Energy-efficient start-up optimization via digital twin for a vegetable broth sterilization process

Marcello Maria Bozzini[a], Marco Menegon[b], Alberto di Loreto[b], Gaia Lunari[a], Silvia Serena Mariani[a], Mattia Vallerio[a], Laura Piazza[b], Flavio Manenti[a,*]

[a] Dipartimento di Chimica, Materiali e Ingegneria Chimica ''Giulio Natta'', Politecnico di Milano, Piazza Leonardo da Vinci, 32, Milan, 20133, Italy
[b] Department of Environmental Science and Policy, Università degli Studi di Milano, Via Mangiagalli, 25, Milan, 20133, Italy

A R T I C L E I N F O

Keywords:
Food engineering
Digital twin
Data reconciliation
Process simulation
Optimization

A B S T R A C T

The food industry is undergoing a digital transformation driven by the need for greater sustainability, efficiency, and data integration. This study presents a methodology for implementing a digital twin in an industrial food manufacturing process, using the vegetable broth production line as a case study. The workflow integrates process analysis, sensor data collection, and data reconciliation to improve the reliability of process variables and enable accurate simulation. The reconciled data were used to develop a dynamic model in commercial software, capable of simulating different operating conditions. Two start-up strategies, cold start-up and pre-heating, were compared, revealing that pre-heating reduces steam consumption by 62% and start-up time by 63%. These results demonstrate the potential of digital twins in optimizing operational efficiency and energy use in the food industry. Future developments may include real-time data acquisition, integration with control systems, and the use of AI for predictive maintenance and process optimization.

1. Introduction

Nowadays, the food industry faces the challenge of meeting the ever-growing demand for food, driven by global population growth (Gilland, 2002). Additionally, as highlighted in recent UN reports, the agri-food sector makes an important contribution to climate change, primarily due to food waste and excessive water use (United Nations, 2025). In this context, food industries must rethink their production processes to enhance efficiency and align with the Sustainable Development Goals (SDGs) (Premalatha et al., 2011). Industry 4.0 marks a new era of digitalization, introducing innovative tools that enhance the management and efficiency of processes (Yadav et al., 2022). Various studies suggest that the adoption of these technologies will enable the transition to smart manufacturing, characterized by a high degree of automation through the widespread use of remote sensing, real-time data acquisition, monitoring, and advanced visualization tools (Chen et al., 2020). As we enter the new era of Agrifood 4.0, there is a growing need to collect and analyze vast amounts of data to generate deeper insights (van den Berg et al., 2013). This will enhance various aspects of the food supply chain, including process control, supply chain management, safety, traceability, quality, and process yields (Hill and Scudder, 2002). At the same time, these advancements promote sustainability and contribute to the reduction of food waste (Hassoun et al., 2023). One effective way to leverage data for process optimization is by integrating it into a ''digital twin'', a virtual replica of a real-world physical process (Maheshwari et al., 2023). Digital data truly unlocks its value when it is used to understand and predict process behavior, enabling smarter and more efficient operations (Zhang et al., 2023). To this purpose, at the core of any digital twin implementation in manufacturing, there are mathematical models representing a process or product (Koulouris et al., 2021). These models can take the form of a complex multi-physics simulation, capturing the detailed behavior of a product or individual process (Verboven et al., 2020). Moreover, the Digital Twin can leverage both online and offline data gathered from the production environment through Internet of Things (IoT) platforms that enable the connection between sensors in the process plant and the Internet. IoT systems inherently gather data from heterogeneous sources, with several types of sensors offering different levels of precision, data ranges, units, and device specifications. These variations, coupled with inherent uncertainties, can result in poor data quality, potentially leading to incorrect outcomes from the Digital Twin (Choi et al., 2022). As suggested by Galeazzi et al. (2023), accurate data collection and precise quantification of material and energy consumption require the use of data reconciliation, a model-based technique

[l] This article is part of a Special issue entitled: 'Digital Twins' published in Journal of Food Engineering.

* Corresponding author.

E-mail address: flavio.manenti@polimi.it (F. Manenti).

https://doi.org/10.1016/j.jfoodeng.2025.112822

Received 29 June 2025; Received in revised form 19 September 2025; Accepted 22 September 2025
Available online 30 September 2025

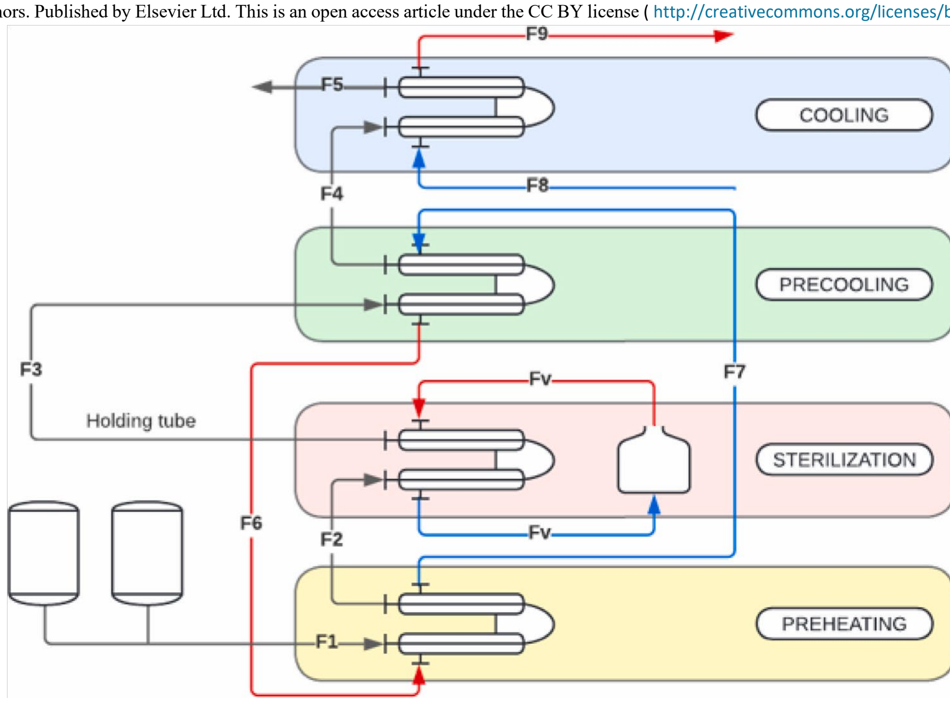


Fig. 1. Process flow diagram of the process showing the four main processing steps: Preheating, Sterilization, Precooling, and Cooling. The gray arrows F1, F2, F3, F4, and F5 represent the broth flow through the system. Red arrows represent heating service fluids, while blue arrows represent cooling service fluids.

that minimizes measurement errors by leveraging redundancies (Zhang et al., 2001). Compared to relying on raw data, this approach enables a more precise evaluation of process variables, thus enhancing the efficiency and reliability of the digital twin (Rodríguez et al., 2023).

