducibility of results and facilitate reuse by other researchers in the future. All technologies used in our solution are open source. Fig. 3 illustrates the system implementation, and the prototype implementation is available in the Zenodo at Béjar Hernández et al. (2025). This ensures that other researchers will be able to examine in detail every aspect of the technologies we use.

The *Shared Repository* is a single process that takes control over the bulk of data processing operations. It takes the role of a data lake, in which the open data and proprietary data are processed and stored. It must be an object storage system capable of retaining all types of data: structured, semi-structured, and unstructured data. We selected Minio,[5] an open source, Amazon S3-compatible system that is lightweight and easy to use. This data lake serves as the central information hub of the system. It is compartmentalised in different buckets and folders to separate incoming and processed data. Our system allocates a dedicated S3 bucket for each digital twin. Additionally, there is a shared bucket containing common open data such as weather information, NDVI values, and satellite imagery accessible to all digital twins. Each bucket implements the Medallion architecture, a data organisation approach that structures data processing through three distinct layers: bronze (raw, unprocessed data), silver (cleaned and validated data), and gold (business-ready, aggregated data). Processed data will be stored in the Parquet format[6] due to its compression and performance. To process this data, we need something lightweight and easy to use. That is why we decided to build *Data Pipelines* as simple Python scripts that can be executed in a containerised environment. These scripts will be monitored, scheduled and executed by the *Orchestrator*. For this component, we are going to use Temporal,[7] a workflow management system that allows us to create, schedule and monitor *Data Pipelines*. It is quite simple, as it will let us control the data flow of the different pipelines just by adding some Python decorators to the scripts functions. In Temporal workflows, PostgreSQL serves as a persistence backend, storing workflow metadata, execution histories, and state transitions. It ensures durability, reliability, and allows Temporal to efficiently manage long-running and complex workflows.

The *Digital Twin Network* is a multi-process component where every digital twin works separately, as individual agents, that interact with

[5] https://min.io

[6] Apache Parquet is an open source, column-oriented data file format designed for efficient data storage and retrieval.

[7] https://temporal.io

each other. The digital twins are containerised applications that can be deployed in a cloud environment or in a local machine. We are going to use Docker[8] to containerise the digital twins and Docker-compose to orchestrate them. In order to keep things simple, they are going to use the same *Data Pipelines* and Orchestrator components, using Python scripts and Temporal to manage the data flow. These flows primarily focus on ingesting necessary and clean data from the *Shared Repository*. They run the training of the models and make predictions, simulations and recommendations. It is completely open, so the developers can create their own models and deploy them in the *Digital Twin Network*. This environment provides custom necessary libraries such as predefined adapters to communicate with external systems. However, internal communications are pre-established and the developer cannot modify them. Historical data, recommendations and simulation results will be stored in a no-SQL database, MongoDB,[9] because it is well known and gives the ability to store data incrementally without the need of a predefined schema and can act like an event dispatcher when new data is added. To avoid bottlenecks, sensor data will be obtained from the physical asset and stored in MongoDB in time series collections, to be able to perform quick queries and analysis. As the frequency of the data is not high, it is not necessary to use a time series database like InfluxDB,[10] which would cause unnecessary overhead in the system. This data is pushed from the physical asset to the digital twin through the *Communication Bus* component, using Benthos[11] which is a simple Go static binary that uses a YAML configuration file to route messages between different systems, processing them on the fly. Benthos is very lightweight with a low memory footprint and allows us to integrate custom plug-ins to adapt it to our needs. It is also used to send feedback messages generated by the digital twin to the physical asset to perform certain actions.

Models are trained using historical digital twin data offline through tools like Jupyter notebooks, R, or any preferred platform. These models and their different versions are stored in the *Model Registry*, in this case, Mlflow,[12] a platform to manage the complete machine learning life cycle, including experimentation, reproducibility and deployment. It is a good tool to track the models and its versions, as well as testing and deploying them in the *Digital Twin Network*. Mlflow also allows storing and managing custom models for application extensibility.[13] The models are served through a REST API,[14] which Mlflow provides.

The digital twins can communicate with each other using a combination of temporal workflows and MongoDB Change Streams.[15] There is an observer workflow for each digital twin that watches different collections looking for significant changes in sensor data, incoming weather conditions, or rapidly spreading pests. For instance, when the observer realises, using a MongoDB change stream, that a plague of one digital twin is spreading fast, then it automatically sends feedback to the physical systems to treat them and sends a warning to near digital twins. Current implementation is rule-based but planned to evolve to AI-based decisions. The *Data Gateway* is a simple and fast REST API that receives the messages from the digital twins and sends them to the corresponding digital twin. It is also responsible for connecting each digital twin with the *Consumption Layer*. It is built with Gin,[16] a lightweight web framework for Go, which is easy to use and maintained by Google. Furthermore, a Go terminal client is provided to do quick write and read operations of the system. This client is only available for the administrators.

The *Consumption Layer* is a simple web application that displays processed information to the user. It can be built using any web framework, but we chose Svelte.[17] The system implements a simple dashboard with several KPIs, a map showing the location of the digital twins, along with up-to-date information, predictions, recommendations, and simulation results. The developer also has a private entry point to the system, using a Go client, to do quick write and read operations of the system.

