10th International Food Operations & Processing Simulation Workshop, 007
21st International Multidisciplinary Modeling & Simulation Multiconference

2724-0355
doi: 10.46354/i3m.2024.foodops.007

A Conceptual Framework for Predictive Digital Dairy Twins: Integrating Explainable AI and Hybrid Modeling

Dana Jox[1], Darius Hummel[2], Jörg Hinrichs[2] and Christian Krupitzer[1,*]

[1]Department of Food Informatics & Computational Science Hub, University of Hohenheim, Fruwirthstr. 21, Stuttgart, 70599, Germany
[2]Department of Soft Matter Science and Dairy Technology, University of Hohenheim, Garbenstr. 21, Stuttgart, 70599, Germany

[*]Corresponding author. Email address: christian.krupitzer@uni-hohenheim.de

Abstract

This vision paper presents a new framework for developing a digital twin for the dairy industry. Unforeseen quality and process challenges that are often specific to each dairy company may arise during processing and must be resolved. Various approaches can be used to solve the different problems, such as food technology experiments or modeling. The framework addresses specific challenges, such as the inconsistency in raw materials and the intricate biochemical dynamics of dairy products. It enables predictive detection and management of critical process states in the products. We use a hybrid modeling approach that integrates white-box models that explain the physical, chemical, and biological mechanisms of dairy processes with black-box machine learning models to improve predictive accuracy and process optimization. The framework centers around implementing explainable artificial intelligence (XAI) to connect empirical food science with advanced data-driven models. XAI is integrated to ensure transparency and interpretability of predictions and decisions made by the digital twin, supporting better decision-making processes within the dairy industry. Our proposed model anticipates quality changes, identifies potential deviations in real time, and elucidates the underlying causes.

Keywords: Digital dairy twin; Physics-informed ML; Explainable ML modeling; Food process analysis

1. Introduction

Dairy processing, and food processing in general, is complex and typically involves numerous processing steps that can vary from company to company. During processing, unforeseen challenges and anomalies can arise that may compromise product quality and the process itself and, therefore, must be resolved. These anomalies are specific to each product and even to each processing plant. For example, in the case of a fermented milk product/dessert/preparation, not only is the composition of the raw milk subject to seasonal fluctuations, but the composition of the starter culture and ingredients such as fruit preparations, spices, and stabilizers are also important during processing. Furthermore, the sequence of process steps and the selected process parameters determine the techno-functional properties such as flow behavior, particle size, voluminosity of the fermented matrix (e.g., Heck et al., 2021; Küçükçetin et al., 2009; Mokoonlall et al., 2016) and these, in turn, correlate with sensory properties such as creaminess (e.g., Cayot et al., 2008; Dickinson, 2018; Frøst and Janhøj, 2007; Sonne et al., 2014).

An increasing amount of structured and semi-structured data is being collected during food processing. This is driven by increasing customer demands, regulatory requirements (e.g., traceability), and advances in process and information technology. Data mining through machine learning (ML) — for computer-aided identification of relationships in data — and intelligent sensors, which are already being used successfully in other industries, have potential in food and dairy processing, e.g., for detection of critical conditions, prediction of quality parameters, or process optimization (Freire et al., 2024; Krupitzer and Stein, 2023).

In the data mining process, different models can be used: black-box, white-box, and grey-box models. Black-box models refer to models in which the relationships between input parameters and outcomes are complex, making it challenging to understand the exact processes through which the final results are derived (e.g., ML models). White-box models provide an explanation of the (physical and chemical) relationships of the model parameters and are the basis for simulations. Grey-box models combine a partial theoretical structure with data to complete the model.

In this context, digital twins are becoming more attractive for companies. One advantage of digital twins is their ability to include real-time and real-world data to identify and manage unexpected states within the food supply chain (Melesse et al., 2023). One challenge of digital twins for food production is that they must consider the processing steps and the chemical, physical, and/or (micro)biological properties as raw materials and processing vary (Henrichs et al., 2022). Therefore, it is important to consider the biochemical and physical properties of the products in an integrated manner to enhance the analytical potential and to analyze the production processes holistically. The implementation of digital twins is majorly complicated due to the absence of a standard methodology that describes transferring information from the physical to the virtual object (Henrichs et al., 2022).

This vision paper specifies a general approach for creating specific digital food twins in the dairy industry to analyze critical process conditions, i.e., critical process parameters or product properties that represent anomalies. We review the challenges to be tackled to implement inherent explainable ML models and a combination of physical (white-box) modeling and ML (black-box) for specific dairy products.

This paper proposes a framework for developing a digital twin for the dairy industry and is organized as follows: Section 2 discusses the state of the art of food and data processing, and digital food twins. In Section 3, we describe our vision of a hybrid digital twin combining white-box and black-box modeling. Sections 4 discusses the research challenges that may arise especially for integrating the XAI (eXplainable Artificial Intelligence) component into the framework. Finally, conclusions, limitations, and future research are briefly discussed in Section 5.

2. State of the art

Currently, white-box approaches that model physical or experimental relationships are widely used. They effectively describe relationships between product characteristics and process steps. However, they are limited by their product-specific nature and can suffer from a lack of transferability to the specific processing of food formulations due to the necessary abstraction of the models. Black-box methods, based on data-driven model identification like ML, face challenges due to a lack of explainability, causality, and IT skills among company employees (Rohleder and Minhoff, 2019). Moreover, companies using these methods often prioritize data over domain expertise, leading to unclear interpretations of correlations found. Currently, there is no integration of both — white-box and black-box — approaches for real-time analysis in the processing of fermented and shelf-stable foods. The following sections focus on the state of the art regarding food processing (Section 2.1), data processing (Section 2.2), and digital food twins (Section 2.3).