1.1. Objective of the work

This research aims to demonstrate the efficacy of commercially available process simulators for modeling and optimizing processes in the food industry. This paper analyzes an industrially relevant food process and develops an effective and reliable methodology for the implementation of digital twins in food manufacturing. The case study highlights the challenges the food industry faces in collecting reliable data and underscores the importance of data reconciliation in improving data quality, as required by ISO 8000 and ISO 23247 standards (Rivas et al., 2017; Caiza and Sanz, 2024), ultimately leading to the successful implementation of the digital twin.

The remainder of the paper is organized as follows: Section 2 presents the methodology, including process analysis, data collection, data cleaning, and data reconciliation. Section 3 reports the results of the digital twin implementation and the comparison of different operating strategies. Section 4 discusses the main findings, outlines possible applications, and highlights future research directions.

2. Material and methods

2.1. Case study This study focuses on the production line of vegetable broth, where the operational phases and process conditions were examined to identify stages that could be optimized from an energy perspective.

2.2. Process and sterilization phase

Fig. 1 shows the process flow diagram of the analyzed process. The initial phase of vegetable broth preparation involves mixing ingredients within insulated tanks (F1 within Fig. 1) at a temperature of 323 K.

Sterilization represents the core operation of the process, ensuring its overall suitability for human consumption. This is performed using a shell-and-tube heat exchanger. The sterilization system features a closed-loop preheating and precooling cycle, utilizing water as service fluid. This cycle preheats and precools the broth, respectively, before and after the sterilization phase, recovering heat from the sterilization phase. Sterilization is carried out using a countercurrent ultra-high temperature (UHT) thermal treatment with saturated steam at 431 K and 6 bar, supplied by an external boiler. The broth is heated to 406 ± 3 K and then maintained at this temperature for 36 s, a duration ensured by the flow rate within a 22-m-long holding tube. After the holding phase, the broth undergoes a cooling phase, achieved countercurrently with cold water at 291 K and a flow rate of around 16.000 l/h, which lowers the broth temperature to 308 K. The steam flow to the sterilizer is variable and controlled by a modulating valve. This adjustment is essential to compensate for fouling, caused by calcium deposits on the inner walls of the tubes, which gradually reduce heat exchange efficiency. A temperature sensor, positioned at the end of the holding phase, detects any drop in broth temperature and adjusts the valve opening to maintain the broth temperature consistent at 406 ± 3 K. During a 6-h work shift, fouling causes a temperature drop that requires progressively increasing the valve opening from 20% at the beginning to 80%–90% by the end of the shift. A Cleaning in Place (CIP) system restores the initial operating conditions.

2.3. Digital twin development

Fig. 2 shows the digital twin development methodology, which was applied to this process. This methodology involves analyzing the production process, collecting process variable data, cleaning and organizing the information into a dataset, applying the data reconciliation method, and ultimately developing the digital twin using specific software.

2.4. Process analysis

In this phase, all technical documentation and the existing process layout are analyzed in detail to collect all relevant information: (i) unit dimensions, (ii) online and offline sensors location, (iii) standard operating conditions, (iv) production recipes. This information is then used to define a Data Collection strategy.

Table 1
Raw dataset collected by merging technical plant design data with inline and offline measurements.

Variable	Value	Measuring device
Broth flow rate	4200 Lh^{-1}	Inline flow meter
Pre-cooling water flow rate	14000 Lh^{-1}	Plant design
Pre-heating water flow rate	14000 Lh^{-1}	Plant design
Cooling water flow rate	16000 Lh^{-1}	Plant design
Steam temperature	431 K	Plant design
Steam flow rate	1000 kg h^{-1}	Plant design
Sterilization temperature	406 K	Inline temperature sensor
Cooling water temperature	291 K	Offline thermometer
Pre-heating water temperature	365 K	Offline thermometer
Pre-cooling water temperature	323 K	Offline thermometer
Packaging broth temperature	308 K	Inline temperature sensor
Steam pressure	6bar	Plant design

2.5. Sensors and data collection

The digital twin development methodology includes collecting all variables necessary to describe the production process. In this case study, this information was obtained using a combination of inline sensors, offline sensors, and technical plant data. Table 1 illustrates the dataset collected. Inline sensors were used to monitor temperature and flow rate during the sterilization and cooling phases of the broth. Offline sensors were employed to measure the temperatures of the closed-loop heat recovery circuit and the service water used for cooling. For service fluid flow rates, technical plant design values were used due to the absence of sensors for direct measurement. The vapor flow rate measure comes from the technical datasheet of the boiler. However, it must be considered that the boiler supplies other minor sections of the plant.

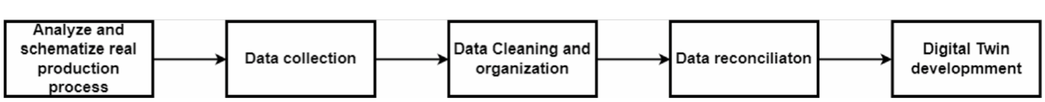


Fig. 2. Digital twin development methodology. The Proposed methodology comprises five steps: Process analysis, Data collection, Data cleaning and organization, Data reconciliation and digital twin development.

2.6. Data cleaning and organization

Industrial data is far from ready to use. Once the data is collected in raw format, i.e., from the plant control device (e.g., Programmable Logic Controller (PLC), Distributed Control System (DCS), Supervisory Control and Data Acquisition (SCADA)), it first needs to be organized in standard table or datasheet format to be further visualized and analyzed correctly. The most common table format consists of organizing variables per column and each row is then an independent time sample of each variable. The following step is the removal of irrelevant data. The most obvious one is the removal of all the data when the plant is not in operation or it is undergoing maintenance or cleaning procedure (Salano et al., 2025). Another common example is the segregation of data related to different products that are manufactured in the same plant. This step is necessary to model each production process independently. Finally, the last step is data cleaning. This step is necessary to eliminate outliers, gross errors, and irrelevant data for the defined scope. This step can be performed with several different algorithms; the most efficient one for industrial data is the Isolation Forest (IF) algorithm (as shown by Salano et al. in their application of this methodology to an industrial dataset (Salano et al., 2025)).

