engineering proceedings

MDPI

Proceeding Paper

DigiFoodTwin: Digital Biophysical Twins Combined with Machine Learning for Optimizing Food Processing †

Christian Krupitzer * and **Tanja Noack**

Food Informatics Department & Computational Science Lab, University of Hohenheim, 70599 Stuttgart, Germany; noack.tanja@outlook.com

* Correspondence: christian.krupitzer@uni-hohenheim.de; Tel.: +49-711-459-23664

† Presented at the 1st International Electronic Conference on Processes: Processes System Innovation, 17–31 May 2022; Available online: https://ecp2022.sciforum.net.

Abstract: Production processes must allow high flexibility and adaptivity to ensure food supply. This includes reacting to disruptions in the supply of ingredients, as well as the varying quality of ingredients, e.g., seasonal fluctuations of raw material quality. Digital twins are known from Industry 4.0 as a method to model, simulate, and optimize processes. In this vision paper, we describe the concept of a digital food twin. Due to the variability of these raw materials, such a digital twin has to take into account not only the processing steps, but also the chemical, physical, or microbiological properties that change the food independent of the processing. We propose a model-based learning and reasoning loop, which is known from self-aware computing (SeAC) systems in the so-called learn–reason–action loop (LRA-M loop), for modeling the input for the LRA-M loop of food production, not as a pure knowledge database, but data that are generated by simulations of the bio-chemical and physical properties of food. This work presents a conceptual framework on how to include data provided by a digital food twin in a self-aware food processing system to respond to fluctuating raw material quality and to secure food supply and discusses the applicability of the concept.

Keywords: digital twin; food processing; Industry 4.0; self-awareness computing systems; artificial intelligence

Citation: Krupitzer, C.; Noack, T. DigiFoodTwin: Digital Biophysical Twins Combined with Machine Learning for Optimizing Food Processing. *Eng. Proc.* **2022**, *19*, 42. https://doi.org/10.3390/ECP2022-12623

Academic Editor: Dariusz Dziki

Published: 19 May 2022

1. Introduction

The term Industry 4.0 refers to current technological changes in the environment of industrial production enabled by advances in information technology. The focus of Industry 4.0 is the smart factory, i.e., the connection of cyber–physical production systems with Internet of Things (IoT) technology, as well as intelligent data analysis. A core element of Industry 4.0 is the digital twin: a virtual model of a product, the machines, or the production process created with data collected by sensors that enables simulations or real-time analyses of the status of production. As a digital twin integrates real-time data, it provides a detailed simulation model that can support decision-making.

The use of digital twins seems beneficial in food processing for various reasons. The Coronavirus pandemic demonstrated the vulnerability of food supply resilience. To ensure the supply of food, production processes must allow high flexibility and adaptivity, which require traceability. The survey "Die Ernährung 4.0—Status Quo, Chancen und Herausforderungen" (Nutrition 4.0—Status Quo, Opportunities and Challenges) by the digital association Bitkom and the Federation of German Food and Drink Industries (BVE) showed that 70% of the more than 300 companies surveyed in the food industry consider end-to-end traceability from the origin of the goods to the customer to be an important scenario for the current decade [1]. Various types of sensors exist to support this. However, the potential is far from being exploited. Furthermore, product quality is influenced by different quality levels of input materials. Especially in the case of seasonal fluctuations of this raw material quality, an adjustment of the parameters in the production process is essential. Introducing new products that are related to existing ones is also a challenge in food processing. Introduction processes of new products could be simplified by a digital twin of already existing products. The digital twin is able to learn the correct process parameters for production and is used as a knowledge foundation within a self-adaptive system [2]. All those application scenarios show the potential of digital twins in the food supply chain.

However, a digital twin of food production has additional specific requirements compared to digital twins of the production of material goods. Due to the variability of raw materials, these cannot be based only on the processing steps, but must also take into account the chemical, physical, or (micro)biological properties of the food. This vision paper aims to provide a concept that complements the typical, retrospective analysis of machine and process data with short-term (detection of potential problems) and medium-term data analysis approaches (planning and optimization), as well as product-related analysis for achieving a proactive decision-making of adaptation in food production and tracking the current state of production at any time. In contrast to common Industry 4.0 approaches, this paper aims to include a product-related data analysis. While Industry 4.0 approaches often focus on the analysis of machine data, this paper describes a product-related data analysis as well. Such an analysis can be the foundation for an adaptive system that is able to control the process, autonomously react to changes, and continuously improve its performance through learning. Consequently, such a concept helps to better (i) understand the behavior of a food production process, (ii) predict critical situations, and (iii) determine a new plan.