5.1. Deployment

All components are compatible with both Docker and Kubernetes. For our prototype, we use Docker with Docker Compose, which is ideal for small-sized projects and simple deployments. As the system scales, the natural progression is to Kubernetes—an open-source platform for automating deployment, scaling, and management of containerised applications. This containerisation approach enables easy deployment anywhere, including cloud environments. To monitor system performance and quickly identify potential issues, we have integrated Grafana[18] and Prometheus[19] for comprehensive observability and decision-making support.

The parts of the system most susceptible to becoming serious bottlenecks are the *Workflows*. The Minio Warehouse can scale horizontally using a concept called Server Pools[20] and the *Data Gateway* can leverage asynchronous communication, multiple instances, and load balancing to prevent overloading. The *Workflows* can be scaled horizontally and vertically, as Temporal is designed to be scalable and fault-tolerant, capable of handling long-running and parallel workflows.

6. System validation with two digital twin prototypes

To validate our proposal's applicability, we have developed an example scenario featuring two adjacent holdings with pistachio crops (Béjar Hernández et al., 2025). Our platform, based on the autonomous digital twins concept (Verdouw et al., 2021), employs human-supervised models capable of making autonomous decisions and sending commands to actuators. To test these capabilities, we are using a combination of data sources: historical NDVI and weather data from REST endpoints, synthetic activities and harvest data as files, and random sensor data from a public MQTT endpoint. This setup allows us to train a harvest prediction model weekly and enable autonomous decision-making. To showcase the communication between digital twins and their ability to respond to environmental changes, we have simulated a flood scenario. In this simulation, the digital twins detect a potential flood and automatically suspend irrigation for 12 h, demonstrating their capability to adapt to unforeseen circumstances and protect the crops. To create a digital twin in our platform, the following steps should be followed:

1. Create a new user through the Go CLI by providing username, password, and role permissions.
2. Navigate to the dashboard and click the add button located at the bottom of the sidebar.
3. Add a new digital twin as a GeoJSON object to leverage MongoDB's geospatial query capabilities for location-based functionality.

[8] https://www.docker.com
[9] https://www.mongodb.com
[10] https://www.influxdata.com
[11] https://www.benthos.dev
[12] https://mlflow.org
[13] https://mlflow.org/docs/latest/traditional-ml/creating-custom-pyfunc/notebooks/introduction
[14] REST API is a interface that two computer systems use to exchange information securely over the internet https://aws.amazon.com/what-is/restful-api.
[15] https://www.mongodb.com/docs/manual/changeStreams/
[16] https://gin-gonic.com
[17] https://svelte.dev
[18] https://grafana.com
[19] https://prometheus.io
[20] Minio server pools group servers for distributed storage, ensuring scalability, redundancy, and high availability https://blog.min.io/server-pools-streamline-storage-operations/

4. Create the digital twin and execute open data workflows either manually or through scheduled automation to obtain historical weather and NDVI data.
5. Configure a new stream ETL pipeline in Benthos to connect SCADA sensors or other data sources, which will automatically appear in the dashboard for real-time monitoring and alerts.
6. Develop custom ETL processes for phytosanitary treatments, farming activities, or harvesting data (structured or unstructured), allowing farmers to upload information periodically on-demand through the dashboard interface, or automatically from a external source, orchestrated by Temporal.
7. Create and train AI models using the collected data with Mlflow integration. Once the model is ready, deploy it using Mlflow serve as a REST API for inference operations.
8. Implement a weekly Temporal workflow for automated predictions and establish an additional workflow for periodic model retraining to maintain accuracy over time.
9. Utilise the simulation tab to create new scenarios using Temporal long-running workflows, with the ability to control and modify simulations dynamically through Temporal signals.

Two adjacent pistachio digital twins were generated to evaluate inter-twin communication in close proximity. The creation protocol incorporates a workflow that identifies the nearest meteorological station identifier for daily weather data acquisition. To facilitate the development of artificial intelligence models, data ingestion is essential to ensure accurate modelling and real-time decision-making for both digital twins. Three data ingestion methodologies were implemented: manual input via the web user interface, automated ingestion utilising the workflows Orchestrator, and external system data push through the *Communication Bus*. This study focuses on the first two methodologies.

Initially, an external REST API service was integrated to ingest historical NDVI and meteorological data. This process employs an asynchronous workflow scheduled for weekly execution to incorporate new data. For experimental purposes, the workflow was manually triggered to expedite the process. Subsequently, example files containing harvest and holding activities data, in JSON and Excel formats, were uploaded. These files undergo processing via distinct temporal workflows, with the resulting information stored in the corresponding digital twin database within the *DT Storage*.

Following data integration, a simulated temperature sensor was incorporated to augment the digital twin with sensor data. The *Communication Bus* was utilised to establish a connection with a public Mosquitto MQTT endpoint (https://test.mosquitto.org). This endpoint is served by Eclipse Mosquitto and used by the community to push and pull data without authentication for testing purposes. Numerical values were filtered within specified parameters, processed, and stored in the digital twin database. Real-time database streaming functionality transmitted changes to the frontend interface.