2.7. Data reconciliation

After the removal of irrelevant data and outliers, the data collected through the sensors underwent a data reconciliation procedure to ensure that the input data for the digital twin development accurately reflected the plant's operative conditions. The data reconciliation allows for refining the measurements collected by on-field sensors to satisfy the mass and energy balances characterizing the analyzed process (Özyurt and Pike, 2004). It consists of an error minimization problem, which ensures that reconciled data are both accurate and consistent with the underlying physical laws governing the process. Data reconciliation was performed using an iterative approach to minimize discrepancies between the measurements and the balances governing the studied process. As the process is linear and stationary, the mass balances were not used, as not necessary (see Fig. 1). One energy balance was defined for each heat exchanger, namely:

1. **Energy balance on the pre-heating**

$$\dot{F}_1 c_{p,B}(T_2 - T_1) = \dot{F}_6 c_{p,W}(T_6 - T_7) \tag{1}$$

2. **Energy balance on the sterilizer**

$$\dot{F}_2 c_{p,B}(T_3 - T_2) = \dot{F}_v \Delta H_s \tag{2}$$

3. **Energy balance on the pre-cooling**

$$\dot{F}_3 c_{p,B}(T_3 - T_4) = \dot{F}_7 c_{p,W}(T_6 - T_7) \tag{3}$$

4. **Energy balance on the cooling**

$$\dot{F}_4 c_{p,B}(T_4 - T_5) = \dot{F}_8 c_{p,W}(T_9 - T_8) \tag{4}$$

where $\dot{F}_i$ are the material streams depicted in Fig. 1, T_i are the temperatures associated with each material stream, $c_{p,B}$ is the broth specific heat at constant pressure, $c_{p,W}$ is the water specific heat at constant pressure, $\dot{F}_v$ is the steam flow circulating in the sterilization section and ΔH_s is the steam heat of vaporization. A constraint on the temperature of the broth flow exiting the sterilization phase was posed to ensure that it stayed within the defined range (134 ± 3 °C). Python was used within the Visual Studio Code editor to write two optimization scripts aimed at minimizing two error functions, defined as follows:

1. **Minimization of the Relative Square Error (RSE)**

$$\min \phi = \sum_{i=1} \frac{\left(x_i^r - x_i\right)^2}{x_i^2} \tag{5}$$

2. **Minimization of the Absolute Error (RSE)**

$$\min \phi = \sum_{i=1} \left|x_i^r - x_i\right| \tag{6}$$

The RSE was used to normalize errors of different orders of magnitude. Each minimization function was applied to two versions of the dataset, differentiated by the unit of measure of the stream flows, defined first in kilograms per hour (kg/h) and then in quintals per hour (q/h). This approach brings all the variables to the same order of magnitude, improving the reliability of the data reconciliation results.

2.8. Digital twin

The physical process was digitally reconstructed using AVEVA Dynamic Simulation. The equipment was sized according to the physical

Table 2 Data reconciliation results.

Reconciled value [RSE][a]	Reconciled value [AE][a]	Reconciled value [RSE][b]	Reconciled value [AE][b]
Flow rates			
F_{br} = 4237.1 [0.049%]	F_{br} = 4234.8 [−0.004%]	F_{br} = 42.4 [0.145%]	F_{br} = 42.4 [0.000%]
F_{pr} = 13999.8 [−0.001%]	F_{pr} = 14000.1 [0.001%]	F_{pr} = 141.2 [0.892%]	F_{pr} = 140.0 [0.000%]
F_{cr} = 16000.0 [−0.000%]	F_{cr} = 16000.1 [0.000%]	F_{cr} = 159.7 [−0.184%]	F_{cr} = 160.0 [−0.000%]
F_{vr} = 993.7 [−0.631%]	F_{vr} = 664.6 [−33.536%]	F_{vr} = 5.8 [−42.212%]	F_{vr} = 3.8 [−61.540%]
Temperatures			
T_{1r} = 305.9 [−2.306%]	T_{1r} = 313.3 [−0.032%]	T_{1r} = 309.6 [−1.121%]	T_{1r} = 313.2 [0.000%]
T_{2r} = 285.7	T_{2r} = 325.7	T_{2r} = 336.7	T_{2r} = 359.4
T_{3r} = 405.0 [−0.136%]	T_{3r} = 405.5 [−0.004%]	T_{3r} = 406.0 [0.111%]	T_{3r} = 405.6 [0.000%]
T_{4r} = 425.2	T_{4r} = 393.1	T_{4r} = 379.0	T_{4r} = 359.4
T_{5r} = 309.1 [0.313%]	T_{5r} = 308.3 [0.054%]	T_{5r} = 307.7 [−0.131%]	T_{5r} = 308.1 [−0.000%]
T_{6r} = 353.8 [−3.122%]	T_{6r} = 365.1 [−0.001%]	T_{6r} = 357.6 [−2.056%]	T_{6r} = 359.4 [−1.588%]
T_{7r} = 359.7 [11.326%]	T_{7r} = 361.5 [11.856%]	T_{7r} = 349.7 [8.213%]	T_{7r} = 345.7 [6.966%]
T_{8r} = 286.9 [−1.471%]	T_{8r} = 291.2 [0.012%]	T_{8r} = 292.7 [0.531%]	T_{8r} = 293.5 [0.813%]
T_{9r} = 317.0 [1.225%]	T_{9r} = 313.2 [0.004%]	T_{9r} = 311.2 [−0.617%]	T_{9r} = 306.8 [−2.031%]

[a] Values in kg/h. [b] Values in q/h.

plant specifications when available or derived from the available information and reasonable assumptions when they were not provided. The specific heat of broth was computed using Choi and Okos' correlations, which are widely used for aqueous food mixtures with a comparable solids content ~5% (Choi, 1986). Similarly, the heat transfer coefficients for the heat exchangers were assumed within a range of 800–1500 [W/m^2 K] (Perry, 1999). These assumptions are consistent with the physical and chemical properties of the broth.

The model inputs consist of the reconciled plant parameters (flow rates, temperatures and setpoints reported in Table 2) and the overall heat-transfer coefficient adopted for each stage. The system is formulated via first-principles dynamic energy balances for the equipment items involved in the process, with boundary conditions given by the reconciled variables above. The system equations and discretization options follow the AVEVA DynSim User Guide, to which we refer for full details and notations. The simulations were executed using the default variable-step implicit integrator with tight tolerances.

3. Results and analysis

In this section, the results obtained from the various steps of the digital twin development and simulation are reported and discussed.