2.1. Food processing

In order to monitor food processes, several sensors are implemented in the processing plants. These sensors serve two purposes: I) To keep the process parameters within defined boundaries and II) to directly monitor the product properties in-line. As in-line sensors are a cost factor and their application for real-time monitoring requires a certain level of digitalization, they are not used yet to a sufficient extent.

Dairy processing includes many semi-continuous processes. Therefore, the individual operations can be divided into batch processes, such as fermentation, and continuous processes, such as heating and homogenization. Further, most processes involve intermediate storage of the products to ensure product availability for posterior processing steps to minimize plant downtimes. Batch and continuous processes demand different sensor properties. During batch processes, product deposition on the sensors can lead to inaccurate measurement results, as in contrary no or very slow product movement is required for exact pH-measurement. For the accurate implementation of a digital twin, the reliability of the sensors and the accuracy of the measurement serving as data sources are essential. Most limits in applicability of sensors are set by their CIP (Cleaning In Place) and SIP (Sterilization In Place) capability as well as the necessity and frequency of calibration.

Table 1 gives an overview of various in-line sensors, their application, and relevant performance aspects such as calibration intervals and SIP/CIP capability. However, the calibration interval depends on the demanded accuracy. This, in turn, is defined by the purpose of measurement and decreases in descending order as safety and plant efficiency, regulatory guidelines, supervisory control, or monitoring and optimization.

The following describes the processing of stirred yogurt as an example of a sensor application. To make full-fat yogurt, the fat content of the milk has to be standardized. Therefore, the milk is separated into skimmed milk and cream. Using NIR, the fat content of cream and milk can be

Table 1. Sensor type and application in dairy processing evaluated according to CIP/SIP capability, calibration frequency, and costs for implementation.

Sensor	Application	CIP/SIP capability	Calibration frequency	Costs	Established
Temperature	fermentation, heating, cooling	+	+	+	Yes
Pressure	phase transition, fouling, level control, mechanical processing	+	+	o	Yes
Turbidity	cleaning, liquid phase composition	+	+	o	Yes
Conductivity[1)]	cleaning, liquid phases properties	+	+	o	Yes
Ultrasound	density, composition, fouling	+	+	o/-	Yes
Volume-/mass-flow-	volume-, mass-flow	+	+	o	Yes
pH[1)]	fermentation, cleaning	o/-	-	o/-	Yes
Viscosity[1)]	flow behavior, texture	+	o	o/-	Yes
NIR[1)]	chemical composition	+	-	-	In parts
FBRM[*]	particle size, particle concentration	+	None	-	No
Raman[1)]	chemical composition	+	-	-	No

[1)]Has to be implemented with an additional temperature sensor for accurate measurement.
[*]FBRM - Focused Beam Reflectance Measurement
\+ - high CIP/SIP capability, low calibration frequency, low costs
o - medium CIP/SIP capability, medium calibration frequency, medium costs
\- - low CIP/SIP capability, high calibration frequency, high costs

determined, and both are mixed according to the mass ratio. Further, to obtain the desired gel firmness, the protein content of the milk can be increased by either evaporation, membrane filtration, or the addition of milk powder.

For food safety issues the milk has to be at least pasteurized to inactivate pathogenic microorganisms (Lewis and Deeth, 2009). Temperature control and documentation are critical to fulfill the legal requirements. In addition, whey proteins are heat-denatured in order to increase the gel stability of the fermented product. The extent of denaturation hereby can be controlled by the applied temperature-holding-time (Anema, 2007; Dannenberg and Kessler, 1988). The milk is then homogenized to obtain an even fat distribution in the bulk product and to prevent phase separation. The particle size of the fat globules can be adjusted by the applied homogenization pressure (Kessler, 2002). Subsequently, the milk is cooled to a fermentation temperature of about 40 °C depending on the starter culture and desired cultivation time (Lucey and Singh, 1997). During fermentation pH-value is monitored and the yogurt is then cooled to about 20 °C in order to prevent further acidification (Küçükçetin, 2008; Weidendorfer, 2009). Stirred-type yogurt is then filled. Previous to filling, e.g., the particle size of the yogurt can be controlled using focused beam reflectance measurement (FBRM) (Heck et al., 2023) or its viscosity using in-line differential pressure measurement, volume flow, and the dimensions of the plant (Mönch-Tegeder et al., 2015).

After processing, the plant has to be cleaned. At first, it is rinsed with water to remove residual product. In order to save water, the electrical conductivity of the rinsing water can be measured. Subsequently, the plant is cleaned with alkaline and acidic detergent solution. Especially in the heating section a layer of primarily minerals and proteins can build up (so called fouling), this layer causes a pressure drop over a defined pipe section, that decreases during cleaning (de Jong, 1997; Huppertz and Nieuwenhuijse, 2022). To make sure no detergent residues remain inside the plant, conductivity is measured. In addition, NIR can be used (Vasafi et al., 2021).

2.2. Challenges of data-driven process analysis

Industry 4.0 approaches aim to collect data through sensors and analyze them intelligently using ML algorithms (Usuga Cadavid et al., 2020; Züfle et al., 2022). This involves various data sources, such as raw materials, machinery, or customer data, which can optimize production planning with ML (Cioffi et al., 2020). Another application is the predictive maintenance of machines (Krupitzer et al., 2020). However, both cases mainly focus on the process and machines' perspective. Dogan and Birant (2021) give an overview over machine learning and data mining in manufacturing and their advantages and challenges. Numerical simulations are commonly used in the food industry to simulate products and processes (e.g., Abdul Ghani et al., 1999; Hartmann, 2002; Montanari et al., 2022). However, these approaches involve abstractions that create uncertainties about the validity of the models. These uncertainties can significantly affect the transferability of the models to larger machines (Rauh and Delgado, 2011).

