Review

Digital Twins and Industry 4.0 Technologies for Agricultural Greenhouses

Naftali Slob [ID] and William Hurst *[ID]

Information Technology Group, Wageningen University and Research, Leeuwenborch, Hollandseweg 1, 6706 KN Wageningen, The Netherlands
* Correspondence: will.hurst@wur.nl

Abstract: Digital twin (DT) is one of the key concepts of the fourth industrial revolution (Industry 4.0). A DT is a data-driven, digital replica of a real-world object or environment, including all the states and behaviour of the real-life counterpart. Based on research across multiple domains, DTs show great potential to improve both efficiency and sustainability. In addition, a DT is able to remove key constraints of human observations and interactions. Yet, the technology is still in its infancy. Thus, this article presents a literature search across five different databases focusing on suitable approaches for data coupling and the stages of DT implementation integration with Industry 4.0 technologies. The review process follows the systematic literature review (SLR) methodology. The selected articles cover a wide range of DT implementations across the agricultural industry, ranging from arable farming to aquaponics. Findings include identification of the stages needed to create the DTs, the data coupling processes, and how Industry 4.0 technologies (e.g., cloud-based technologies, IoT, and AI subfields) are integrated. Findings presented in this article will support others in the design of DTs and integration of Industry 4.0 technologies for agricultural greenhouses.

Keywords: digital twin; greenhouse; agriculture; smart farming; internet of things

1. Introduction

The world population is growing; estimations show that by 2050, there will be more than 9 billion people living on the Earth [1]. Moreover, due to urbanisation, 66% of the world population will live in cities. This will lead to numerous challenges concerning housing and food production. As the population grows, but the available area does not, agricultural practice needs to be able to produce more food with fewer resources. Thus, the need for efficient food-production systems is apparent. One possible method of gaining more efficiency lies in precision agriculture. The aim of precision agriculture is to use temporal and spatial information of field or animal produce in order to improve the productivity of the farm [2]. A growing area of technology to achieve this is through virtualisation of all the objects on a farm. For the virtualisation, a digital twin can be used [3]. Digital twins (DT) offer a novel way of representing a physical object or system in a digital environment. A DT is a data-driven, digital replica of a real-world object or environment, which can be used for decision support and systems analysis [4]. Once a DT is created, the application can be linked to the real object/environment using the Internet of Things (IoT). With this link, a DT is able to replicate the states and behaviour of the physical object in a digital environment [5]. This removes key constraints about human observations (time and place) since the observation can be done in the digital environment [3,6]. For instance, these observations can be achieved by means of augmented or mixed reality, as outlined in our previous work in [7]. The human collaborative process can be enhanced when using additional technologies such as data analytics, machine learning, artificial intelligence, and automation. This could lead to a learning and adapting DT, which can make choices based on simulated scenarios and data from the past [8] (with the main

difference between simulations and DTs being the bi-directional communication). The DT and the physical object communicate during their life span, for instance, updating values based on actual measurements [9].

DTs are currently employed across a wide variety of disciplines for diverse applications. In the manufacturing industry, for example, DTs are used for production planning and design, maintenance, product lifecycle, manufacturing, layout planning, and process design [10], whereas in the energy sector, the approach is used to forecast the life expectancy of wind turbines [11]. Further, DTs are readily used to evaluate the ergonomics of the workplace, detecting the welding completeness and automated cybersecurity testing in the automotive industry [7,12–14], and in the healthcare domain, as detailed by Du, et al. [15], who implemented DTs resulting in an increased accuracy, recall, and F1 score of stem cell detections.

In addition to outlining four categories of DT, Verdouw, et al. [3] identified five main classes for use cases in agriculture, including energy consumption analysis, system failure analysis and prediction, real-time monitoring, optimization/update, and a technology integration tool. However, ref. [16] only identified two case study examples that described the application of DTs in greenhouses.

A greenhouse, according to Stanghellini, et al. [17], is a permanent cover structure for crops, is tall enough to enter, and has some means of controlling the environment. Greenhouses are used to protect crops from the outside environment to ensure good crop yield even if conditions are harsh. More advanced greenhouses (high-tech) are equipped with computer-controlled environmental control, consisting of a network of actuators and sensors. When placed in a temperate climate (e.g., the Netherlands) these high-tech greenhouses consume a great deal of energy to control the indoor climate [18]. In order to save energy, a wide variety of equipment is used, i.e., heat exchangers, aquifers, and windows for natural ventilation. During daily operation, decisions need to be made about the use of this equipment [19]. As previously highlighted, DTs are used to perform energy analysis and are able to assist in the decision-making process. Therefore, it is possible to infer that these two applications of DTs seem to be ideal for greenhouses.