The resulting monitoring page (Fig. 4) displays the newly integrated information. Historical harvest data is presented in the upper section, with static digital twin information positioned on the left. A map illustrating NDVI values, sourced from the Copernicus Data Space Ecosystem,[21] is included with dynamic modification capabilities. A chart depicts the previous 30 days of data, while cards below show real-time sensor streams, currently limited to temperature readings.

With the digital twin database populated, a simple machine learning model was developed to forecast harvest dates and quantities for the current year. Data used included year, day, month, precipitations, temperatures, wind speed, wind gusts, NDVI and the yield. A Random Forest Regressor algorithm was trained offline to forecast yield, month and day. It was then registered in the *Model Registry* server. We could have used any kind of model, including LSTM, SVM or even transformers, but we prefer to focus on the architecture. This model is accessible via a REST endpoint for testing and simulation purposes. A workflow was established to periodically retrain the model with new data and forecast optimal harvest dates and expected yields. The most recent model version is retrieved from the *Model Registry*, with model hyper-parameters, metrics, and version information visible in Fig. 6.

The workflow, which can be manually initiated, is illustrated in Fig. 5, delineating all constituent steps. Resultant data is stored in the digital twin database and visualised on the monitoring page (Fig. 4). The prediction is marked by a vertical golden line on the chart, indicating the optimal harvest time and expected yield.

Digital twin simulations were implemented as long-running workflows, utilising a simplified rule-based model to emulate growth and development over time. The simulation integrates time, digital twin, meteorological, and phytopathological components.

Simulations are started via a user interface that allows for setting initial conditions. Fig. 7 displays an active simulation, showing control functions, a visual representation of the digital twin's state, and a real-time JSON state representation. The interface also presents harvest outcomes graphically and tabulates key metrics including tree viability and yield data. The *Workflows* component enables real-time interaction with the simulation through signal mechanisms,[22] allowing dynamic parameter modifications. Simulation states and results are stored in the digital twin database and can be retrieved in real-time upon data insertion.

A flooding simulation was implemented to evaluate inter-twin communication and physical asset feedback mechanisms. Using the *Communication Bus*, a virtual water height sensor connected to a Mosquitto MQTT endpoint triggered automated responses when levels exceeded 60 mm. The system analysed threats in real-time, initiating actions such as suspending irrigation for 12 h and alerting adjacent digital twins. Actions were logged on the web application's Activities page, categorised as either automatic (system-initiated) or manual (user-initiated) commands, with status tracking. This prototype demonstrates the capacity for digital twins to autonomously respond to environmental stimuli and communicate potential threats to neighbouring entities.

6.1. Results

We monitored system performance through Docker over three deployment-to-validation cycles, capturing average metrics shown in Table 1. The platform consumed approximately 60% CPU and 2.5 GB RAM across all services on average. Benthos demanded the highest CPU usage due to its continuous data streaming operations with Mosquitto. Memory consumption was dominated by Temporal, which hosts Python data pipelines (mostly inactive except for long-running workflows), and MLflow, which requires significant memory for its model registry, artifact storage, and visualisation capabilities. We could try to lower this memory footprint with future versions.

The system illustrates potential for automated threat response and inter-twin communication in agricultural settings. Future implementations could incorporate machine learning models to enhance predictive capabilities and decision-making accuracy.

7. Conclusion

Our solution addresses four critical gaps in agricultural digital twin technology. We tackle the affordability barrier for European agriculture's predominantly small farms by avoiding resource-intensive infrastructure, which exceed agricultural needs where real-time processing is less critical due to open-field farming's natural pace. Addressing the modelling complexity challenge, our data-driven approach

[21] https://documentation.dataspace.copernicus.eu/APIs/SentinelHub/OGC/WMS.html

[22] Signals are a mechanism for sending asynchronous messages to workflows https://docs.temporal.io/encyclopedia/workflow-message-passing.

Fig. 4. Guide: monitoring page with new NDVI and weather information to show.