2.3. Digital (food) twins

Generated data can be represented as a digital twin. Digital twins are used to simplify planning by providing a data-driven representation of the product, processes, and machinery. Solutions are offered by providers of production machines. And a concept exists to make digital twins interoperable between (industry) standards (Da Rocha et al., 2022). Tao et al. (2019) provide an overview of digital twins in industrial manufacturing. However, the focus is typically on process optimization in the manufacturing of goods, such as textiles or machinery. In these cases, the products do not change on their own but only through process steps. However, for food products, changes due to physico-chemical and microbiological processes (without obvious representation of the induced changes by the machine operating parameters) are also important and need to be reflected in the digital twin (Krupitzer et al., 2022).

Concepts for digital twins already exist in food processing (Udugama et al., 2023; Krupitzer et al., 2022) and also in the field of dairy (Konstantinidis et al., 2023; Werner et al., 2020). However, they differ in their conceptual ideas. Also, their applicability has yet to be proven, as only one of them has been implemented and focuses only on a management system, not including functionalities like forecasting (Werner et al., 2020).

A promising approach uses a mechanistic model based on multi-physics modeling and simulation (Henrichs et al., 2022). This technique can fully represent the real-world object and, therefore, should be used for prediction (Henrichs et al., 2022). Various projects integrate physical models and numerical simulations to predict changes in food more accurately. For a comprehensive overview, see (Defraeye et al., 2021).

The application of physics-informed ML is especially relevant for creating predictive models for food. This approach integrates ML tools with knowledge-based guided learning to find physically consistent solutions. These hybrid models integrate both physics-based and data-driven models, utilizing different structures or configurations depending on the task/problem (Purlis, 2024). Hybrid modeling aims to create models that explicitly incorporate physical knowledge while using a limited amount of data (Gutschi et al., 2019). This approach can increase accuracy and interpretability with less data than purely data-driven models and uses fewer resources than a purely physics-based approach (Bradley et al., 2022). In the food industry, physics-informed ML is a relatively new topic, but one that is promising (Purlis, 2024).

For digital food twins, when combined with ML, there are two approaches, including explainability to the decision-making of the models: inherent explainability of models and non-explainable models (deep neural networks) extended by an XAI component (Krupitzer et al., 2022). XAI focuses on techniques and algorithms that provide humans with an explanation of how the algorithm identified the result, hence, with insight into the reasoning behind a decision. An XAI component can be integrated for this purpose, e.g., based on food science models and/or simulations (Krupitzer and Stein, 2023). In Krupitzer et al. (2022), we suggest mechanistic modeling and ML for a digital food twin combined with production data, e.g., sensor data, raw material data, and expert knowledge. Such a digital twin could predict and forecast future changes and events in the food and the process (Henrichs et al., 2022).

2.4. Research questions

Our approach combines both white-box and black-box models in a digital food twin using a gray-box method: We focus on data-driven methods (representing black-box approaches) that can be integrated into an XAI approach since such approaches also provide an explanation of how the result was determined. Existing models as well as observations from experiments in the pilot plant (representing white-box approaches) are integrated for the purpose of explanation. The feedback via experiments in the pilot plant integrated into an XAI approach for data analysis represents a new approach. Our methodology should contribute to answer the following questions:

- What ML models are best suited to make process predictions while achieving explainability?
- How can internal data of the production process and external data such as expert knowledge be integrated into a digital food twin?

3. Methodology

Our vision is a hybrid digital twin that combines traditional modeling and simulation of biochemical and physical food properties (white-box approach) with ML analysis for integrating specific information (black-box approach). Unlike common Industry 4.0 approaches for digital twins that mainly analyze machine data (i.e., processes), our concept also includes the simulation of internal product states and validation in the pilot plant. Figure 1 shows a summary of the proposed framework which is described in the following two subsections.

3.1. Integration of simulation and machine learning

To demonstrate the importance of a product perspective, we consider the example of yogurt fermentation. The traditional digital twin concept, as commonly used in Industry 4.0, would not be sufficient because it relies on process data (primarily from machines) to control the production process. Without machine actions, the state of the product in the digital twin will not change. However, yogurt production is based on fermentation, which alters the product and is difficult to monitor in detail. Therefore, data from the process alone cannot adequately describe the process as it occurs within the product. To obtain a more accurate understanding, we supplement the sparse process information with established scientific models that describe the behavior of the bacteria. However, the model alone would not be complete because it is abstract, and, for example, each batch of starter culture has its variations, similar to milk, whose properties vary throughout the year due to various factors such as different feeding.

The digital food twin should contain data from the production site, including process data, physical measurements of intermediate and final products, as well as chemical and microbiological data. This includes sensor, machine, and processing data such as temperature, pressure, or pH value. Raw material data and expert knowledge on handling production problems are also integrated. These data can be used to create a unique product code, or “fingerprint,” of quality characteristics that can be analyzed