3.1. Data reconciliation

The algorithm was adapted to minimize the Relative Square Error (RSE) and the Absolute Error (AE) functions, as defined in the Materials and Methods section. The reconciliation process was performed on raw stream and temperature data. The reconciliation step was applied twice to two versions of the dataset. The two versions are distinguished by the different units of measure of the raw stream data: kilograms per hour [kg/h] and quintals per hour [q/h]. This was done to test the reconciliation algorithm with different orders of magnitude. In particular, the modified dataset with the flow measures reported in quintals per hour shows all variables with the same order of magnitude. This should result in a more even distribution of the final error among all the variables. The dataset with this modification will be referred to as the normalized dataset, the other as the original dataset. Hence, the data reconciliation resulted in four datasets, emerging from the application of the two versions of the algorithm to the two datasets (see Table 2). Therefore, we have:

1. Original dataset with MSE minimization
2. Original dataset with AE minimization
3. Normalized dataset with MSE minimization 4. Normalized dataset with AE minimization The reconciliation carried out on the normalized dataset by minimizing the MSE (Dataset N°3) is the most consistent, as it respects the physical constraints and correctly reflects the operative conditions. The relative error with respect to the measured variables is minimized, while the reconciled vapor flow rate assumes a realistic value. It is important to underline that the measured value of the vapor flow reflects the overall capacity of the boiler unit, which supplies several sections of the site. This dataset was used as input for the construction of the digital model.

3.2. Start-up scenario

This study compares two start-up strategies reported in Fig. 3 for a broth sterilization process using a digital model developed in AVEVA Dynamic Simulation (see Fig. 4).

The cold start-up approach involves the simultaneous activation of all controllers, while the plant pre-heating approach preconditions service fluids and equipment before introducing the broth. Simulations show that the plant pre-heating strategy significantly reduces the time required to reach steady-state sterilization conditions, highlighting its potential to improve operational efficiency in food manufacturing processes. The process simulations were carried out following operational setpoints, and a controller on the steam valve regulated the input to achieve the target sterilization temperature. Output variables were benchmarked against reconciled plant data to validate consistency between predicted and expected behavior.

3.2.1. Cold start-up

In the cold start-up scenario, all controllers were activated simultaneously, operating in automatic mode. The system required approximately 32 min to reach the steady sterilization temperature of 406 K (133 °C), due to the slow approach to operating conditions as the controller progressively regulated the vapor flow to reach the target temperature. During this phase, the steam flow rate increased progressively until stabilizing around 610 kg/h. As shown in Fig. 5, the sterilization temperature (red line) and steam flowrate (blue line in Fig. 6) rise synchronously, confirming thermal coupling through the heat exchanger. The broth flow rate and the broth vapor fraction in the process tube (respectively, blue and gray lines in Fig. 5) are also plotted; the inclusion of the latter helps explain the initial temperature oscillations. During the first 3 min of start-up, broth vaporization is observed, primarily due to insufficient flow rate. The smooth temperature ramp reflects a controlled increase in thermal input by the controller. Fig. 6 illustrates the convergence between steam and condensate temperatures, indicating effective heat transfer and condensation. This is evidenced by the narrowing gap between the vapor stem temperature (red line) and the condensation temperature (red dotted line), while by 12 min (peak steam flow 740kgh⁻¹). the steam flow rate (blue line) steadily increases up to 610 kg/h. The green vertical dotted line at 32 min indicates the end of the start-up phase of the process.

3.2.2. Plant pre-heating start-up

In the plant pre-heating start-up scenario, both the service fluids and the sterilizer were pre-heated and stabilized for approximately 6 min before introducing the broth into the system. Subsequently, the steam valve was manually fully opened to accelerate the heating phase, resulting in a peak steam flow rate of about 740 kg/h. This strategy enabled the system to reach the steady sterilization temperature of 405 K in approximately 12 min (green vertical line, end of start-up phase), significantly reducing the start-up time compared to the cold startup condition. The temperature profile in Fig. 7 illustrates the initial stabilization phase, during which steam heats the sterilizer until it reaches 405 K at minute 6. This is followed by a sharp temperature drop to 340 K upon the introduction of the broth, which is promptly corrected through manual adjustment of the steam valve. The broth then gradually heats up, reaching the target operating temperature of 405 K by minute 12. The solid red line represents the temperature immediately downstream of the sterilizer. The vapor fraction of the broth in the process tube (gray line) is also shown to help interpret the flow dynamics (blue line) and the temperature variations observed in the discontinuity region around minute 6. The yellow dotted line marks the opening of the broth line valve. As in previous cases, during the first 2 min, the broth vaporizes due to the insufficient flow.

Fig. 8 shows the steam flow rate (blue line), along with the vapor temperature (red line) and the condensed steam temperature (dotted red line). Following the manual activation of the steam valve, the flow rate increases sharply. Once the broth begins to circulate, the system stabilizes at approximately 740 kg/h, as the controller is switched to automatic mode and the valve remains fully open.

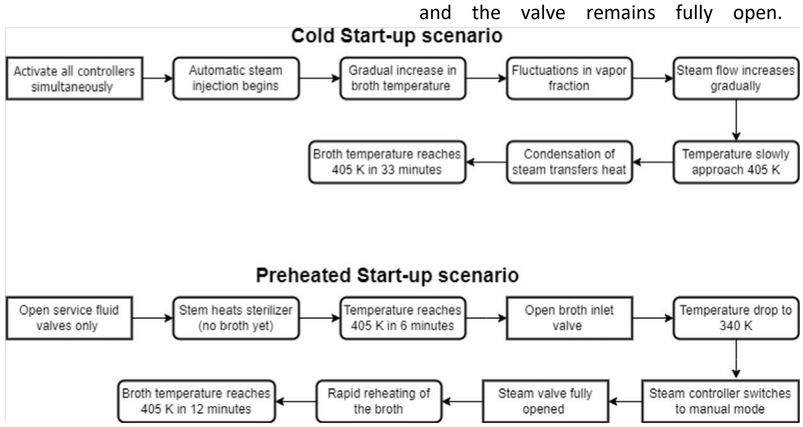


Fig. 3. Sequence of operations for the two start-up strategies. **Cold start-up:** all control loops are enabled in automatic; steam injection ramps up and condenses in the exchanger, producing a gradual rise in broth temperature with small vapor-fraction oscillations. The sterilization set-point (405K) is reached after 33 min as the steam flow stabilizes around 610kgh⁻¹. **Plant pre-heating start-up:** service fluids pre-heat the sterilizer to 405K in 6 min (no broth). When the broth valve opens, the temperature briefly drops to 340K; the steam controller is switched to manual and the valve fully opened, yielding rapid reheating to the set-point

3.2.3. Energy consumption comparison

Fig. 9 illustrates the vapor flow rate in both scenarios. The faster start-up results in a lower consumption of steam for the start-up period, as reported in Table 3. As can be seen, the proposed pre-heating strategy not only reduces the start-up time by 20 min, but also results in a reduction of 167 kg of steam. Taking into account that this plant is started up every day for 200 days a year, the proposed strategy could lead to an annual saving of 33.4 tons of steam and fresh water for this production line alone.