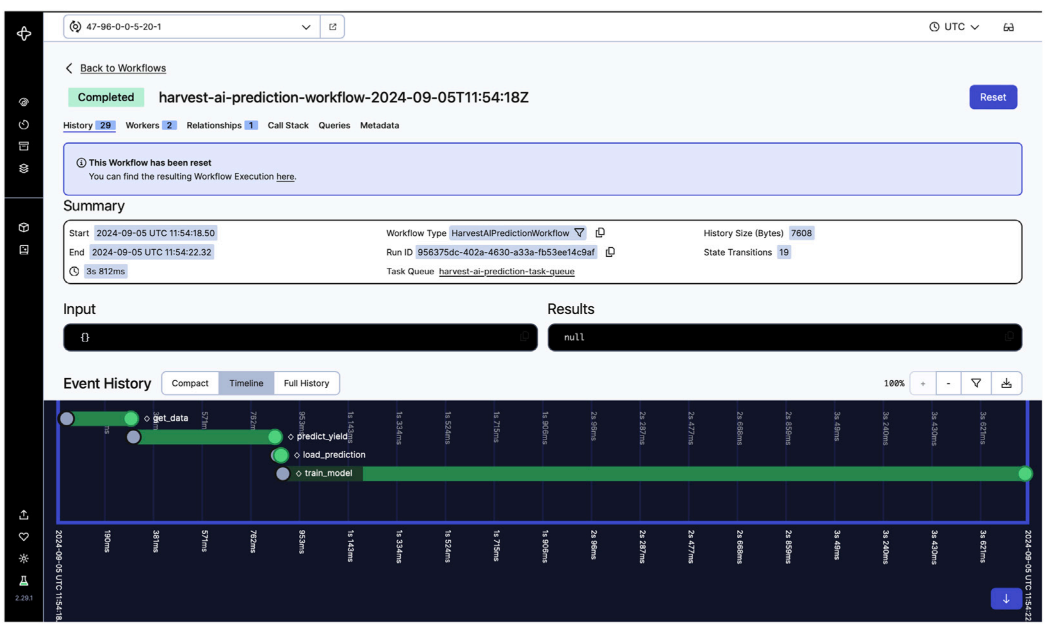

Fig. 5. Guide: Prediction temporal workflow, execute every week or on demand, where the different stages can be seen.

replaces physical modelling approaches that struggle with open-field agriculture's irregular external conditions. Unlike controlled manufacturing environments, farming's unpredictable external factors make traditional modelling extremely challenging, which our solution handles through simplicity and extensibility while maintaining core functionality. We tackle the scarcity of open-source solutions adapted to agriculture's slower innovation pace, providing accessible tools for gradual technology adoption. Most critically, our open-source platform bridges the significant gap between technology and farmers by providing developers with a foundation to create accessible solutions for non-technical agricultural stakeholders, fostering community-driven improvements that benefit the entire agricultural sector.

The case study implementation demonstrated the platform's efficiency on modest hardware configurations (2 vCPUs, 4 GB RAM), operating with average CPU usage of 60%, RAM consumption of 2.5 GB, and deployment time of approximately one minute. These performance metrics validate our approach viability for small holdings with limited resources, effectively lowering adoption barriers that have historically hampered digital transformation in open-field agriculture.

Future work should focus on several key areas:

Enhanced data integration. Exploring lightweight natural language processing AI to intelligently scrape and interpret agricultural texts would improve the system's ability to incorporate diverse data sources. Additionally, implementing a knowledge database using graph or vector

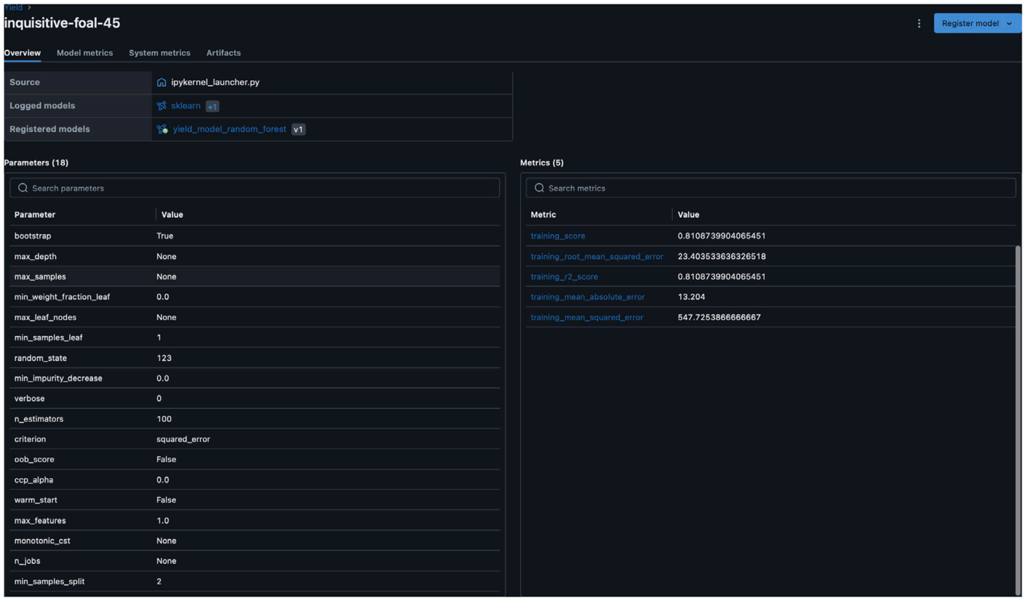

Fig. 6. Guide: Mlflow web user interface, where model properties can be tracked and monitored.