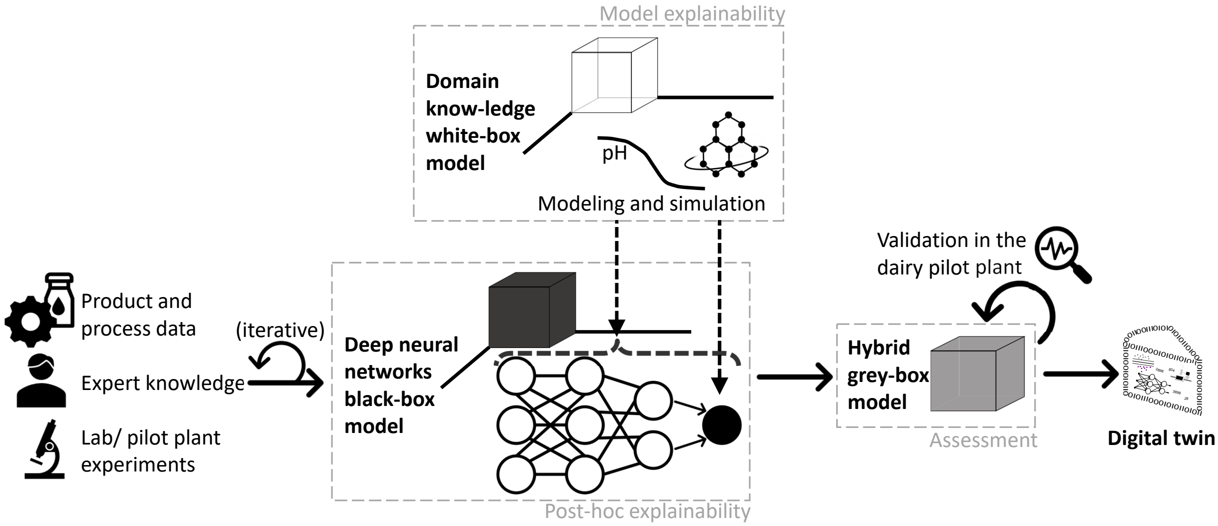

Figure 1. Conceptual summary of the vision of the proposed digital dairy twin framework. The data cover processing, machines, and physical, chemical, and microbiological measurements of raw materials and (intermediate) products. They are the input for the deep neural network. (Food) scientific models, consisting of modeling and (numerical) simulation of biochemical and physical food properties, including internal product states, enrich and explain the neural network. Experiments in the dairy pilot plant validate the resulting model, whose data can also be input into this hybrid model. The digital twin is created using the model and post-hoc explainability and assessment.

in real-time. This can then be used to compare the current production with the "normal state" (historical data) to detect critical process conditions and/or make predictions about quality anomalies that could lead to quality deviations, such as reduced shelf life. Using various simulation methods based on chemical-physical models and numerical simulations from food science, the digital twin provides information about the actual food processing and supports food processing operations with real-time feedback.

3.2. Using an XAI component

To create the digital twin, we intend to use explainable artificial intelligence. Four axes can contribute to the explainability of all levels of the AI process (Ali et al., 2023):

1. Data explainability: Feature engineering and other summarizing and analyzing techniques can be used to decide which data to use for training models.
2. Model explainability: Model explainability targets the understanding of the internal structures of the model. Some ML methods are inherently explainable, such as decision trees or random forests. However, these methods may have limitations when dealing with large datasets and do not support automatic feature extraction, as is the case with deep learning methods.
3. Post-hoc explainability: Determination of which features were significant and used to make decisions.
4. Assessment: Assessment approaches evaluate and compare explanatory approaches by using various desired properties, such as explanatory power.

Our approach focuses on a combination of the second to fourth axes. However, it is also important to consider data explainability, for example, during data selection and pre-processing. For complex approaches such as deep neural networks, the idea is to use a second component, the XAI component, which tries to explain the results using domain knowledge in the form of white-box models. We focus on more complex models, as higher accuracy is expected with more complex models. The challenge is that the degree of accuracy and the degree of interpretability of a model are often contradictory (Ali et al., 2023).

Figure 2 illustrates our XAI approach. We want to transform the partially existing "black box" of production data into a flexible "grey box" by means of a software method based on XAI models, and integrating simulations and the latest scientific models (white box). The results obtained through model explainability and post-hoc explainability methods can be used to conduct specific experiments in the laboratory/ pilot plant that are relevant to the dairy product. The data derived from these experiments can subsequently be integrated into the current XAI model. Therefore, our approach enables the (iterative) incorporation of external information such as raw material and final product properties (chemical, microbiological, physical, and sensory) obtained from laboratory tests, storage tests, and market feedback (e.g., complaints). These hybrid approaches can enhance the explanatory power of black-box approaches and reduce the ambiguity of white-box models by integrating the collected data. The assessment can improve learning and systematically analyze models, which helps to understand and validate them.

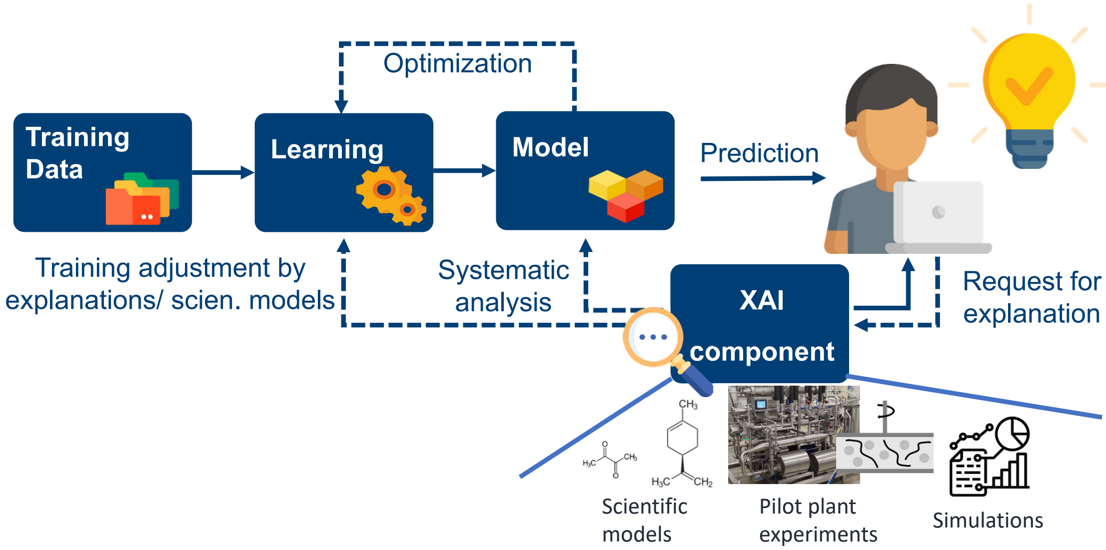

Figure 2. Prediction process using machine learning with an additional XAI component to explain the results to the users (based on Krupitzer et al., 2022).