It should be noted again that the steam flow data were retrieved from a shared boiler unit, representing an upper-bound limit of that variable. This could introduce uncertainty in the steam consumption quantification. The data reconciliation procedure was applied to minimize this bias. Moreover, the relative comparison of the two start-up procedures remains unaffected, as both scenarios rely on the same baseline. In addition to the uncertainty related to the boiler, some other variables were not available online and were therefore estimated through data reconciliation. To improve the reliability of the model, it is recommended to install additional measurement sensors (e.g., temperature upstream of the boiler). A more accurate model would lead to more precise results. However, the current model is still suitable for offline optimization purposes.

Table 3

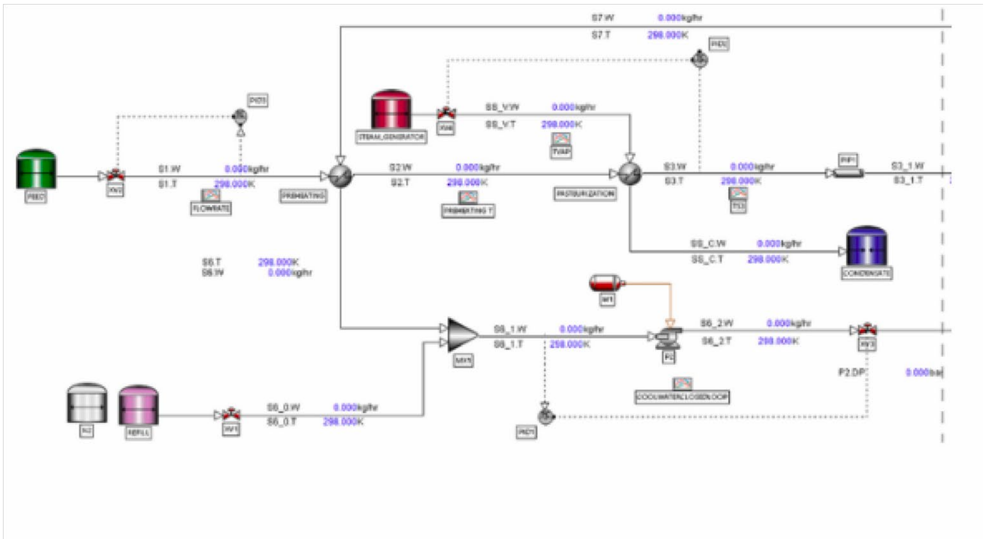
Steam consumption for start-up.

	Steam [kg]	Time [min]
Cold Start-up	269.54	32
Preheating	102.17	12
Difference	−167.37	20
Difference %	−62%	−63%

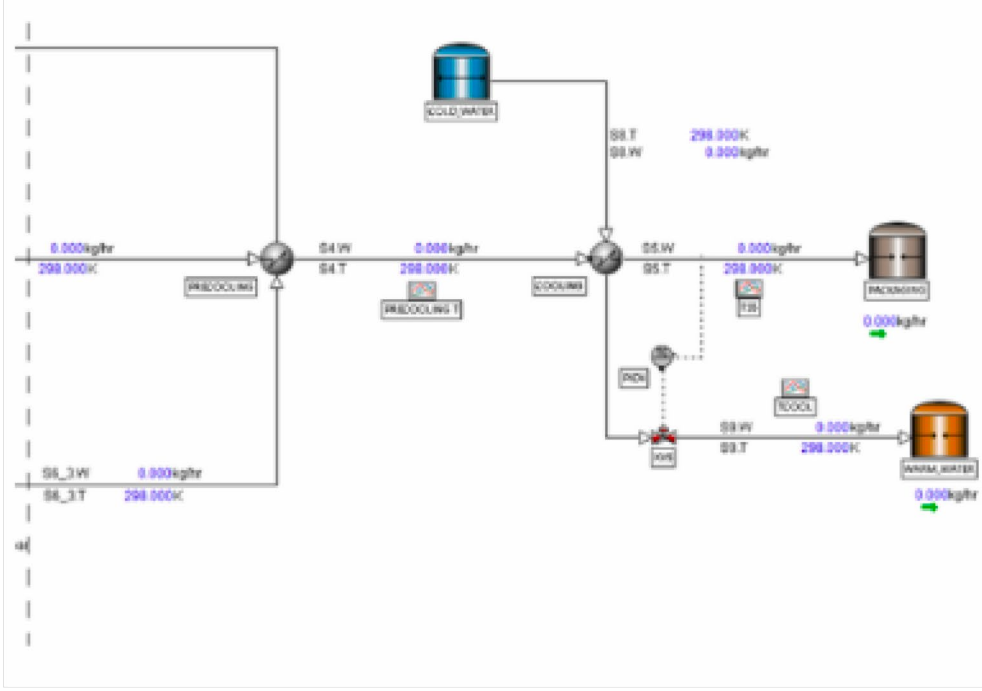
4. Discussion

4.1. Possible applications and future development

The model developed in this study represents a valuable simulation tool for evaluating different operating conditions, such as variations in broth temperature, flow rate, and start-up strategies. It also provides a foundation for predictive maintenance, as the integration of fouling parameters enables the anticipation of cleaning or replacement needs, minimizing downtime and preventing failures. To further enhance the model, more accurate and continuous data acquisition is required, which includes the deployment of smart sensors and a real-time data communication infrastructure. Additionally, integration with the plant’s control system and operator training are essential for effective implementation in an industrial context. Another promising line of development concerns energy efficiency. By simulating different process configurations, it would be possible to optimize the use of utilities such as steam or cooling water, helping reduce costs and environmental impact. For future developments, the integration of artificial intelligence (AI) and machine learning (ML) technologies could allow the system to learn from historical data, recognize patterns, and further improve its predictive and adaptive capabilities. A possible AI/ML implementation roadmap will be explained in the next subsection.



(a) First plant section includes the Preheating and the Sterilization units



(b) Second plant section includes the Precooling and Cooling units

Fig. 4. Broth sterilization process scheme in AVEVA dynamic simulation.