Based on the aforementioned discussion, this research aims to provide an overview of what technologies are used for the creation of an agricultural DT for greenhouses. There are other works focusing on this topic; for example, Ariesen-Verschuur, et al. [20] provided a comprehensive review of DT applications in greenhouse horticulture. Their approach focused on applications of DT technology. However, the research in this article, instead presents findings on suitable approaches for data coupling, the stages of DT implementation, and Industry 4.0 technologies. To achieve this, a literature search on DTs was conducted focusing on DT implementation and Industry 4.0 technology. As such, the following research questions were defined to gain a clear overview of the state-of-the-art DT application in agriculture. (1) *What data are typically coupled to agricultural digital twins?* (2) *What are the stages of implementation for creating a digital twin in scientific literature?* (3) *What Industry 4.0 technologies could be implemented in a greenhouse DT?* (3a) *What cloud-based technologies are implemented?* (3b) *What AI subfields are used alongside a DT?*

The remainder of this article is as follows. Section 2 provides a background discussion on DTs and their categorization. Section 3 outlines the methodology employed for the literature search, and Section 4 presents the evaluation of related articles. The article is concluded in Section 5.

2. Background

Based on the investigation by Verdouw, et al. [3], six categories of DTs were identified, namely (i) monitoring, (ii) predictive, (iii) prescriptive, (iv) autonomous, (v) imaginary, and (vi) recollection. Monitoring DTs visualise the current and historic behaviour and state of the object or environment that it replicates. The digital representation is based on sensor data and a meta model. Typically, these sensor data are enriched with external data to further enhance the virtual representation. Predictive DTs go a step further: instead of

showing the current or the past state, it shows the (possible) future state. Similar to the monitoring DTs, current and historic sensor data are employed. However, these data serve a prediction model to generate the future state of the real-world object or environment. A prescriptive DT shows the effect of certain interventions on the future DT, and these interventions are conducted in the present DT. An autonomous DT operates independently and does not need any intervention by humans; it has full control over the real-life object. An imaginary DT describes an object that does not yet exist in real-life, and this includes information needed to create the object. Recollection DTs store all the historical data of the physical object; therefore, this type of DT is also referred to as the digital memory. These historical data can be used to improve the next generation of objects [21].

2.1. DT Interventions

Additionally, there are typically two types of interventions: (i) reactive and (ii) proactive [3]. Reactive interventions address a current problem found within monitoring DT solutions; an example could be that the current temperature in a greenhouse is too high. A proactive intervention is based on problems identified by a forecast generated by the predictive DT. For instance, the predictive DT predicts that the temperature in the greenhouse will be too high in three hours.

The prescriptive DT provides information and insight into the different interventions and what the effects are, yet the final decision is still taken by a human operator, and the intervention is executed in the physical world. This is not the case for the autonomous DTs, which have full control over the object without human interventions. The outcome of the prescriptive DT (intervention) is remotely implemented using the actuators of the real-life object or environment.

Moreover, autonomous DTs can be enriched with self-learning, where embedded algorithms evaluate the results of the suggested interventions [22]. The complex interactions between the different DT types and a greenhouse are visualised in Figure 1. Within the diagram, the coloured nodes (yellow, red, green, blue) are specific to the different DTs, and the black nodes are common across all DT types. The diagram is inspired by articles by Verdouw et al. [3,16].

2.2. DT and Industry 4.0

Both according to industry and academia, DTs are finding a footing as one of the key technologies within the fourth industrial revolution (also known as Industry 4.0) [23]. For example, Hofmann and Rüsch [24] identified the following Industry 4.0 concepts based on their systematic literature review (SLR): cyber-physical systems (CPS), Internet of things (IoT), internet of services (IoS), and smart factory. CPS are systems that connect the virtual and physical world together on a network base; some DTs, therefore, could be considered as CPS. IoT is a mainstay technology within the Industry 4.0 domain. One of the more general definitions of IoT is an environment in which objects, using (relatively) small devices that are connected to the internet, can be turned into smart things [25]. This also leads to the concept of IoS, which refers to the service society that we live in these days and making these services more readily accessible through the internet. Smart factory combines the power of the previously mentioned concepts (CPS, IoT, and IoS). A smart factory is based upon decentralised production where human, resource, and machine communicate seamlessly.

Figure 1. Network graph of 4 DT stages, including monitoring, predictive, prescriptive, and autonomous.

At the time of conducting this investigation, no literature review articles were found that focussed on the technologies, data, or the development of (specifically) greenhouse DTs. However, Pylianidis, et al. [4] conducted a literature review on the application of DTs in agriculture focusing on the potential of the DT in agriculture. The authors concluded that most DTs were still in their primary stage, and therefore, the DTs did not yet offer the same benefits as found in other domains. Furthermore, they suggested starting implementing DTs with simpler functionalities and gradually adding more complex parts. Verdouw, et al. [3] analysed what DTs can contribute to the advancement of smart farming. They defined the concept and introduced a typology of DTs based on literature. In addition, they proposed a conceptual framework for the implementation and design of DTs. This framework was validated and applied in a case study. Based on the study, they defined a DT as a dynamic representation of a real-life object that mirrors its states and behaviour across its lifecycle. This application can be used to monitor, analyse, and simulate current and future states of and interventions. Further, this can involve the use of data integration, artificial intelligence, and machine learning.

Furthermore, Sreedevi and Santosh Kumar [26] found a variety of applications of DTs in agriculture based on ten articles. Based on the knowledge they gained through the review, they discussed how the application of similar techniques was to be applied in a case study on aquaponics.

Following this discussion, in the following sub-section, a literature review of DTs in the agricultural domain was provided first by selecting databases; then, defining the search strategy, defining the search string, and defining the selection criteria were performed before the actual search was started. Collecting and selecting papers, quality assessment, extracting data, and synthesizing data describe the steps performed during the search for papers.