4. Challenges and discussion

Multiple research challenges may arise for using XAI for process predictions. In Krupitzer et al. (2022), we discussed use cases and challenges for hybrid digital twins, but did not address the XAI component in detail. Therefore, it is important to consider the best-suited XAI approach for data analysis using ML algorithms, which includes identifying physical-chemical models for the XAI component and developing methodologies that integrate domain knowledge into ML models. Validating data and integrating data from laboratory or pilot plant experiments into a digital twin requires rigorous data quality checks and the use of sophisticated data integration techniques. Additionally, it is necessary to investigate how models that include an XAI component can be generalized and adapted to various datasets and scenarios in the pursuit of creating a general dairy twin.

Hence, we discuss those challenges in the following and sketch possible solutions focusing on white-box (Section 4.1), black-box (Section 4.2), and grey-box (Section 4.3) challenges. Subsequently, we will further discuss the general importance and relevance of our approach (Section 4.4).

4.1. White-box challenges

Creating digital dairy twins using white-box models is difficult due to the complexity of food and its processing (e.g., Heck et al., 2021; Küçükçetin et al., 2009; Mokoonlall et al., 2016; Olivares et al., 2013). However, inherently explainable ML models have an advantage as their explanatory power is built-in. The challenge lies in data selection and creating features that form properties for the ML (Freire et al., 2024). Expert and ML knowledge can improve the explainability of results through feature selection and preprocessing.

Data collection and selection are important to create a digital twin for a specific dairy product (Henrichs et al., 2022). It is necessary to define the process line(s) and associated raw materials, intermediates, and finished products. When selecting those data, the focus should be on identifying all quality-related parameters and, hence, choosing (and, if necessary, installing) appropriate analysis methods for the physical and chemical characterization of the products. One exemplary product property, varying frequently, is the texture, which can be analyzed with rheological measurements. This also includes regularly occurring textural anomalies (which may require countermeasures during production or can be seen in complaints) and should be done in collaboration with the company and based on relevant literature (e.g., Heck et al., 2021). After the features are generated from the raw data, the ML algorithms are used to explore the properties and relationships of these features in order to learn ML models that represent the relationships, such as the relationship between the raw data and the quality characteristics of the final product. The processes of many dairy products are already described in detail, and many relationships are known. Hence, they can be used to validate the found relationships.

To determine the most suitable model, it is necessary to first select potentially suitable ML algorithms and configure their parameters (hyperparameter tuning, e.g., with optimization approaches (Brownlee, 2011)). These algorithms should then be validated using standardized test procedures, particularly k-fold cross-validation. Ensemble tree approaches, such as Random Forest (RF) and Gradient Boosting (e.g. XGBoost), can be utilized for this purpose. These algorithms offer high performance in multidimensional scenarios with comparatively short training times. For instance, RF has produced robust results in various domains (Züfle et al., 2019, 2022; Krupitzer et al., 2018). Our recommendation service for selecting an algorithm for time series prediction (Züfle et al., 2019) could serve as a foundation for creating an ML pipeline that automatically selects the algorithm and its parameters based on the data patterns. Therefore, the ML pipeline produces appropriate ML algorithms and parameter configurations to train the ML models for the data patterns. Afterwards, the optimal model is selected.

4.2. Black-box challenges

Deep learning methods are very popular due to their high performance and ability to build features autonomously. Artificial neural networks in the form of recurrent neural networks can account for even sequential dependencies between data points. Hence, they can be used even in the absence of physico-chemical models. However, they require large amounts of data, lack transparency, and identifying which features are most relevant for the model can be challenging, especially when dealing with complex and high-dimensional data from food processes.

To achieve a sufficient amount of data, data can be created artificially or must be collected. For example, the SMOTE approach (Chawla et al., 2002) artificially increases the amount of data (data augmentation). For this purpose, we suggest using our previous work on data processing automation (Züfle et al., 2019, 2022).

For permanent data collection, in-line sensors could monitor dairy product properties. The production goal differs from the research. It aims to produce a matrix or final product with minimal variance despite variations in raw materials and process conditions. Texture properties of fermented dairy products are seldom used in quality control or in-line processing (e.g., Mönch-Tegeder et al., 2015; Weidendorfer and Hinrichs, 2011) due to laboratory costs, leading to reliance on legally required chemical and microbiological analyses and qualitative sensory evaluations by staff. In general, physical product data (e.g., viscosity or particle size) are rarely collected systematically in companies, although they would be suitable in addition to chemical markers to analytically differentiate the end products (e.g., Janhøj et al., 2006; Hartmann et al., 2015a,b,c; Schenkel et al., 2014) and to predict possible quality deviations (e.g., Vasafi et al., 2021). Hence, access to texture information during processing could allow for adjustments to compensate for matrix variances.

Post-hoc explainability approaches the challenge of transparency and feature identification. For example, layer-wise relevance propagation (LRP) (Wu et al., 2022) and model-agnostic methods like shapley additive explanations (SHAP) (Lundberg and Lee, 2017) and local interpretable model-agnostic explanations (LIME) (Peng et al., 2022), both of which can interpret the results, can be used to explain the models. Those post-hoc approaches, especially the LIME method, are primarily optimized for image data. Hence, significant challenges are present when adapting to other data types or integrating with white-box models. The adaptation of those post-hoc approaches is not straightforward and requires careful consideration of compatibility and interoperability.