4.2. Conceptual roadmap for AI/ML integration Building on the reconciled data and validated model, a pragmatic AI/ML layer can be added in four steps: (i) *physics-informed soft sensors* that estimate unmeasured states (e.g., effective heat-transfer coefficients, fouling resistance, broth vapor fraction) by learning residuals around first-principles balances; (ii) *predictive maintenance and anomaly detection* using temporal models trained on reconciled *good* operation (e.g., time-to-Clean In Process forecasts and unsupervised detectors of sensor drift or process deviations); (iii) an *energy-aware supervisory optimizer* that embeds a learned surrogate of the sterilizer (e.g., Gaussian processes or neural ODEs constrained by mass/energy balances) within MPC or safe RL to compute start-up trajectories and steam-valve manipulations that minimize time and utilities, with policies validated in the twin before advisory deployment to the DCS; and (iv) *transfer learning across recipes/lines* to accelerate adaptation, with performance tracked on start-up time and steam use. This roadmap leverages assets produced in this work (reconciled historian, labeled operating modes, validated model) and requires only incremental instrumentation and logging to close the loop.

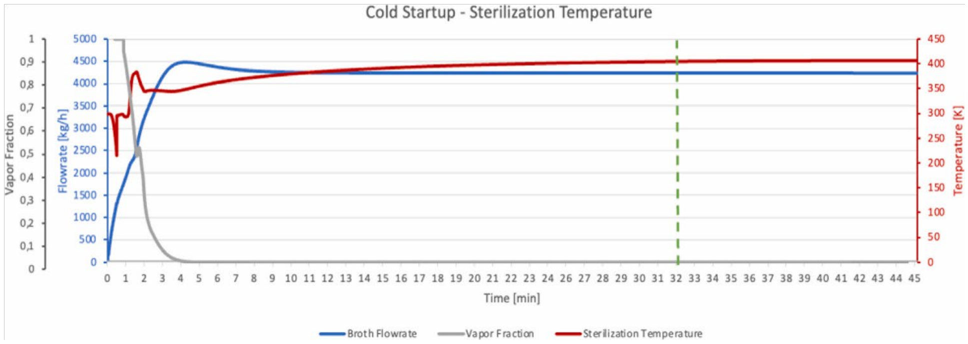


Fig. 5. Cold start-up scenario dynamics. Broth flow rate (blue, left axis), broth vapor fraction in the process tube (gray, left axis), and sterilization temperature (red, right axis) versus time. With all controllers in automatic, the temperature ramps smoothly toward the set-point (405K). During the first ∼3 min, insufficient broth flow causes transient vaporization and small temperature oscillations; the broth vapor fraction then decays to zero as flow stabilizes. The green vertical dashed line at 32 min marks the reaching of the temperature set-point and the end of the start-up phase.

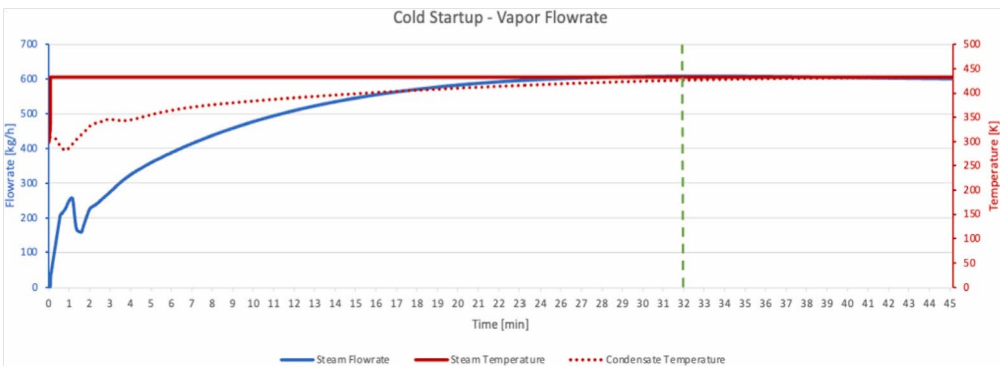


Fig. 6. Cold start-up steam-side behavior. Steam flow rate (blue, left axis) shows brief controller transients in the first ∼2 min, then ramps smoothly and stabilizes at 610kgh⁻¹. The steam temperature (solid red, right axis) remains nearly constant while the condensate temperature (red dotted) rises and progressively approaches the steam temperature, indicating effective condensation and heat-transfer stabilization. The green vertical dashed line at 32 min marks the reaching of the temperature set-point and the end of the start-up phase.

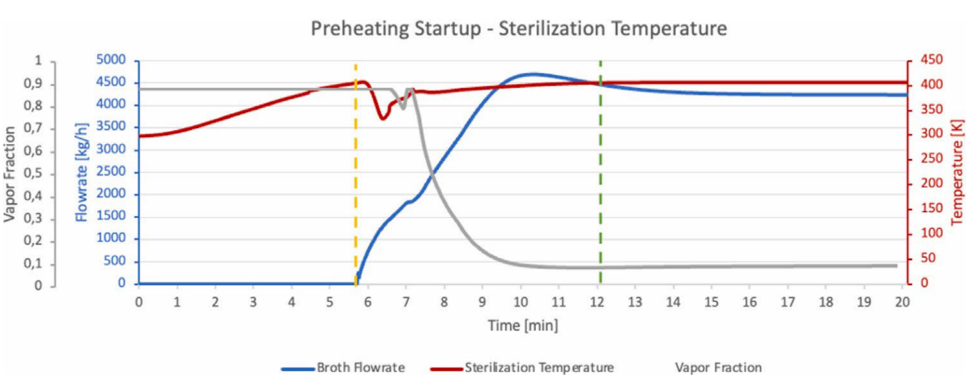


Fig. 7. Pre-heating start-up—sterilization temperature and broth dynamics downstream of the sterilizer. Plotted are broth flow rate (blue, left axis), broth vapor fraction in the process tube (gray, left axis), and sterilization temperature (red, right axis) versus time. The plant is pre-heated without broth until the sterilizer reaches 405K at ∼6 min; the yellow dashed line marks the opening of the broth valve. On admission of broth, the temperature briefly drops to ∼340K and then recovers rapidly (manual, fully opened steam valve), reaching the set-point by ∼12 min; the green dashed line indicates the end of the start-up phase. During the first ∼2 min after broth introduction, insufficient flow causes transient vaporization (gray), which decays to nearly zero as flow stabilizes.

5. Conclusions

Despite some limitations related to the availability and quality of the experimental data, which considerably slowed the model development, the data reconciliation process proved to be a valuable tool.