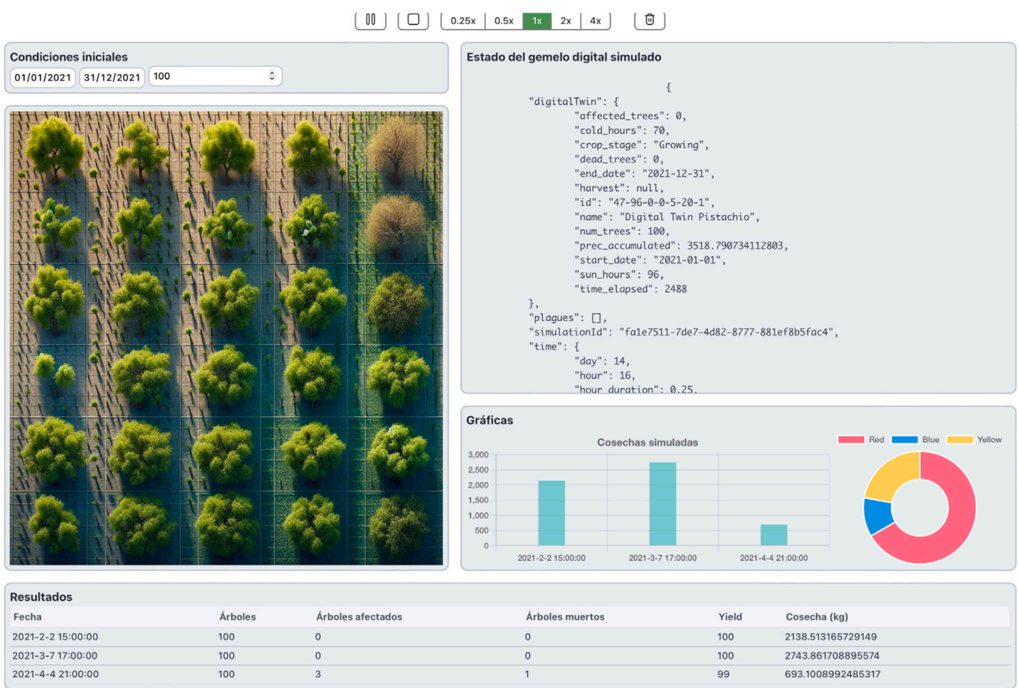

Fig. 7. Guide: New simulation of digital twin with id *47-96-0-0-5-20-1*.

databases would enable semantic relationships between agricultural entities and concepts.

Vertical interlinking and sub-twin creation. While our current system focuses on horizontal interlinking (Digital Twin Network), creating digital sub-twins for specific components such as machinery or individual crops would provide more granular insights and precise modelling.

Platform standardisation. Developing a stronger functionality core specifically tailored to agricultural domain requirements would specialise the infrastructure through domain-specific data transformations, agricultural modelling capabilities, and context-aware processing optimised for farming operations. Establishing standardised protocols for digital twin interconnection would enable seamless communication, enhancing interoperability and scalability across Digital Twin Network.

Cloud-based digital-twin-as-a-service model. If the self-hosted model proves successful, developing a cloud-based service would make the technology more accessible to a wider range of users, at least components that need to scale independently, such as *Data pipelines* and *Model registry*.

Table 1
Digital Twin Platform average resource usage using the Docker Desktop application. The platform was deployed using Docker with resources limited to 2 VCPUs and 4GB of RAM, using around 60% CPU and around 2.5GB of RAM on average.

Container name	Avg. CPU %	Avg. RAM. Usage
digital-twin-dev-mflow-1	0.01%	397.7MiB
digital-twin-dev-frontend-1	0.00%	97.94MiB
digital-twin-dev-gin-gonic-1	0.07%	109MiB
digital-twin-temporal-ui	0.04%	15.94MiB
digital-twin-dev-minio-1	0.00%	99.39MiB
digital-twin-dev-postgresql-1	0.09%	317.9MiB
digital-twin-dev-mongo-1	1.35%	252.8MiB
digital-twin-dev-certbot-1	0.00%	25.39MiB
digital-twin-dev-benthos-1	55.46%	173.5MiB
digital-twin-temporal	1.67%	238.6MiB
digital-twin-temporal-worker	2.45%	760.1MiB
TOTAL	**61.14%**	**2488.26MiB**

Progressive data sharing framework. Creating a framework that helps farmers maintain control over their data sharing while digitising operations would address privacy concerns. This could include tiered data-sharing options with clear benefits at each level, robust anonymisation techniques, and trust-building mechanisms.

CRediT authorship contribution statement

Jorge Laguna: Writing – original draft, Validation, Software, Methodology. **Mario E. Suaza-Medina:** Validation, Formal analysis, Data curation. **Rubén Béjar:** Writing – review & editing, Methodology. **Javier Lacasta:** Writing – review & editing, Conceptualization. **F. Javier Zarazaga-Soria:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Funding sources

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

Declaration of competing interest

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

Acknowledgements

This paper is part of the R&D projects T59_23R and PROY_T04_24 supported by the Aragon Regional Government (Spain). The work of Mario Suaza has been partially supported by the Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS 885/2020).

Data availability

Data will be made available on request.

References

AboElHassan, A., Sakr, A.H., Yacout, S., 2023. General purpose digital twin framework using digital shadow and distributed system concepts. Comput. Ind. Eng. 183, 109534. http://dx.doi.org/10.1016/j.cie.2023.109534, URL https://JMASISON2021https://www.sciencedirect.com/science/article/pii/S0360835223005582.

Ariesen-Verschuur, N., Verdouw, C., Tekinerdogan, B., 2022. Digital twins in greenhouse horticulture: A review. Comput. Electron. Agric. 199, 107183. http://dx.doi.org/10.1016/j.compag.2022.107183.

Attaran, M., Celik, B.G., 2023. Digital twin: Benefits, use cases, challenges, and opportunities. Decis. Anal. J. 6, 100165. http://dx.doi.org/10.1016/j.dajour.2023.100165, URL https://www.sciencedirect.com/science/article/pii/S277266222300005X.