3. Materials and Methods

The search for papers for this literature study were conducted over five databases including *Scopus* (SC), *Web of Science* (WoS), *Springer Link* (SL), *ScienceDirect* (SD), and *IEEE Xplore* (IEEE). The use of this selection of databases provided a comprehensive coverage of the field, with a focus on journals rather than conference papers. Further, the databases used were selected based on their relevance to agriculture and the informatics domain.

3.1. Search Protocol

Based on the knowledge gained during the background and related work section, it became clear that there was a limited number of papers regarding agricultural DTs. Therefore, to start with, a refined search string was first defined as in (1).

$$(\text{``digital twin'' AND agriculture}). \tag{1}$$

Based on the first retrieved papers, this was later extended to (2).

$$(\text{``digital twin *'' OR digital-twin *) AND (agriculture * OR farm *)} \tag{2}$$

This search string was the basis for all the databases; however, each database has a specific search string. The selected articles from the database search were used to perform a snowballing review process. During the process of snowballing, the reference list of the selected article was reviewed for additional articles. The articles found through this process were also checked using the selection criteria and quality assessment.

The databases used have different search conventions, and therefore, each database had a specific search string and specific settings, defined in Table 1.

Table 1. Search protocol.

Database	Search String	Specific Search Settings
Scopus	("digital twin *" OR digital-twin *) AND (agriculture * OR farm *)	Done within: Article title, Abstract, and Keywords. Document type was set to article or conference paper
Web of Science	("digital twin *" OR digital-twin *) AND (agriculture * OR farm *)	Was performed in the topic (title, abstract, and keywords).
Springer Link	"digital twin" AND agriculture AND farm	Content type was set to be either an article or conference paper
Science Direct	("digital twin" OR "digital-twin") AND (agricultur OR farm)	Was performed in the search field: "Find articles with these terms". The article type was set to research articles.
IEEE Xplore	("All Metadata": "digital twin *" OR "All Metadata": digital-twin *) AND ("All Metadata": agriculture * OR "All Metadata": farm*)	Metadata include the abstract, title text, and the indexing words.

3.2. Selection Criteria

All the found papers were checked based upon pre-defined selection criteria to ensure only useable papers are selected. If one of the exclusion criteria (Table 2) was true, the paper was excluded from this study.

Articles were collected from the described databased using the search strings and by performing snowballing. The list of selected articles was gathered in an excel file, and this file was later used to keep track of the quality assessment and the data extraction. For each individual article, the following information was gathered, presented in Table 3.

Table 2. Exclusion Criteria.

Exclusion If:
The publication is not about the application of a DT in agriculture.
The publication is not written in English.
The publication was published before 2017.
Full text is not available.
The publication is not an article (e.g., survey).
The publication is a duplicate study.

Table 3. Data Extracted.

Data per Article
Authors
Title
Year of publication
Database from which article was retrieved
Data that was used to create DT
The steps taken to create the DT
Cloud-based technologies that were used/linked to the DT
What AI subfields were implemented (what algorithms, if specified)

The articles found through either searching the databases with a search string or through snowballing were checked against the previously defined selection criteria. Only the selected articles were used in the following steps of the review.

3.3. Quality Assessment

Next, a quality assessment was performed on the selected papers to ensure their quality. As proposed in the study of Kitchenham, et al. [27], eight questions were used to assess the quality and suitability of the articles for analysis, outlined in Table 4.

Table 4. Quality Assessment.

The Eight Quality Assessment Questions
Are the aims of the study clearly stated?
Are the scope and experimental design of the study defined clearly?
Are the variables in the study likely to be valid and reliable?
Is the research process documented adequately?
Are all the study questions answered?
Are the negative findings presented?
Are the main findings regarding creditability, validity, and reliability stated?
Do the conclusions relate to the aim of the purpose of the study?

Articles were scored based on the answer to each question, where 1 point (yes), 0 points (no), or 0.5 points (somewhat) were allocated after review of the article. The highest achievable score is 8; if an article had less than 4 points (50%), it was excluded from the study. The awarded score is not necessarily a complete reflection of the quality but rather the suitability for this investigation.

3.4. Data Extraction

Relevant data from the selected papers were extracted in order to answer the research questions. This was achieved by reading all papers in full and writing down the useful information. The data were gathered in a spreadsheet file to obtain an overview of the whole study.

3.5. Data Synthesis

The data collected during the extraction were diverse and needed to be synthesized to gain a holistic overview. This means that, per research question, umbrella terms were used to categorise the different answers of different articles. The defined umbrella terms per research question can be found in Table 5 below.

Table 5. Research Question and Umbrella Terms.