Specifically, AE minimization was more effective when dealing with data of different orders of magnitude, while RSE minimization did better on data of similar orders of magnitude.

The validated model was used to simulate two distinct start-up strategies. The first, referred to as the ‘‘cold start-up’’, involved the simultaneous activation of all system controls, allowing the plant to gradually reach operating conditions. The second strategy, defined as ‘‘plant preheating start-up’’, consists of a more controlled approach: service fluids were circulated, and the heat exchangers brought to temperature before product introduction. This latter scenario significantly

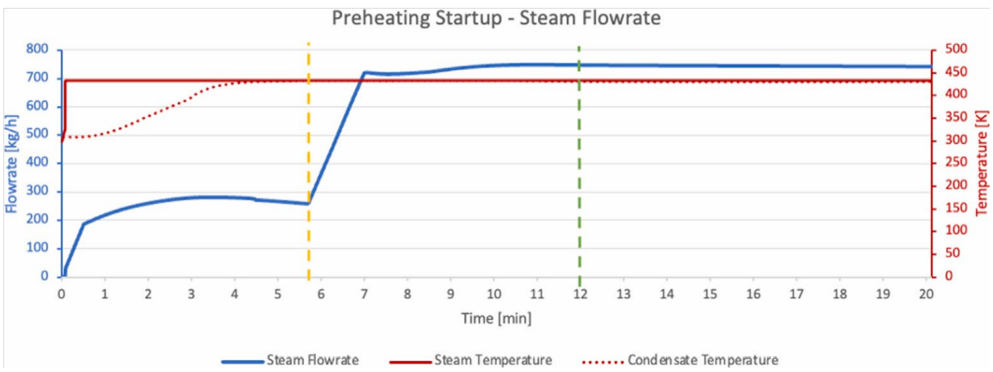


Fig. 8. Pre-heating start-up steam-side behavior. Steam flow rate (blue, left axis), steam temperature (solid red, right axis), and condensate temperature (red dotted, right axis) versus time. During pre-heating (0–6 min), the steam flow is moderate (~250–300kgh^{-1}) while the condensate temperature rises toward the steam temperature. At ~6 min (yellow dashed line), the broth line opens and the steam valve is placed in manual and fully opened, causing a step in steam flow to ~700–740kgh^{-1}. The condensate temperature then approaches the steam temperature as heat transfer stabilizes; by ~12 min (green dashed line), the start-up is complete and the steam flow has nearly stabilized.

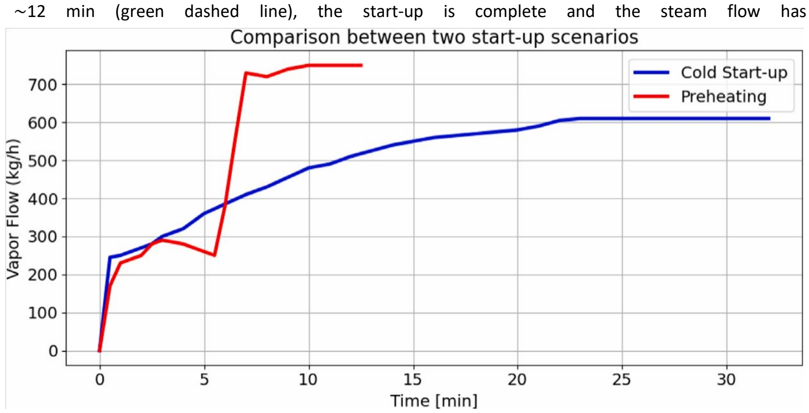


Fig. 9. Comparison of vapor (steam) flow rate during start-up for the two strategies. The cold start-up (blue) ramps gradually and reaches the operating condition in ~32 min with a stabilized flow of ~610kgh^{-1}, whereas the pre-heating strategy (red) attains the target in ~12 min with a brief peak near ~740kgh^{-1}. Despite the higher peak, the shorter start-up reduces the integral of the flow curve by 167kg of steam over the start-up period (see Table 3). Assuming one daily start-up over 200 production days, this corresponds to an annual saving of ~33.4t yr^{-1} of steam and fresh water.

reduced the time required to reach sterilization conditions, achieving the target temperature in 12 min, compared to approximately 32 min for the cold start-up. The optimization of the start-up procedure also leads to a decrease in the steam consumption of around 62% with respect to the base case.

The proposed DT workflow, data reconciliation, first-principles modeling, and historian-driven parameter identification, is scalable to *batch* operations, multi-product plants, and other thermal treatment units (e.g., pasteurization, HTST/UHT, blanching, tubular/plate exchangers). Extending it mainly requires (i) adapting the dynamic balances to batch/recipe timing, (ii) re-estimating heat-transfer and fouling parameters for each equipment/product, and (iii) encoding productspecific quality constraints (maximum temperature, time–temperature limits) to prevent organoleptic degradation such as loss of color or taste. The principal limitation is the availability of reliable product and equipment information. Thermophysical properties and quality limits may necessitate targeted tests before deployment. With these inputs, the same toolchain supports changeovers and multi-product operation with minimal modifications.

CRediT authorship contribution statement

Marcello Maria Bozzini: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Marco Menegon:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Alberto di Loreto:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Gaia Lunari:** Writing – original draft, Methodology, Formal analysis. **Silvia Serena Mariani:** Writing – original draft, Methodology, Formal analysis. **Mattia Vallerio:** Writing – original draft, Supervision, Methodology, Conceptualization. **Laura Piazza:** Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Flavio Manenti:** Writing – original draft, Supervision, Funding acquisition, Conceptualization.

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.

Acknowledgments

This work was funded by the European Union – Next Generation EU, Mission 4, Component 1, CUP D53D23004290006 – PRIN 2022 Project, code 2022XJ4Z42. The authors would like to acknowledge Cleca s.p.a., Italy for their collaboration.

Data availability

Data will be made available on request.

References

Caiza, G., Sanz, R., 2024. Immersive digital twin under ISO 23247 applied to flexible manufacturing processes. Appl. Sci. 14 (10), 4204. http://dx.doi.org/10.3390/ app14104204, URL https://www.mdpi.com/2076-3417/14/10/4204.

Chen, Y., Yang, O., Sampat, C., Bhalode, P., Ramachandran, R., Ierapetritou, M., 2020. Digital twins in pharmaceutical and biopharmaceutical manufacturing: A literature review. Processes 8 (9), 1088. http://dx.doi.org/10.3390/pr8091088, URL https://www.mdpi.com/2227-9717/8/9/1088.