Béjar Hernández, R., Lacasta Miguel, J., Laguna Argüelllo, J., Suaza Medina, M.E., Zarazaga Soria, F.J., 2025. Digi2Farm. http://dx.doi.org/10.5281/zenodo.15697049.

Blinowski, G., Ojdowska, A., Przybylek, A., 2022. Monolithic vs. Microservice architecture: A performance and scalability evaluation. IEEE Access 10, http://dx.doi.org/10.1109/ACCESS.2022.3152803, 1–1.

Budaev, D., Lada, A., Simonova, E., Skobelev, P., Travin, V., Yalovenko, O., Voshchuk, G., Zhilyaev, A., 2018. Conceptual design of smart farming solution for precise agriculture. Int. J. Des. Nat. Ecodynamics 13 (3), 307–314. http://dx.doi.org/10.2495/DNE-V13-N3-307-314.

Chamorro-Padial, J., García, R., Gil, R., 2024. A systematic review of open data in agriculture. Comput. Electron. Agric. 219, 108775. http://dx.doi.org/10.1016/j.compag.2024.108775, URL https://www.sciencedirect.com/science/article/pii/S0168169924001662.

Chaux, J.D., Sanchez-Londono, D., Barbieri, G., 2021. A digital twin architecture to optimize productivity within controlled environment agriculture. Appl. Sci. 11 (19), http://dx.doi.org/10.3390/app11198875, URL https://www.mdpi.com/2076-3417/11/19/8875.

De Alwis, S., Hou, Z., Zhang, Y., Na, M.H., Ofoghi, B., Sajjanhar, A., 2022. A survey on smart farming data, applications and techniques. Comput. Ind. 138, 103624. http://dx.doi.org/10.1016/j.compind.2022.103624, URL https://www.sciencedirect.com/science/article/pii/S0166361522000197.

European Union Eurostat, 2023. Farms and farmland in the European union - statistics. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Farms_and_farmland_in_the_European_Union_-_statistics.

Fawad, M., Salamak, M., Chen, Q., Uscilowski, M., Koris, K., Jasinski, M., Lazinski, P., Piotrowski, D., 2025. Development of immersive bridge digital twin platform to facilitate bridge damage assessment and asset model updates. Comput. Ind. 164, 104189. http://dx.doi.org/10.1016/j.compind.2024.104189, URL https://www.sciencedirect.com/science/article/pii/S0166361524001179.

Gabriel, A., Gandorfer, M., 2023. Adoption of digital technologies in agriculture—an inventory in a european small-scale farming region. Precis. Agric. 24, 68–91. http://dx.doi.org/10.1007/s11119-022-09931-1, URL https://link.springer.com/article/10.1007/s11119-022-09931-1.

Goldstein, A., Fink, L., Ravid, G., 2021. A framework for evaluating agricultural ontologies. Sustainability 13 (11), http://dx.doi.org/10.3390/su13116387, URL https://www.mdpi.com/2071-1050/13/11/6387.

Gomes Alves, R., Souza, G., Maia, R., Tran, A., Kamienski, C., Soininen, J.-P., Aquino Junior, P., Lima, F., 2019. A digital twin for smart farming. In: 2019 IEEE Global Humanitarian Technology Conference. GHTC, IEEE, pp. 1–4. http://dx.doi.org/10.1109/GHTC46095.2019.9033075.

Grieves, M., Vickers, J., 2017. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. Transdiscipl. Perspect. Complex Syst.: New Find. Approaches 85–113. http://dx.doi.org/10.1007/978-3-319-38756-7_4.

Haghshenas, A., Hasan, A., Osen, O., Mikalsen, E.T., 2023. Predictive digital twin for offshore wind farms. Energy Informat. 6 (1), 1. http://dx.doi.org/10.1186/s42162-023-00257-4.

Howard, D., Ma, Z., Mazanti Aaslyng, J., Nørregaard Jørgensen, B., 2020. Data architecture for digital twin of commercial greenhouse production. In: 2020 RIVF International Conference on Computing and Communication Technologies. RIVF, pp. 1–7. http://dx.doi.org/10.1109/RIVF48685.2020.9140726.

Hussain, W.A., 2023. 15 best digital twin companies to look for in 2023. https://metaverseinsider.tech/2023/03/08/best-digital-twin-companies/. (Accesed: 02 October 2023).

Idoje, G., Dagiuklas, T., Iqbal, M., 2021. Survey for smart farming technologies: Challenges and issues. Comput. Electr. Eng. 92, 107104. http://dx.doi.org/10.1016/j.compeleceng.2021.107104, URL https://www.sciencedirect.com/science/article/pii/S0045790621001117.

Iranshahi, K., Brun, J., Arnold, T., Sergi, T., Müller, U.C., 2025. Digital twins: Recent advances and future directions in engineering fields. Intell. Syst. Appl. 26, 200516. http://dx.doi.org/10.1016/j.iswa.2025.200516, URL https://www.sciencedirect.com/science/article/pii/S2667305325000420.

Kalyani, Y., Collier, R., 2023. Hypermedia multi-agents, semantic web, and microservices to enhance smart agriculture digital twin*. In: 2023 IEEE International Conference on Pervasive Computing and Communications Workshops and Other Affiliated Events (PerCom Workshops). pp. 170–171. http://dx.doi.org/10.1109/PerComWorkshops56833.2023.10150413.