Research Question	Description
1	Outdoor climate, indoor climate, soil, plant/crop, animal, water, images, machines, other
2	Textual, mathematical/physics, agent network, digital model, 3D model, linking data, simulation, cloud interaction
3, 3A, and 3B	Database, interface, API, communication, computation, computer vision, machine learning, deep learning, neural network

4. Results

After checking the selection criteria and applying the quality assessment (Figure 2), 22 articles were selected and used in the literature study (see Table 6). The findings are discussed in the following sub-sections, where the research questions are employed as headers.
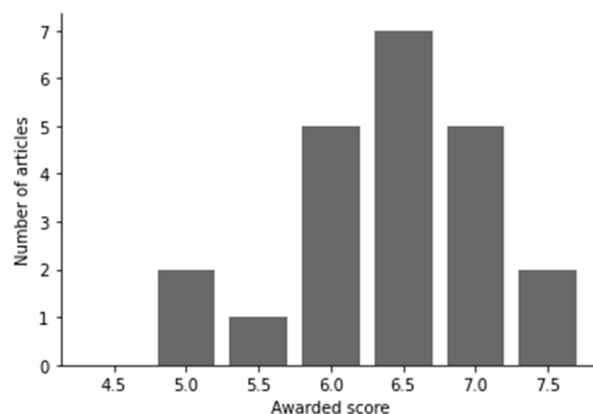
Figure 2. Quality score distribution of the selected articles.

Table 6. Selected Articles for Review.

Citation	Title
Pylianidis, et al. [4]	Introducing digital twins to agriculture
Verdouw, et al. [3]	Digital twins in smart farming
Laryukhin, et al. [28]	The multi-agent approach for developing a cyber-physical system for managing precise farms with digital twins of plants
Moshrefzadeh, et al. [29]	Towards a distributed digital twin of the agricultural landscape
Chaux, et al. [9]	A digital twin architecture to optimize productivity within controlled environment agriculture
Jans-Singh, et al. [30]	Digital twin of an urban-integrated hydroponic farm
Angin, et al. [31]	Agrilora: A digital twin framework for smart agriculture
Tekinerdogan and Verdouw [8]	Systems Architecture Design Pattern Catalog for Developing Digital Twins
Ghandar, et al. [32]	A Decision Support System for Urban Agriculture Using Digital Twin: A Case Study with Aquaponics

Table 6. *Cont*.

Citation	Title
Machl, et al. [33]	Planning Agricultural Core Road Networks Based on a Digital Twin of the Cultivated Landscape
Kampker, et al. [34]	Business Models for Industrial Smart Services—The Example of a Digital Twin for a Product-Service-System for Potato Harvesting
Monteiro, et al. [35]	Towards Sustainable Digital Twins for Vertical Farming
Jo, et al. [36]	Smart Livestock Farms Using Digital Twin: Feasibility Study
Alves, et al. [37]	A digital twin for smart farming
Ahmed, et al. [38]	Digital Twin Technology for Aquaponics: Towards Optimizing Food Production with Dynamic Data Driven Application Systems
Skobelev, et al. [39]	Development of Digital Twin of Plant for Adaptive Calculation of Development Stage Duration and Forecasting Crop Yield in a Cyber-Physical System for Managing Precision Farming
Foldager, et al. [40]	Towards a Digital Twin—Modelling an Agricultural Vehicle
Skobelev, et al. [41]	Development of models and methods for creating a digital twin of plant within the cyber-physical system for precision farming management
Skobelev, et al. [42]	Multi-agent approach for developing a digital twin of wheat
Neethirajan and Kemp [43]	Digital Twins in Livestock Farming
Jo, et al. [44]	Energy Planning of Pigsty Using Digital Twin
Howard, et al. [45]	Data Architecture for Digital Twin of Commercial Greenhouse Production

4.1. What Data Is Typically Coupled to Agricultural Digital Twins?

In total, the articles used data regarding 52 different parameters. These different parameters were categorized under the aforementioned nine umbrella terms to gain a better overview. Figure 3 displays the data groups used per article; if at least one parameter of a certain group is used, the square is black. In addition, this Figure 3 visualises the number of times groups are used together in one article. The single most used group is "outdoor climate", which is used by twelve articles [3,8,9,28–31,39,41,42,44,45].
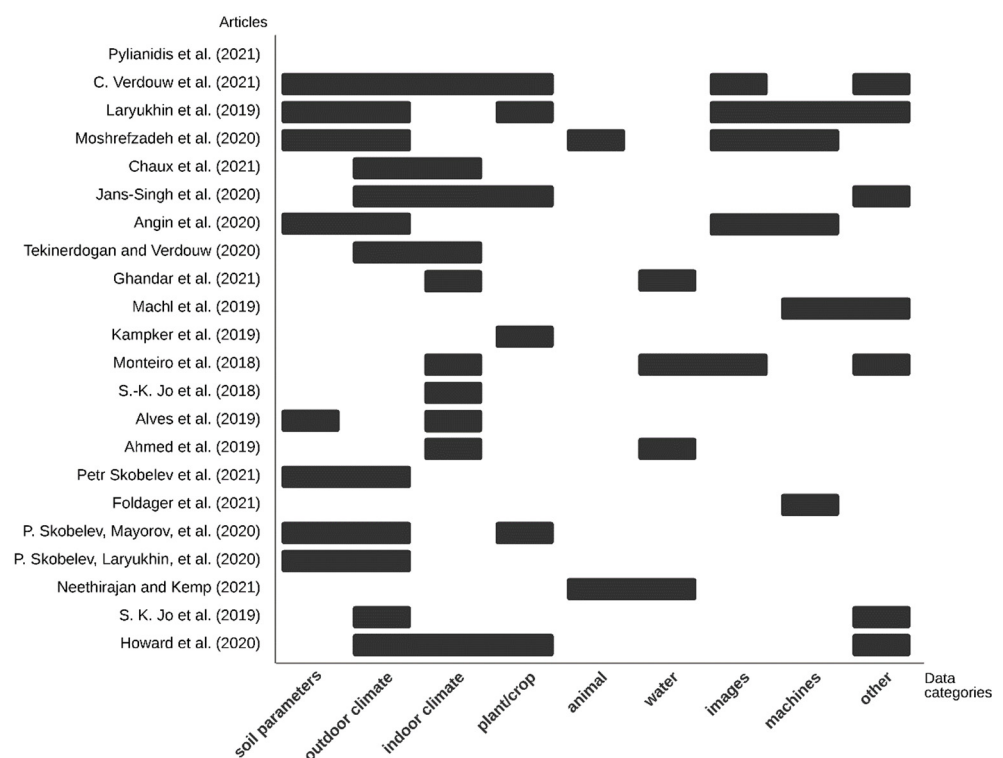