4.3. Grey-box challenges

For the grey-box model, the aim is to integrate the existing knowledge from companies (e.g., expert knowledge of employees) and research (e.g., scientific models) into the XAI component to explain the results of machine/deep learning algorithms. This can be done in two phases.

The first phase focuses on modeling the required domain knowledge using white-box models. The goal is to integrate existing food science knowledge with physical simulation models, results from newly implemented texture analysis methods, and dedicated laboratory experiments.

In the second phase, we focus on the ML perspective and compare two methods for XAI. On the one hand, artificial neural networks are used for deep learning combined with post-hoc explainability approaches. On the other hand, they are compared to hybrid systems that combine rule-based selection methods with ML. The rules are derived from the white-box models. The XAI component uses these rules to control the learning process and to keep the decision traceable. Since the generation of rules is very individual, this approach is the least automatable but probably delivers the best results, as the results of deep learning methods can be explained.

A validation in the laboratory is conducted to explain the uncertainties of the data-driven ML methods from the black-box models and to generate knowledge to increase the explainability of the ML algorithms through experiments. First, products are manufactured with process conditions, raw materials, company ingredients, etc., and tested for reproducibility (3 to 5 repetitions). This provides reference data from the pilot plant. Then, anomalies in the data can be explained, or ML analyses can be verified, e.g., by varying process conditions below and above the reference. The main purpose of the experiments is to validate the analysis results of the white- and black-box ML models. Of course, a difference in scale (factory vs. pilot plant) has to be considered. We plan to apply the explanations of the XAI approach to find out how to generalize from pilot plant to factory for handling the differences in scale.

Since the validation of the results in the laboratory can change the ML methods, especially regarding the XAI components, the results have to be verified with production data in cooperation with the company. On the one hand, we focus on evaluating the quality of the ML algorithms (sensitivity, specificity, mean square error, etc.) and their resource consumption (time, CPU, RAM, etc.). For this purpose, new data collected but not used in the training phase can be analyzed. Furthermore, an empirical evaluation of the comprehensibility of the ML algorithms' explanations for the employees should be implemented.

4.4. General importance and relevance

There is still no general concept of a digital dairy twin that includes predictive capabilities through combining machine learning with domain knowledge and explicitly addresses explainability. Our approach fills this gap. It is a relevant new concept because it includes and improves the traceability of the digital dairy twin's decisions through iterative data and knowledge integration. For example, learning the white-box ML model requires the creation of features. If the performance of the model is insufficient, the features are revised. Similarly, data from experiments in the pilot plant influence the grey-box models for the XAI components.

It is possible to transfer the model of one digital dairy twin representing a specific product to another that represents a dairy product with similar processes using transfer learning (Lisa Torrey and Jude Shavlik, 2010). The scientific models of the digital dairy twin must be changed for different dairy products or processes. However, the general structure and procedure for establishing the digital dairy twin can be adapted.

5. Conclusions

Our framework provides the practical basis for comprehensive analysis of existing data with XAI incorporating expert knowledge. It can be used to extract raw material data and process parameters for the optimization of production processes in real-time. In addition, it can support root cause analysis and accelerate the scaling up of new product variants/ideas on existing process lines. Hence, the implementation of our approach results in two concepts:

- a concept for recommendation services that decide which ML algorithm is best suited for a given set of characteristics, and
- a framework for automated analysis using XAI.

This enables the use of ML methods and the interpretation of their results even without knowledge of data mining or data science. Therefore, our framework allows answering the research questions for the most appropriate ML models to make process predictions while achieving explainability using an XAI component and integrating additional data into the digital food twin. Hence, our approach constitutes the implementation in the industry, which can utilize our methodology as a foundation for automating the steps of ML analysis on their own. The automated data analysis offers further advantages for dairy companies:

1. Variations in product characteristics can be quickly identified using the digital dairy twin, which also reduces the effort required for root cause analysis.
2. The time-to-market of products can be reduced through sensitivity-oriented adjustments in the process, using knowledge from historical data, laboratory experiments, or pilot plant results.
3. Proactive detection of critical system states enables continuous verification, as conclusions about critical system parameters can be drawn from collected data, but also from complaints.

Certain prerequisites must be met within the industry to use our approach: Many companies are already collecting data. However, data is often merely archived, and it must be evaluated whether the IT infrastructure already enables real-time data provision. Our approach can be implemented using open-source software, allowing for adaptation by individual companies. This enables the cost-effective implementation of a decision-support tool for selecting parameters in food processing operations. In addition to the need for extensive relevant data, other limitations may include unsuccessful pilot plant verification, inability to deploy the XAI component, or overall unsuccessful modeling.

Based on our approach, optimized control (during operation) and, in the future, targeted control (in the event of a raw material/formulation change or when a new product is introduced to the line) of process parameters based on intelligent data analysis can be facilitated. Therefore, our framework can be viewed as the initial phase of a strategic transformation towards adaptive food production.

Funding

The IGF Project (AiF 22840 N) of the Research Association of the German Food Industry (FEI) was supported via AiF within the program for promoting the Industrial Collective Research (IGF) of the German Federal Ministry for Economic Affairs and Climate Action (BMWK), based on a resolution of the German Parliament.

References

Abdul Ghani, A. G., Farid, M. M., Chen, X. D., and Richards, P. (1999). Numerical simulation of natural convection heating of canned food by computational fluid dynamics. *Journal of Food Engineering*, 41(1):55–64.

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., and Herrera, F. (2023). Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion*, 99:101805.