Lan, H.-Y., Ubina, N.A., Cheng, S.-C., Lin, S.-S., Huang, C.-T., 2023. Digital twin architecture evaluation for intelligent fish farm management using modified analytic hierarchy process. Appl. Sci. 13 (1), http://dx.doi.org/10.3390/app13010141, URL https://www.mdpi.com/2076-3417/13/1/141.

Lang, Y., Zhang, Y., Sun, T., Chai, X., Zhang, N., 2025. Digital twin-driven system for efficient tomato harvesting in greenhouses. Comput. Electron. Agric. 236, 110451. http://dx.doi.org/10.1016/j.compag.2025.110451, URL https://www.sciencedirect.com/science/article/pii/S0168169925005575.

Luzolo, P.H., Elrawashdeh, Z., Tchappi, I., Galland, S., Outay, F., 2024. Combining multi-agent systems and artificial intelligence of things: Technical challenges and gains. Internet Things 28, 101364. http://dx.doi.org/10.1016/j.iot.2024.101364, URL https://www.sciencedirect.com/science/article/pii/S2542660524003056.

Lytos, A., Lagkas, T., Sarigiannidis, P., Zervakis, M., Livanos, G., 2020. Towards smart farming: Systems, frameworks and exploitation of multiple sources. Comput. Netw. 172, 107147. http://dx.doi.org/10.1016/j.comnet.2020.107147, URL https://www.sciencedirect.com/science/article/pii/S1389128620301201.

Mao, H., Zhou, L., Ying, R., Pan, D., 2021. Time preferences and green agricultural technology adoption: Field evidence from rice farmers in China. Land Use Policy 109, 105627. http://dx.doi.org/10.1016/j.landusepol.2021.105627, URL https://www.sciencedirect.com/science/article/pii/S0264837721003501.

Masison, J., Beezley, J., Mei, Y., Ribeiro, H., Knapp, A.C., Sordo Vieira, L., Adhikari, B., Scindia, Y., Grauer, M., Helba, B., Schroeder, W., Mehrad, B., Laubenbacher, R., 2021. A modular computational framework for medical digital twins. Proc. Natl. Acad. Sci. 118 (20), e2024287118. http://dx.doi.org/10.1073/pnas.2024287118, URL https://www.pnas.org/doi/10.1073/pnas.2024287118.

Matta, A., Lugaresi, G., 2024. An introduction to digital twins. In: 2024 Winter Simulation Conference. WSC, pp. 1281–1295. http://dx.doi.org/10.1109/WSC63780.2024.10838793.

Mirbod, O., Choi, D., Schueller, J.K., 2025. From simulation to field validation: A digital twin-driven sim2real transfer approach for strawberry fruit detection and sizing. AgriEngineering 7 (3), http://dx.doi.org/10.3390/agriengineering7030081, URL https://www.mdpi.com/2624-7402/7/3/81.

Mitsanis, C., Hurst, W., Tekinerdogan, B., 2024. A 3D functional plant modelling framework for agricultural digital twins. Comput. Electron. Agric. 218, 108733. http://dx.doi.org/10.1016/j.compag.2024.108733, URL https://www.sciencedirect.com/science/article/pii/S0168169924001248.

Ogunsoto, O.V., Olivares-Aguila, J., ElMaraghy, W., 2025. A conceptual digital twin framework for supply chain recovery and resilience. Supply Chain Anal. 9, 100091. http://dx.doi.org/10.1016/j.sca.2024.100091, URL https://www.sciencedirect.com/science/article/pii/S2949863524000347.

Purcell, W., Neubauer, T., 2023. Digital twins in agriculture: A state-of-the-art review. Smart Agric. Technol. 3, 100094. http://dx.doi.org/10.1016/j.atech.2022.100094, URL https://www.sciencedirect.com/science/article/pii/S2772375522000594.

Pylianidis, C., Osinga, S., Athanasiadis, I.N., 2021. Introducing digital twins to agriculture. Comput. Electron. Agric. 184, 105942. http://dx.doi.org/10.1016/j.compag.2020.105942, URL https://www.sciencedirect.com/science/article/pii/S0168169920331471.

Redeker, M., Weskamp, J.N., Rössl, B., Pethig, F., 2021. Towards a digital twin platform for industrie 4.0. In: 2021 4th IEEE International Conference on Industrial Cyber-Physical Systems. ICPS, pp. 39–46. http://dx.doi.org/10.1109/ICPS49255.2021.9468204.

Robles, J., Martín, C., Díaz, M., 2023. OpenTwins: An open-source framework for the development of next-gen compositional digital twins. Comput. Ind. 152, 104007. http://dx.doi.org/10.1016/j.compind.2023.104007, URL https://www.sciencedirect.com/science/article/pii/S0166361523001574.

Ruzzante, S., Labarta, R., Bilton, A., 2021. Adoption of agricultural technology in the developing world: A meta-analysis of the empirical literature. World Dev. 146, 105599. http://dx.doi.org/10.1016/j.worlddev.2021.105599, URL https://www.sciencedirect.com/science/article/pii/S0305750X2100214X.