Figure 3. Data types per article [3,4,8,9,28–45].

For example, Skobelev, et al. [42] used temperature, humidity, and precipitation data in their wheat growth model. The least-used category is "animal"; this was only used by two articles. Neethirajan and Kemp [43] suggested using animal body temperature when creating a DT of a livestock farm, and Moshrefzadeh, et al. [29] used animal GPS data in their design of a distributed digital twin. The most co-occurring groups are "soil parameters" and "outdoor climate"; seven articles use these two groups together [3,28,29,31,39,41,42]. Articles contained an average of 5.2 different types of data, belonging to an average 2.6 groups.

4.2. What Are the Stages of Implementation for Creating a Digital Twin in Scientific Literature?

The stages were grouped under the aforementioned different umbrella terms. Articles that used the textual stage explicitly stated that the initialisation process involved a written model. For example, Monteiro, et al. [35] specifically stated they used a goal-orientated requirements language model. The articles that used mathematics/physics used equations defining the interaction of the model. The digital model stage is where software (often not stated) was used to conduct the calculation or visualisation process. The linking data stage involves gathering data and linking to the model.

The identified stages and their links are visualised using a chord diagram in Figure 4. As can be seen, the groups "textual and mathematics/physics" are often the starting stage, with no arcs going ending at these nodes. Simulation and cloud interaction were often the end of the research, with no arc leaving these nodes. The stage that was most often used is the link data stage. This is a logical consideration based on the fact that most of the DTs are data-based.
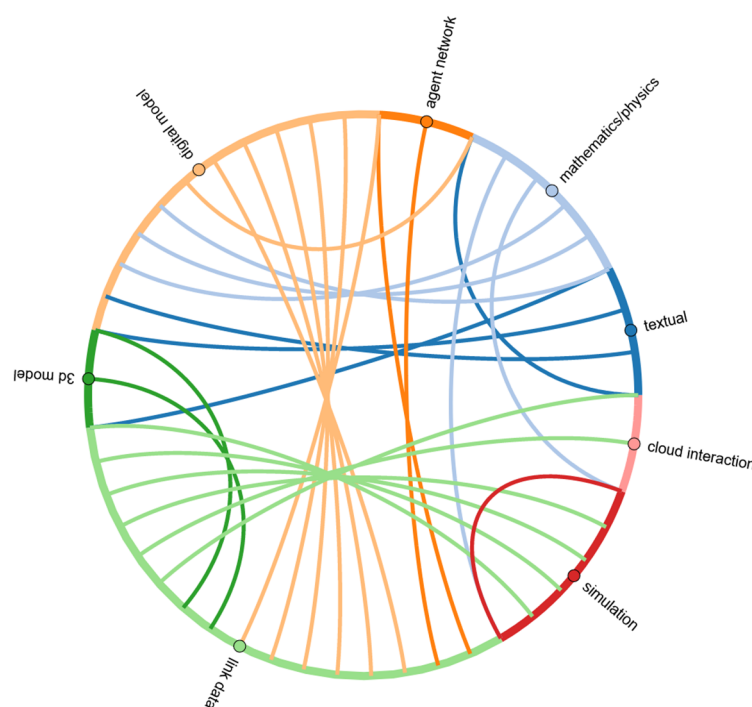
Figure 4. Chord diagram showing the different stages and the links between.

4.3. What Industry 4.0 Technologies Could Be Implemented in a Greenhouse DT?

4.3.1. What Cloud-Based Technologies Are Implemented?

Out of the 22 articles, 3 did not specify if they used any cloud-based technologies [4,8,36], and the other 19 articles used at least one type of cloud-based tech. Some articles also used multiple cloud-based technologies, and this is why the sum of the number of articles in Table 7 exceeds the total number of articles (22). An example of a DT that uses multiple cloud-based technologies can be found in the article of Alves, et al. [37]. The

authors developed a framework for a DT of an agricultural field. Since the field would be at a remote location, communication passes through the cloud, data would be directly stored in the online database, and the online interface gives the user a clear overview of the gathered data.

Table 7. Number and percentage of papers using certain cloud-based technology.