The remainder of the paper is structured as follows. Next, Section 2 describes current approaches in the literature. Afterwards, Section 3 presents our concept for a digital food twin. Then, Section 4 discusses research challenges for the implementation of our concept. Finally, Section 5 concludes this paper.

2. Materials and Methods

This section presents several approaches and concepts that we identified in the literature and that are relevant to the field of digital twins for the food processing industry.

Smart factory in the food industry. Current approaches in Industry 4.0 focus on the intelligent collection of data with technology from the IoT and their analysis with machine learning algorithms [3]. This includes a variety of data sources, including raw material data, machine data, or customer data. In particular, production planning can be optimized with machine learning in this context [4]. Another use case is predictive maintenance of machines [5,6]. However, the focus is primarily on the view of the process and the machines. Internal processes in the food industry are not included, and the view of the product is limited to identifying products with bar or QR codes. Proactive adaptation improves system performance as it forecasts adaptation concerns (e.g., through identification of patterns in historical data) and reacts either by preparing an adaptation or adapting [7]. Real-time data of production sites would help to realize proactive adaptation and dynamic adjustment when a disruption takes place.

Digital twins in the food sector. Digital twins can be classified into six types—(i) imaginary, which simulate reference objects, (ii) a digital twin, which monitors in real-time the state and behavior of an object, (iii) predictive, which projects future states and behaviors of an object, (iv) prescriptive and (v) autonomous digital twins (using artificial intelligence), and (vi) a recollection digital twin with historical data [8]. However, there are still few concepts for digital twins specialized for food processing. Further, in a recent review [9], we showed that agri-food digital twins are limited to specific aspects (e.g., animal monitoring, crop management, or hydroponics), rather than generically applicable throughout the value chain. Most closely related to our concept, the *smartFoodTechnologyOWL* initiative investigates the transferability of the digital twin concept to food processing. The focus is on mapping the process for better control of cyber–physical production systems. In order to make quality control of food safer and more efficient, their goal is to continuously generate a "virtual image" of the product during production.

Other projects focus on the integration of physical models to better predict the changes to the food through its processing. In [10], the authors describe the integration of physical, biochemical, and microbiological processes. However, this type of digital twin often lacks the data-driven perspective of the processes, and [10] propose to include real-time coupling of sensor data with the digital twin. That would help to foresee problems and proactively react to them. However, the focus is not on adapting the production process based on the gained information, nor on processing the data for predicting critical events. Digital twins are used in production for monitoring a production process [11]. Autonomous systems can respond to changes in state during ongoing operation, while digital twins can integrate a variety of data such as environment data, operational data, and process data [11,12]. Today, food process modeling has mostly pure design optimization and cost targets, but there is a great potential in reducing inter-product variability, achieving higher transparency, and reducing the use of resources [13].

Sensors and indicators: With the help of *indicators*, the presence or absence of a substance, reactions between different substances, or the concentration of a particular substance can be detected. Indicators show the analysis results by direct changes (usually different color intensities) and are placed inside or outside the packaging. Different types of indicators exist. The most common types are time–temperature indicators, which show that critical temperatures have been reached; freshness indicators, which monitor the quality of food products based on microbiologically motivated or chemical changes in the products; and gas indicators, which detect changes in the atmosphere of the package. In contrast to immutable indicators that cannot be reused once they have changed their state, *sensors* that are either integrated into the food packaging or in the environment can detect temperature, humidity, pressure on food, or vibrations (accelerometers). Specific sensors such as gas sensors or biosensors measure the concentration of certain gases such as carbon dioxide (CO_2) or hydro-sulfuric acid, which allow conclusions to be drawn about perishability. The CO_2 concentration can be measured using non-dispersive infrared (NDIR) sensors or chemical sensors; infrared sensors, as well as electrochemical, ultrasonic, and laser technologies are used to detect the oxygen concentration. Another type of sensors is biosensors based on receivers made of biological materials such as enzymes, antigens, hormones, or nucleic acids. These are used, for example, to identify pathogens such as salmonella, *E. coli*, or listeria. The overview in [14] describes the recent state-of-the-art in sensor and indicator types. Especially, sensors facilitate real-time data collection, which supports building digital twins.