Sharma, V., Tripathi, A.K., Mittal, H., 2022. Technological revolutions in smart farming: Current trends, challenges & future directions. Comput. Electron. Agric. 201, 107217. http://dx.doi.org/10.1016/j.compag.2022.107217, URL https://www.sciencedirect.com/science/article/pii/S0168169922005324.

Sivamani, S., Kwak, K., Cho, Y., 2014. A rule based event-driven control service for vertical farm system. Lect. Notes Electr. Eng. 309, 915–920. http://dx.doi.org/10.1007/978-3-642-55038-6_138.

Skobelev, P., Budaev, D., Gusev, N., Voschuk, G., 2018. Designing multi-agent swarm of UAV for precise agriculture. In: Bajo, J., Corchado, J.M., Navarro Martinez, E.M., Osaba Icedo, E., Mathieu, P., Hoffa-Dbrowska, P., del Val, E., Giroux, S., Castro, A.J., Sanchez-Pi, N., Julian, V., Silveira, R.A., Fernandez, A., Unland, R., Fuentes-Fernandez, R. (Eds.), Highlights of Practical Applications of Agents, Multi-Agent Systems, and Complexity: The PAAMS Collection. Springer International Publishing, Cham, pp. 47–59.

Sung, Y.-M., Kim, T., 2022. Smart farm realization based on digital twin. ICIC Express Lett. 13, 421–427, URL http://www.icicelb.org/ellb/contents/2022/4/elb-13-04-11.pdf.

Tagarakis, A.C., Benos, L., Kyriakarakos, G., Pearson, S., Sørensen, C.G., Bochtis, D., 2024. Digital twins in agriculture and forestry: A review. Sensors 24 (10), http://dx.doi.org/10.3390/s24103117, URL https://www.mdpi.com/1424-8220/24/10/3117.

Tao, F., Sun, X., Cheng, J., Zhu, Y., Liu, W., Wang, Y., Xu, H., Hu, T., Liu, X., Liu, T., Sun, Z., Xu, J., Bao, J., Xiang, F., Jin, X., 2024. makeTwin: A reference architecture for digital twin software platform. Chin. J. Aeronaut. 37 (1), 1–18. http://dx.doi.org/10.1016/j.cja.2023.05.002, URL https://www.sciencedirect.com/science/article/pii/S1000936123001541.

Tao, F., Xiao, B., Qi, Q., Cheng, J., Ji, P., 2022. Digital twin modeling. J. Manuf. Syst. 64, 372–389. http://dx.doi.org/10.1016/j.jmsy.2022.06.015.

Tao, F., Zhang, M., 2017. Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing. IEEE Access 5, 20418–20427. http://dx.doi.org/10.1109/ACCESS.2017.2756069.

Thomopoulos, R., Destercke, S., Charnomordic, B., Johnson, I., Abécassis, J., 2013. An iterative approach to build relevant ontology-aware data-driven models. Inform. Sci. 221, 452–472. http://dx.doi.org/10.1016/j.ins.2012.09.015, URL https://www.sciencedirect.com/science/article/pii/S0020025512006081.

Tržec, K., Kušek, M., Žarko, I.P., 2022. Building an interoperable IoT ecosystem for data-driven agriculture. In: 2022 International Conference on Smart Systems and Technologies. SST, pp. 341–347. http://dx.doi.org/10.1109/SST55530.2022.9954641.

Verdouw, C., Tekinerdogan, B., Beulens, A., Wolfert, S., 2021. Digital twins in smart farming. Agricult. Sys. 189, 103046. http://dx.doi.org/10.1016/j.agsy.2020.103046, URL https://www.sciencedirect.com/science/article/pii/S0308521X20309070.

Wang, L., 2024. Digital twins in agriculture: A review of recent progress and open issues. Electronics 13 (11), http://dx.doi.org/10.3390/electronics13112209, URL https://www.mdpi.com/2079-9292/13/11/2209.

Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.-J., 2017. Big data in smart farming – a review. Agricult. Sys. 153, 69–80. http://dx.doi.org/10.1016/j.agsy.2017.01.023, URL https://www.sciencedirect.com/science/article/pii/S0308521X16303754.

Wulf, C., Wiechmann, C.C., Hasselbring, W., 2016. Increasing the throughput of pipe-and-filter architectures by integrating the task farm parallelization pattern. In: 2016 19th International ACM SIGSOFT Symposium on Component-Based Software Engineering. CBSE, pp. 13–22. http://dx.doi.org/10.1109/CBSE.2016.21.

Xu, Y., Qamsane, Y., Puchala, S., Januszczak, A., Tilbury, D.M., Barton, K., 2024. A data-driven approach toward a machine- and system-level performance monitoring digital twin for production lines. Comput. Ind. 157–158, 104086. http://dx.doi.org/10.1016/j.compind.2024.104086, URL https://www.sciencedirect.com/science/article/pii/S0166361524000149.

Zheng, P., Sivabalan, A.S., 2020. A generic tri-model-based approach for product-level digital twin development in a smart manufacturing environment. Robot. Comput.-Integr. Manuf. 64, 101958. http://dx.doi.org/10.1016/j.rcim.2020.101958, URL https://www.sciencedirect.com/science/article/pii/S0736584519305289.