Cloud-Based Technologies	# of Articles	Percentage of Articles
Database	15	68%
API	8	36%
Interface	6	27%
Communication	4	18%
Computation	2	9%

4.3.2. What AI Subfields Are Used Alongside a DT?

Just over half of the selected articles (12) specified the use of an AI subfield in their research. Table 8 shows the number of articles that used a specified subfield, and three used multiple subfields. A combination of deep learning, computer vision, and neural networks was used to create a DTs framework for precision agriculture. This was based on IOT and real-time cloud-based data integration [31]. Tekinerdogan and Verdouw [8], for example, defined multiple architecture design patterns, and the digital matching pattern specified that machine learning or deep learning is typically used for object recognition or classification. In their feasibility study of a DT of livestock farm, Jo, et al. [44] used machine and deep learning for their modelling and analysis.

Table 8. AI Subfields in Selected Articles.

AI Subfield	# of Articles
Machine learning	10
Deep learning	3
Neural network	2
Computer vision	2

4.4. Discussion

This literature study showed that there is a wide variety of applications of digital twins across the agriculture domain. Articles focussed on the livestock farming, using DTs to reduce energy cost [44], monitoring feed uptake of animals [43], or even improving their living condition [36]. In their article, Laryukhin et al. [28] described the cyber-physical multi-agent system (CP MAS) architecture. A similar approach was also adopted to incorporate into a DT to predict the growth, based on stages, of wheat plants in a field in articles [26,39,41]. These articles described gathering data in the field, which was also described in detail by Alves et al. [37].

Further, multiple articles discussed the application of DTs in aquaponics. Ref. [38] defined a cyber-physical system for aquaponics based on IoT. They used a virtual system to validate and implement their system, whereas [32] used a real-life aquarium to implement their proposed DT as prove of their concept. Something similar was done for a greenhouse by [9]: they applied their DT on a small, desktop-sized prototype of a greenhouse.

Even though all the articles describe applications of DTs, not all stated the needed information to answer all the research questions. When an article did not contain the information needed to answer a specific research question, no answer was filled in. Especially for research question 2, this needs to be considered, as often, articles were not implicitly clear on their process of developing their DT. It is likely that the approaches involved more steps not included in the manuscripts.

However, a different situation occurred whilst gathering data for research question 1. Laryukhin, et al. [28] stated they used notes from agronomists in the field. These notes

included data that belonged to multiple groups but also included additional data that were not specified. This additional data could have been about weather or animals, in which case it would belong to additional groups.

As can be seen in Table 8, ten articles employed machine learning techniques. For example, Jans-Singh, et al. [30] used a random forest algorithm to impute missing temperature data based on outside weather data, the hour of the day, and neighbouring sensor data. Ghandar, et al. [32] tested three different types of algorithms (linear regression, support vector regression, and CART decision trees) to predict weekly plant growth and daily fish growth. However, 8/10 articles that stated using machine learning did not state what algorithm they used. For example, Jo, et al. [36] stated using machine learning for analyses and modelling of various kinds, yet it is not clear what exactly they used it for or what algorithm they employed.

When considering the development of a greenhouse DT, the article by Chaux, et al. [9] is a suitable starting point. The authors defined an architecture for an DT of a controlled environment (greenhouse) and tested it on a small prototype. For their DT, they used outdoor climate data (like the majority of articles) and indoor climate data. In addition to these two groups, Howard, et al. [45] also used plant data in their designed architecture for a commercial greenhouse. Depending on if the plants in the greenhouse grow in soil or not, one might consider using soil data as well.

Alves, et al. [37] showed how this could be done in the field, and the same concepts can be applied in a greenhouse. From the results in Section 4, the common stages of implementation are clear. An approach should either start with equations or with a textual description, but this depends on intended implementation and preference. From there, one would either create a digital or 3D model and link the selected data to this model. This digital model could also be an agent-based model. Cloud interaction could be implemented if wanted/needed, and the use of a database is common (68% of articles). The simulation aspect of the DT can be added, often in combination with machine learning.

5. Conclusions and Future Work

The database search is the biggest threat to validity of this literature study. The five used databases were selected based on their relevance to agriculture and the informatics domain. To the best of the authors' knowledge, these databases cover the vast amount of the published articles. However, there could be other literature types that were not included in this review because a certain database was excluded. Moreover, the defined search strings could have missed relevant papers. By performing a broad search and looking for synonyms before defining the final search strings, the authors limited the potential number of missing papers. In addition, snowballing was performed on the selected papers to find papers that were not directly retrieved using the search string. By following a predefined and widely used procedure, the SLR [27], this research is reproduceable when people are in doubt of the results or want to perform a similar study.

This article gained an overview of the state-of-the-art DT application in agriculture by performing a literature study. The search query across five different databases resulted in 22 articles used for data extraction. A wide variety of DT application in agriculture was found, ranging from predicting wheat growth [42] to controlling an aquaponics system [32]. Consequently, a wide variety of data types were used, and eight umbrella terms were defined: outdoor climate, indoor climate, soil, plant/crop, animal, water, images, machines, and other. The DTs used outdoor climate data the most (12 articles); this was often used in combination with soil data (7 articles). The DTs used, on average, 5.2 different data types, belonging to, on average, 2.6 categories.

The stages of implementation of the DTs were diverse. However, it became apparent that starting with equations or a textual description of the DT is common. Cloud interaction or simulation were often the final step of the DTs, and the most used step was linking data.