Anema, S. G. (2007). Role of kappa-casein in the association of denatured whey proteins with casein micelles in heated reconstituted skim milk. *Journal of agricultural and food chemistry*, 55(9):3635–3642.

Bradley, W., Kim, J., Kilwein, Z., Blakely, L., Eydenberg, M., Jalvin, J., Laird, C., and Boukouvala, F. (2022). Perspectives on the integration between first-principles and data-driven modeling. *Computers & Chemical Engineering*, 166:107898.

Brownlee, J. (2011). *Clever algorithms: Nature-inspired pro-*

gramming recipes. London, revision 2 edition.

Cayot, P., Schenker, F., Houzé, G., Sulmont-Rossé, C., and Colas, B. (2008). Creaminess in relation to consistency and particle size in stirred fat-free yogurt. *International Dairy Journal*, 18(3):303–311.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357.

Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., and de Felice, F. (2020). Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability*, 12(2):492.

Da Rocha, H., Pereira, J., Abrishambaf, R., and Espirito Santo, A. (2022). An interoperable digital twin with the ieee 1451 standards. *Sensors (Basel, Switzerland)*, 22(19):7590.

Dannenberg, F. and Kessler, H.-G. (1988). Reaction kinetics of the denaturation of whey proteins in milk. *Journal of Food Science*, 53(1):258–263.

de Jong, P. (1997). Impact and control of fouling in milk processing. *Trends in Food Science & Technology*, 8(12):401–405.

Defraeye, T., Shrivastava, C., Berry, T., Verboven, P., Onwude, D., Schudel, S., Bühlmann, A., Cronje, P., and Rossi, R. (2021). Digital twins are coming: Will we need them in supply chains of fresh horticultural produce? *Trends in Food Science & Technology*, 109:245–258.

Dickinson, E. (2018). On the road to understanding and control of creaminess perception in food colloids. *Food Hydrocolloids*, 77:372–385.

Dogan, A. and Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166:114060.

Freire, P., Freire, D., and Licon, C. C. (2024). A comprehensive review of machine learning and its application to dairy products. *Critical reviews in food science and nutrition*, pages 1–16.

Frøst, M. B. and Janhøj, T. (2007). Understanding creaminess. *International Dairy Journal*, 17(11):1298–1311.

Gutschi, C., Furian, N., Suschnigg, J., Neubacher, D., and Voessner, S. (2019). Log-based predictive maintenance in discrete parts manufacturing. *Procedia CIRP*, 79:528–533.

Hartmann, C. (2002). Numerical simulation of thermodynamic and fluid-dynamic processes during the high-pressure treatment of fluid food systems. *Innovative Food Science & Emerging Technologies*, 3(1):11–18.

Hartmann, K. I., Bähler, B., and Hinrichs, J. (2015a). Structure-process-function relationship during heating of natural cheese. *European Dairy Magazine*, 27(3).

Hartmann, K. I., Dunkel, A., Hillmann, H., Hansen, D., Schieberle, P., Hofmann, T., and Hinrichs, J. (2015b). Identification of physical properties and volatile and non-volatile compounds for discrimination between different emmental-type cheeses: a preliminary study. *Dairy Science & Technology*, 95(5):701–717.

Hartmann, K. I., Samudrala, R., Hofmann, T., Schieberle, P., Hitzmann, B., and Hinrichs, J. (2015c). Thermophysical parameters applied for instrumental profiling and statistical evaluation of german emmental cheese. *International Dairy Journal*, 49:118–124.

Heck, A., Nöbel, S., and Hinrichs, J. (2023). Inline particle size analysis during technical-scale processing of a fermented concentrated milk protein-based microgel dispersion: Feasibility as a process control. *Dairy*, 4(1):180–199.

Heck, A. J., Schäfer, J., Nöbel, S., and Hinrichs, J. (2021). Fat-free fermented concentrated milk products as milk protein-based microgel dispersions: Particle characteristics as key drivers of textural properties. *Comprehensive reviews in food science and food safety*, 20(6):6057–6088.

Henrichs, E., Noack, T., Pinzon Piedrahita, A. M., Salem, M. A., Stolz, J., and Krupitzer, C. (2022). Can a byte improve our bite? an analysis of digital twins in the food industry. *Sensors*, 22(1).

Huppertz, T. and Nieuwenhuijse, H. (2022). Constituent fouling during heat treatment of milk: A review. *International Dairy Journal*, 126:105236.

Janhøj, T., PETERSEN, C. B., Frøst, M. B., and IPSEN, R. (2006). Sensory and rheological characterization of low–fat stirred yogurt. *Journal of Texture Studies*, 37(3):276–299.

Kessler, H.-G., editor (2002). *Food and bio process engineering: Dairy technology*. Verlag A. Kessler, München, 5th rev. and extended edition.

Konstantinidis, F. K., Balaska, V., Symeonidis, S., Psarommatis, F., Psomoulis, A., Giakos, G., Mouroutsos, S. G., and Gasteratos, A. (2023). Achieving zero defected products in diary 4.0 using digital twin and machine vision. *PETRA '23: Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments*, pages 528–534.

Krupitzer, C., Noack, T., and Borsum, C. (2022). Digital food twins combining data science and food science: System model, applications, and challenges. *Processes*, 10(9):1781.

Krupitzer, C. and Stein, A. (2023). Unleashing the potential of digitalization in the agri-food chain for integrated food systems. *Annual review of food science and technology*.

Krupitzer, C., Sztyler, T., Edinger, J., Breitbach, M., Stuckenschmidt, H., and Becker, C. (2018). Hips do lie! a position-aware mobile fall detection system. In *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE.

Krupitzer, C., Wagenhals, T., Züfle, M., Lesch, V., Schäfer, D., Mozaffarin, A., Edinger, J., Becker, C., and Kounev, S. (2020). A survey on predictive maintenance for industry 4.0. *arXiv preprint arXiv:2002.08224*.