Contribution. In the case of the food supply chain, a detailed model of the supply chain, which integrates real-time data to predict supply chain dynamics, can be a promising concept to respond to unexpected events in the whole supply chain including field, factory, retailer, and consumer. The goal of our project is to create a digital food twin that can be used to track the current state of production at any time. While Industry 4.0 approaches often focus on the analysis of machine data, this project aims at also including a product-related data analysis (e.g., the effects of pressure exerted by machines). Recent work is conducted on self-aware computing (SeAC) systems, especially to extract models from data and use these models to define adaptations of a system or process, as well as on digital twins in the food sector, but not in a combined approach to intelligently generate a digital twin and use this digital twin for reasoning. The main contribution is to provide a framework that includes data provided by a digital food twin in real-time in an SeAC system. The sensor measurements are complemented by forecasting methods, continuous simulation, and critical event prediction to act as a knowledge base for a self-aware learning and reasoning loop (LRA-M loop) and enable adaptive, resilient food processing.

3. System Design

This paper presents and discusses a concept that complements the typical, retrospective analysis of supply chain data with short-term (detection of potential problems) and medium-term data analysis approaches (planning and optimization) to achieve real-time, predictive decision-making of adaptation in the food supply chain. Consequently, such a concept helps to better (i) understand the behavior of a supply chain, (ii) predict critical situations, and (iii) determine a new action plan.

With the help of machine learning and artificial intelligence, the digital twin is generated from production data and additional data sources (e.g., scientific models, process data, or raw material data) to ensure the traceability of the production and the food status, but also to enable the simulation of the variability of the food in the process operation.

Figure 1 shows our concept of the digital twin. In the figure and the following, we focus on the example of a dairy product (e.g., cheese). The digital twin obtains its data from the production site (e.g., sensor, machine, and processing data) and also integrates raw material data, complaints, and knowledge from experts (e.g., about the handling of production issues). Using different simulation methods based on models from food science, the digital twin provides information about the actual food processing and provides real-time feedback to the food process operation, but could also use those simulations based on scientific models to generate forecasts on how the process steps might influence the quality of the product. Accordingly, the digital twin is suitable for retrospective, but also predictive analytics of the process and the quality of the product.

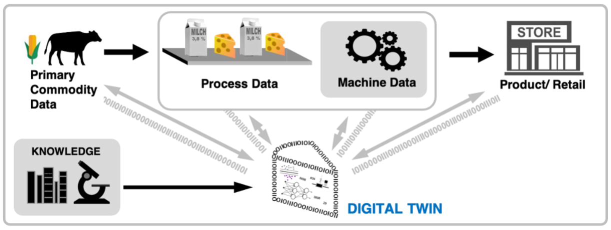

Figure 1. The digital food twin, which integrates data from various sources.

For constructing the digital twin, we rely on machine learning procedures, especially from the field of explainable artificial intelligence (XAI). Such approaches helps to transform the sensor data into a digital twin model, which can be used for simulation. Further, in contrast to approaches based on artificial neural networks (e.g., deep learning), those XAI models are explainable and humans are able to understand and adjust them. This simplifies the integration of expert knowledge in the learning process.

Consequently, using the digital twin as a base for reasoning, processes can be adapted based on the information provided by the digital twin. For controlling the food process operation, the LRA-M loop known from SeAC systems' research of the field of artificial intelligence is used (see Figure 2). Those SeAC systems have two main properties that describe their functionality [15]. First, those systems learn models that capture knowledge about (i) the systems themselves (i.e., their hardware and software, including possible adaptation actions and runtime behavior) and (ii) their environment such as users and other systems, but also environmental parameters that might be relevant. In the case of food production, this can be temperature, humidity, conditions of the transportation, raw material quality, etc. Second, SeAC systems use the information of the models to reason (i.e., to predict, analyze, consider, or plan required adaptations), which enables them to act based on their knowledge and reasoning results. For example, this could be the analysis that some process steps do not provide the target performance, and hence, the system changes different parameters.