Choi, Y., 1986. Effects of temperature and composition on the thermal properties of foods. Transp. Phenom. 99–101, URL https://cir.nii.ac.jp/crid/1570291225046275200.

Choi, S.H., Park, K.-B., Roh, D.H., Lee, J.Y., Mohammed, M., Ghasemi, Y., Jeong, H., 2022. An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation. Robot. Comput.-Integr. Manuf. 73, 102258. http://dx.doi.org/10.1016/j.rcim.2021.102258, URL https:// www.sciencedirect.com/science/article/pii/S0736584521001381.

Galeazzi, A., Girotto, F., Rizzoli, L., Prifti, K., Manenti, F., Piazza, L., 2023. Data reconciliation as a key to enable digitalisation of agrifood industrial processes: A preliminary case study. Int. J. Ind. Sustain. Dev. http://dx.doi.org/10.21608/ijisd. 2023.311435, URL https://journals.ekb.eg/article_311435_0.html.

Gilland, B., 2002. World population and food supply: can food production keep pace with population growth in the next half-century? Food Policy 27 (1), 47–63. http:// dx.doi.org/10.1016/S0306-9192(02)00002-7, URL https://www.sciencedirect.com/ science/article/pii/S0306919202000027.

Hassoun, A., Jagtap, S., Garcia-Garcia, G., Trollman, H., Pateiro, M., Lorenzo, J.M., Trif, M., Rusu, A.V., Aadil, R.M., Šimat, V., Cropotova, J., Câmara, J.S., 2023. Food quality 4.0: From traditional approaches to digitalized automated analysis. J. Food Eng. 337, 111216. http://dx.doi.org/10.1016/j.jfoodeng.2022.111216, URL https://www.sciencedirect.com/science/article/pii/S0260877422002709.

Hill, C.A., Scudder, G.D., 2002. The use of electronic data interchange for supply chain coordination in the food industry. J. Oper. Manage. 20 (4), 375–387. http:// dx.doi.org/10.1016/S0272-6963(02)00017-7, URL https://www.sciencedirect.com/ science/article/pii/S0272696302000177.

Koulouris, A., Misailidis, N., Petrides, D., 2021. Applications of process and digital twin models for production simulation and scheduling in the manufacturing of food ingredients and products. Food Bioprod. Process. 126, 317–333. http://dx. doi.org/10.1016/j.fbp.2021.01.016, URL https://www.sciencedirect.com/science/ article/pii/S0960308521000183.

Maheshwari, P., Kamble, S., Belhadi, A., Venkatesh, M., Abedin, M.Z., 2023. Digital twin-driven real-time planning, monitoring, and controlling in food supply chains. Technol. Forecast. Soc. Change 195, 122799. http://dx.doi.org/10. 1016/j.techfore.2023.122799, URL https://www.sciencedirect.com/science/article/ pii/S0040162523004845.

Özyurt, D.B., Pike, R.W., 2004. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. Comput. Chem. Eng. 28 (3), 381–402. http://dx.doi.org/10.1016/j.compchemeng.2003.07.001, URL https: //www.sciencedirect.com/science/article/pii/S0098135403001960.

Perry, R.H., 1999. Perry's Chemical Engineering, seventh ed..

Premalatha, M., Abbasi, T., Abbasi, T., Abbasi, S.A., 2011. Energy-efficient food production to reduce global warming and ecodegradation: The use of edible insects. Renew. Sustain. Energy Rev. 15 (9), 4357–4360. http://dx.doi.org/10. 1016/j.rser.2011.07.115, URL https://www.sciencedirect.com/science/article/pii/ S1364032111003625.

Rivas, B., Merino, J., Caballero, I., Serrano, M., Piattini, M., 2017. Towards a service architecture for master data exchange based on ISO 8000 with support to process large datasets. Comput. Stand. Interfaces 54, 94–104. http://dx. doi.org/10.1016/j.csi.2016.10.004, URL https://www.sciencedirect.com/science/ article/pii/S0920548916301192.

Rodríguez, F., Chicaiza, W.D., Sánchez, A., Escaño, J.M., 2023. Updating digital twins: Methodology for data accuracy quality control using machine learning techniques. Comput. Ind. 151, 103958. http://dx.doi.org/10.1016/j.compind.2023.103958, URL https://www.sciencedirect.com/science/article/pii/S0166361523001082.

Salano, L., Vallerio, M., Moioli, E., Manenti, F., 2025. Industrial scale biogas reforming modelling and validation. Chem. Eng. J. 510, 160871. http://dx.doi.org/10. 1016/j.cej.2025.160871, URL https://www.sciencedirect.com/science/article/pii/ S1385894725016924.

United Nations Sustainable Development Group: Home, 2025. URL https://unsdg.un. org/. van den Berg, F., Lyndgaard, C.B., Sørensen, K.M., Engelsen, S.B., 2013. Process analytical technology in the food industry. Trends Food Sci. Technol. 31 (1), 27–35. http://dx.doi.org/10.1016/j.tifs.2012.04.007, URL https://www.sciencedirect.com/ science/article/pii/S0924224412000933.

Verboven, P., Defraeye, T., Datta, A.K., Nicolai, B., 2020. Digital twins of food process operations: the next step for food process models? Curr. Opin. Food Sci. 35, 79–87. http://dx.doi.org/10.1016/j.cofs.2020.03.002, URL https://www. sciencedirect.com/science/article/pii/S2214799320300217.

Yadav, V.S., Singh, A.R., Raut, R.D., Mangla, S.K., Luthra, S., Kumar, A., 2022. Exploring the application of Industry 4.0 technologies in the agricultural food supply chain: A systematic literature review. Comput. Ind. Eng. 169, 108304. http://dx.doi.org/10.1016/j.cie.2022.108304, URL https://www.sciencedirect.com/ science/article/pii/S0360835222003631.

Zhang, P., Rong, G., Wang, Y., 2001. A new method of redundancy analysis in data reconciliation and its application. Comput. Chem. Eng. 25 (7), 941–949. http:// dx.doi.org/10.1016/S0098-1354(00)00659-1, URL https://www.sciencedirect.com/ science/article/pii/S0098135400006591.

Zhang, Y., Zhang, Y., Gao, M., Dai, B., Kou, S., Wang, X., Fu, X., Shen, W., 2023. Digital twin perception and modeling method for feeding behavior of dairy cows. Comput. Electron. Agric. 214, 108181. http://dx.doi.org/10.1016/j.compag.2023.108181, URL https://www.sciencedirect.com/science/article/pii/S0168169923005690.