Küçükçetin, A. (2008). Effect of heat treatment of skim milk and final fermentation ph on graininess and roughness of stirred yogurt. *International Journal of Dairy Tech-*

nology, 61(4):385–390.

Küçükçetin, A., Weidendorfer, K., and Hinrichs, J. (2009). Graininess and roughness of stirred yoghurt as influenced by processing. *International Dairy Journal*, 19(1):50–55.

Lewis, M. J. and Deeth, H. C. (2009). Heat treatment of milk. In Tamime, A. Y., editor, *Milk Processing and Quality Management*, Society of Dairy Technology Series, pages 168–204. John Wiley & Sons, New York, NY.

Lisa Torrey and Jude Shavlik (2010). Transfer learning. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, pages 242–264. IGI Global.

Lucey, J. A. and Singh, H. (1997). Formation and physical properties of acid milk gels: A review. *Food Research International*, 30(7):529–542.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Melesse, T. Y., Franciosi, C., Di Pasquale, V., and Riemma, S. (2023). Analyzing the implementation of digital twins in the agri-food supply chain. *Logistics*, 7(2):33.

Mokoonlall, A., Nöbel, S., and Hinrichs, J. (2016). Postprocessing of fermented milk to stirred products: Reviewing the effects on gel structure. *Trends in Food Science & Technology*, 54:26–36.

Mönch-Tegeder, M., Lemmer, A., Hinrichs, J., and Oechsner, H. (2015). Development of an in-line process viscometer for the full-scale biogas process. *Bioresource technology*, 178:278–284.

Montanari, R., Volpi, A., Solari, F., and Lysova, N. (2022). Performance assessment, through cfd simulation, of an ultraviolet reactor for the processing of fruit juices. In *Proceedings of the 8th International Food Operations and Processing Simulation Workshop, FoodOPS*.

Olivares, M. L., Berli, C. L. A., and Zorrilla, S. E. (2013). Rheological modelling of dispersions of casein micelles considered as microgel particles. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 436:337–342.

Peng, H., Yang, R., Wang, Z., Li, J., He, L., Yu, P. S., Zomaya, A. Y., and Ranjan, R. (2022). Lime: Low-cost and incremental learning for dynamic heterogeneous information networks. *IEEE Transactions on Computers*, 71(3):628–642.

Purlis, E. (2024). Physics-informed machine learning: the next big trend in food process modelling? *Current Food Science and Technology Reports*, 2(1):1–6.

Rauh, C. and Delgado, A. (2011). Limitations of mathematical modelling and numerical simulation of industrial and laboratory high-pressure processes. *High Pressure Research*, 31(1):126–130.

Rohleder, B. and Minhoff, C. (2019). Die Ernährung 4.0 – Status Quo, Chancen und Herausforderungen. Technical report, Bitkom und Bundesvereinigung der Deutschen Ernährungsindustrie.

Schenkel, P., Hartmann, K. I., Samudrala, R., and Hinrichs, J. (2014). Characterization of commercial cheese regarding thermophysical properties and application of multivariate statistical analysis to elaborate product mapping. *Journal of Texture Studies*, 45(6):440–451.

Sonne, A., Busch-Stockfisch, M., Weiss, J., and Hinrichs, J. (2014). Improved mapping of in-mouth creaminess of semi-solid dairy products by combining rheology, particle size, and tribology data. *LWT - Food Science and Technology*, 59(1):342–347.

Tao, F., Zhang, H., Liu, A., and Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4):2405–2415.

Udugama, I. A., Kelton, W., and Bayer, C. (2023). Digital twins in food processing: A conceptual approach to developing multi-layer digital models. *Digital Chemical Engineering*, 7:1–9.

Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., and Fortin, A. (2020). Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing*, 31(6):1531–1558.

Vasafi, P. S., Paquet-Durand, O., Brettschneider, K., Hinrichs, J., and Hitzmann, B. (2021). Anomaly detection during milk processing by autoencoder neural network based on near-infrared spectroscopy. *Journal of Food Engineering*, 299(1–2):110510.

Weidendorfer, K. (2009). *Die Gelpartikelsuspension Rührjoghurt: Mechanisch-dynamische Eigenschaften, Inline-Partikelgrößenmessung und Apparate zum Glätten der Textur: Zugl.: Hohenheim, Univ., Diss., 2008*. Lebensmitteltechnologie. Verl. Dr. Hut, München, 1. edition.

Weidendorfer, K. and Hinrichs, J. (2011). On-line microgel particle size measurement in stirred skim milk yoghurt. *Milchwissenschaft*, 66(2):152–155.

Werner, R., Beugholt, A., Takacs, R., Geier, D., Becker, T., Sollacher, R., Mauermann, M., Weißenberg, N., Roest, M., and Istaith, J. (2020). Standardized digitalization of an existing pudding production: Integration of a digital twin management system. *International Dairy Magazine*, (5):22–25.

Wu, H., Huang, A., and Sutherland, J. W. (2022). Layerwise relevance propagation for interpreting lstm-rnn decisions in predictive maintenance. *The International Journal of Advanced Manufacturing Technology*, 118(3-4):963–978.

Züfle, M., Bauer, A., Lesch, V., Krupitzer, C., Herbst, N., Kounev, S., and Curtef, V. (2019). Autonomic forecasting method selection: Examination and ways ahead. In *2019 IEEE International Conference on Autonomic Computing (ICAC)*. IEEE.

Züfle, M., Moog, F., Lesch, V., Krupitzer, C., and Kounev, S. (2022). A machine learning-based workflow for automatic detection of anomalies in machine tools. *ISA transactions*, 125:445–458.