The LRA-M loop uses ongoing learning about the environment in combination with reasoning for the next actions of the system. For the ongoing learning process, the empirical observations are used. The learning process analyzes the observations, and the gained knowledge is stored using models. The knowledge from the models and the given goals is used by the reasoning process to determine the next actions that the system should take to achieve these goals. The generated models can be complemented by other models, which, e.g., describe biological, physical, or chemical relations that influence the food. These actions can affect the behavior of the system and have an impact on the environment as well. The LRA-M loop is adapted as we want to include knowledge provided by the previously introduced digital twin into the framework. Thereby, the knowledge provided by the digital twin is not only a simple knowledge database, but processed data, which are generated using critical event prediction or different machine learning approaches. The main goal is that the SeAC system provides recommendations to the user on how to react to or adjust the parameters autonomously.

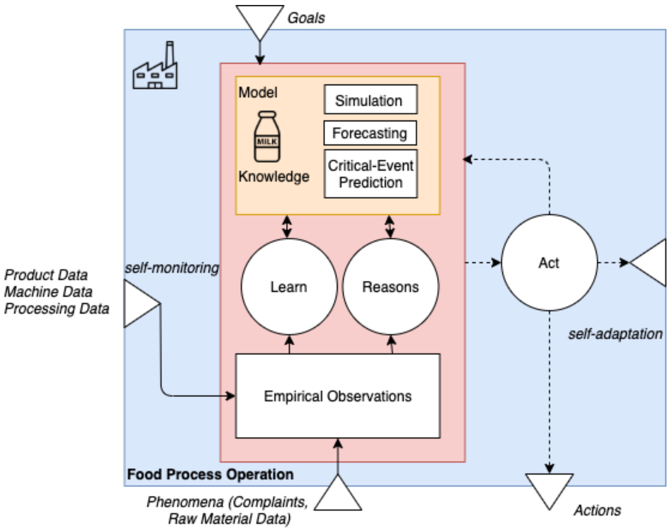

Figure 2. Conceptual framework on how to include data provided by a digital food twin into a self-aware learning and reasoning loop. Adapted from [15].

4. Discussion

Food production processes are particularly vulnerable, as the quality of raw materials varies depending on the season, and in addition, internal biological and chemical properties have to be taken into account. This information has to be included in the food process operation to secure a consistent high food quality and reduce food waste during production. Up to now, there is no food process operation that includes data provided by a digital twin as real-time input within an adaptive system to control the food processing. The concept of digital twins could improve this reasoning on how to adapt the process (e.g., machine parameters) based on the quality or properties of the raw material.

The digital twin concept could also support various functionalities of the food supply chain. Especially, the possibility to simulate various aspects and, through that, predict a critical situation in advance (e.g., cold chain violations) help to proactively react to and adapt the process. This work presents the underlying concept that shows how processed data (e.g., raw material, machine data, etc.) are used as the input for the manufacturing site to adapt production processes based on predicting critical situations.

Further, the digital twin can help to decrease the time to market for new products and support the scale-up of the production of new products. In theory, it would be possible to use the digital twin of a product with similar properties or a similar food matrix, adjust this digital twin, and use it as a base to learn the required adjustments in the product process (e.g., new configurations of machines) for hastening the scale-up of new products. Similarly, it is feasible to use the digital twin information for the determination of the potential shelf life of a new product based on the observations of similar products and the adjustments of a corresponding existing digital twin for the new product.

5. Conclusions

In this paper, we discussed the idea of using biophysical digital twins—composed of data from the process (collected by sensors), raw materials of the products, but also scientific models from food science—to capture and simulate the state of a food product and process during food processing. Such a digital twin would have several benefits; especially, it can be the base for reasoning on process adjustment and adaptations. This paper described the idea of integrating XAI procedures to improve the construction of the digital twins and integrating human knowledge—transferring the black box of machine learning to a gray box. Further, the paper described how SeAC systems can support adaptive food processing.

In our research group, we made the first steps towards our vision. Obviously, there are several challenges we still have to tackle. These include a general applicable model for describing the properties of the digital twin, which can be applied to different categories of food products. Further, we currently are building the digital twins manually. We are working on solutions that automate the construction of digital twins, as well as the analysis of the modeled food, similar to solutions from the area of machine learning, e.g., AutoML, or based on our previous works [6,7]. Additionally, we already have several parts for a system that can adapt the process from previous work and research projects—we are currently working on integrating and adjusting them for food processing.