Out of the 22 selected articles, 3 did not implement any cloud-based technology, while the other 19 did. Out of the 22 articles, 68% of them used a database for their DT, and 36%

used an API. Ten articles used machine learning in their DT; however, eight out of these ten did not specify what algorithm.

Based on all the collected information, we proposed a design for a DT of a greenhouse. Further research should focus on applying and testing the proposed design. In addition, the documentation of the development process and design choices of DTs should be improved, as this ensures that people can recreate and benefit from the research.

References

1. McGranahan, G.; Satterthwaite, D. Urbanisation Concepts and Trends. 2014. Available online: http://VH7QX3XE2P.search.serialssolutions.com/?V=1.0&L=VH7QX3XE2P&S=AC_T_B&C=Urbanisation%20concepts%20and%20trends&T=marc&tab=BOOKS (accessed on 29 October 2021).
2. Balafoutis, A.T.; Beck, B.; Fountas, S.; Tsiropoulos, Z.; Vangeyte, J.; Wal, T.V.D.; Soto-Embodas, I.; Gómez-Barbero, M.; Pedersen, S.M. Smart Farming Technologies—Description, Taxonomy and Economic Impact. In *Precision Agriculture: Technology and Economic Perspectives*; Pedersen, S.M., Lind, K.M., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 21–77.
3. Verdouw, C.; Tekinerdogan, B.; Beulens, A.; Wolfert, S. Digital twins in smart farming. *Agric. Syst.* **2021**, *189*, 103046. [CrossRef]
4. Pylianidis, C.; Osinga, S.; Athanasiadis, I.N. Introducing digital twins to agriculture. *Comput. Electron. Agric.* **2021**, *184*, 105942. [CrossRef]
5. Boschert, S.; Rosen, R. Digital Twin—The Simulation Aspect. In *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and their Designers*; Hehenberger, P., Bradley, D., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 59–74.
6. Verdouw, C.N.; Wolfert, J.; Beulens, A.J.M.; Rialland, A. Virtualization of food supply chains with the internet of things. *J. Food Eng.* **2016**, *176*, 128–136. [CrossRef]
7. Hurst, W.; Mendoza, F.R.; Tekinerdogan, B. Augmented Reality in Precision Farming: Concepts and Applications. *Smart Cities* **2021**, *4*, 1454–1468. [CrossRef]
8. Tekinerdogan, B.; Verdouw, C. Systems Architecture Design Pattern Catalog for Developing Digital Twins. *Sensors* **2020**, *20*, 5103. [CrossRef]
9. Chaux, J.D.; Sanchez-Londono, D.; Barbieri, G. A digital twin architecture to optimize productivity within controlled environment agriculture. *Appl. Sci.* **2021**, *11*, 8875. [CrossRef]
10. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022. [CrossRef]
11. Sivalingam, K.; Sepulveda, M.; Spring, M.; Davies, P. A Review and Methodology Development for Remaining Useful Life Prediction of Offshore Fixed and Floating Wind turbine Power Converter with Digital Twin Technology Perspective. In Proceedings of the 2018 2nd International Conference on Green Energy and Applications (ICGEA), Singapore, 24–26 March 2018; pp. 197–204.
12. Caputo, F.; Greco, A.; Fera, M.; Macchiaroli, R. Digital twins to enhance the integration of ergonomics in the workplace design. *Int. J. Ind. Ergon.* **2019**, *71*, 20–31. [CrossRef]
13. Marksteiner, S.; Bronfman, S.; Wolf, M.; Lazebnik, E. Using Cyber Digital Twins for Automated Automotive Cybersecurity Testing. In Proceedings of the 2021 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Vienna, Austria, 6–10 September 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 123–128.
14. Li, H.; Li, B.; Liu, G.; Wen, X.; Wang, H.; Wang, X.; Zhang, S.; Zhai, Z.; Yang, W. A detection and configuration method for welding completeness in the automotive body-in-white panel based on digital twin. *Sci. Rep.* **2022**, *12*, 1–11. [CrossRef]
15. Du, X.; Liu, M.; Sun, Y. Segmentation, Detection, and Tracking of Stem Cell Image by Digital Twins and Lightweight Deep Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 1–10. [CrossRef]
16. Verdouw, C.; Beulens, A.J.M.; Reijers, H.A.; van der Vorst, J.G.A.J. A control model for object virtualization in supply chain management. *Comput. Ind.* **2015**, *68*, 116–131. [CrossRef]
17. Stanghellini, C.; Ooster, B.v.; Heuvelink, E. *Greenhouse Horticulture*; Wageningen Academic Publishers: Wageningen, The Netherlands, 2019; Volume 1, Chapter 1; pp. 14–25.

18. van den Bulck, N.; Coomans, M.; Wittemans, L.; Hanssens, J.; Steppe, K. Monitoring and energetic performance analysis of an innovative ventilation concept in a Belgian greenhouse. *Energy Build.* **2013**, *57*, 51–57. [CrossRef]

19. van Beveren, P.J.M.; Bontsema, J.; van Straten, G.; van Henten, E.J. Optimal utilization of a boiler, combined heat and power installation, and heat buffers in horticultural greenhouses. *Comput. Electron. Agric.* **2019**, *162*, 1035–1048. [CrossRef]

20. Ariesen-Verschuur, N.; Verdouw, C.; Tekinerdogan, B. Digital Twins in greenhouse horticulture: A review. *Comput. Electron. Agric.* **2022**, *199*, 107183. [CrossRef]

21. Grieves, M.; Vickers, J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 85–113.

22. Porter, M.E.; Heppelmann, J.E. How smart, connected products are transforming competition. *Harv. Bus. Rev.* **2014**, *92*, 64–88.