Author Contributions: Conceptualization, C.K. and T.N.; methodology, C.K. and T.N.; formal analysis, C.K. and T.N.; writing—original draft preparation, T.N.; writing—review and editing, C.K.; visualization, T.N. and C.K.; supervision, C.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: I would like to thank my former student Tanja Noack for her support when preparing a previous version of this text.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IoT	Internet of Things
LRA-M	learn–reason–action-model
SeAC	self-aware computing
XAI	explainable artificial intelligence

References

1. Rohleder, B.; Minhoff, C. Die Ernährung 4.0—Status Quo, Chancen und Herausforderungen. 2019. Available online: https://www.bitkom.org/sites/default/files/2019-03/Bitkom-Charts%20190326%20Digitalisierung%20der%20Ern%C3%A4hrungsindustrie.pdf (accessed on 25 November 2021).
2. Krupitzer, C.; Roth, F.M.; VanSyckel, S.; Schiele, G.; Becker, C. A Survey on Engineering Approaches for Self-adaptive Systems. *Pervasive Mob. Comput.* **2015**, *17*, 184–206. [CrossRef]
3. Cadavid, J.P.U.; Lamouri, S.; Grabot, B.; Pellerin, R.; Fortin, A. Machine learning applied in production planning and control: A state-of-the-art in the era of industry 4.0. *J. Intell. Manuf.* **2020**, *31*, 1531–1558. [CrossRef]
4. Cioffi, R.; Travaglioni, M.; Piscitelli, G.; Petrillo, A.; De Felice, F. Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability* **2020**, *12*, 492. [CrossRef]
5. Krupitzer, C.; Wagenhals, T.; Züfle, M.; Lesch, V.; Schäfer, D.; Mozaffarin, A.; Edinger, J.; Becker, C.; Kounev, S. A survey on predictive maintenance for industry 4.0. *arXiv* **2020**, arXiv:2002.08224.
6. Züfle, M.; Moog, F.; Lesch, V.; Krupitzer, C.; Kounev, S. A Machine Learning-based Workflow for Automatic Detection of Anomalies in Machine Tools. *ISA Trans. J. Autom.* **2021**, *in press*. [CrossRef] [PubMed]

7. Zuefle, M.; Bauer, A.; Lesch, V.; Krupitzer, C.; Herbst, N.; Kounev, S.; Curtef, V. Autonomic Forecasting Method Selection: Examination and Ways Ahead. In Proceedings of the 2019 IEEE International Conference on Autonomic Computing (ICAC), Umea, Sweden, 16–20 June 2019; pp. 167–176. [CrossRef]
8. Verdouw, C.; Tekinerdogan, B.; Beulens, A.; Wolfert, S. Digital twins in smart farming. *Agric. Syst.* **2021**, *189*, 103046. [CrossRef]
9. Henrichs, E.; Noack, T.; Pinzon Piedrahita, A.M.; Salem, M.A.; Stolz, J.; Krupitzer, C. Can a Byte Improve Our Bite? An Analysis of Digital Twins in the Food Industry. *Sensors* **2022**, *22*, 115. [CrossRef] [PubMed]
10. Defraeye, T.; Shrivastava, C.; Berry, T.; Verboven, P.; Onwude, D.; Schudel, S.; Bühlmann, A.; Cronje, P.; Rossi, R.M. Digital twins are coming: Will we need them in supply chains of fresh horticultural produce? *Trends Food Sci. Technol.* **2021**, *109*, 245–258. [CrossRef]
11. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2405–2415. [CrossRef]
12. Rosen, R.; Von Wichert, G.; Lo, G.; Bettenhausen, K.D. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine* **2015**, *48*, 567–572. [CrossRef]
13. Verboven, P.; Defraeye, T.; Datta, A.K.; Nicolai, B. Digital twins of food process operations: the next step for food process models? *Curr. Opin. Food Sci.* **2020**, *35*, 79–87. [CrossRef]
14. Müller, P.; Schmid, M. Intelligent Packaging in the Food Sector: A Brief Overview. *Foods* **2019**, *8*, 16. [CrossRef] [PubMed]
15. Kounev, S.; Kephart, J.; Milenkoski, A.; Zhu, X. *Self-Aware Computing Systems*; Springer: Cham, Switzerland, 2017.