23. Huang, Z.; Shen, Y.; Li, J.; Fey, M.; Brecher, C. A Survey on AI-Driven Digital Twins in Industry 4.0: Smart Manufacturing and Advanced Robotics. *Sensors* **2021**, *21*, 6340. [CrossRef]

24. Hofmann, E.; Rüsch, M. Industry 4.0 and the current status as well as future prospects on logistics. *Comput. Ind.* **2017**, *89*, 23–34. [CrossRef]

25. Fleisch, E. What is the internet of things? An economic perspective. *Econ. Manag. Financ. Mark.* **2010**, *5*, 125–157.

26. Sreedevi, T.R.; Kumar, M.B.S. Digital Twin in Smart Farming: A Categorical Literature Review and Exploring Possibilities in Hydroponics. 2020, pp. 120–124. Available online: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85094606823&doi=10.1109%2fACCTHPA49271.2020.9213235&partnerID=40&md5=1ca0013861a012c80d6ee0b81ed963d6 (accessed on 17 December 2021). [CrossRef]

27. Kitchenham, B.; Charters, S. Guidelines for performing Systematic Literature Reviews in Software Engineering. Keele University, 9/07/2007. Available online: http://cdn.elsevier.com/promis_misc/525444systematicreviewsguide.pdf (accessed on 15 December 2021).

28. Laryukhin, V.; Skobelev, P.; Lakhin, O.; Grachev, S.; Yalovenko, V.; Yalovenko, O. The multi-agent approach for developing a cyber-physical system for managing precise farms with digital twins of plants. *Cybern. Phys.* **2019**, *8*, 257–261. [CrossRef]

29. Moshrefzadeh, M.; Machl, T.; Gackstetter, D.; Donaubauer, A.; Kolbe, T.H. Towards a distributed digital twin of the agricultural landscape. *J. Digital Landsc. Archit.* **2020**, *2020*, 173–186.

30. Jans-Singh, M.; Leeming, K.; Choudhary, R.; Girolami, M. Digital twin of an urban-integrated hydroponic farm. *Data-Centric Eng.* **2020**, *1*, e20. [CrossRef]

31. Angin, P.; Anisi, M.H.; Göksel, F.; Gürsoy, C.; Büyükgülcü, A. Agrilora: A digital twin framework for smart agriculture. *J. Wireless Mobile Netw. Ubiquitous Comput. Dependable Appl.* **2020**, *11*, 77–96.

32. Ghandar, A.; Ahmed, A.; Zulfiqar, S.; Hua, Z.; Hanai, M.; Theodoropoulos, G. A Decision Support System for Urban Agriculture Using Digital Twin: A Case Study With Aquaponics. *IEEE Access* **2021**, *9*, 35691–35708. [CrossRef]

33. Machl, T.; Donaubauer, A.; Kolbe, T.H. Planning agricultural core road networks based on a digital twin of the cultivated landscape. *J. Digital Landsc. Archit.* **2019**, *2019*, 316–327.

34. Kampker, A.; Stich, V.; Jussen, P.; Moser, B.; Kuntz, J. Business Models for Industrial Smart Services—The Example of a Digital Twin for a Product-Service-System for Potato Harvesting. *Procedia CIRP* **2019**, *83*, 534–540. [CrossRef]

35. Monteiro, J.; Barata, J.; Veloso, M.; Veloso, L.; Nunes, J. Towards Sustainable Digital Twins for Vertical Farming. In Proceedings of the 2018 Thirteenth International Conference on Digital Information Management (ICDIM), Berlin/Heidelberg, Germany, 24–26 September 2018; pp. 234–239.

36. Jo, S.-K.; Park, D.-H.; Park, H.; Kim, S.-H. Smart livestock farms using digital twin: Feasibility study. In Proceedings of the 2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, 17–19 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1461–1463.

37. Alves, R.G.; Souza, G.; Maia, R.F.; Tran, A.L.H.; Kamienski, C.; Soininen, J.P.; Aquino, P.T.; Lima, F. A digital twin for smart farming. In Proceedings of the 2019 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, USA, 17–20 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–4.

38. Ahmed, A.; Zulfiqar, S.; Ghandar, A.; Chen, Y.; Hanai, M.; Theodoropoulos, G. Digital Twin Technology for Aquaponics: Towards Optimizing Food Production with Dynamic Data Driven Application Systems. In *Methods and Applications for Modeling and Simulation of Complex Systems*; Springer: Singapore, 2019; pp. 3–14.

39. Skobelev, P.; Mayorov, I.; Simonova, E.; Goryanin, O.; Zhilyaev, A.; Tabachinskiy, A.; Yalovenko, V. Development of Digital Twin of Plant for Adaptive Calculation of Development Stage Duration and Forecasting Crop Yield in a Cyber-Physical System for Managing Precision Farming. In *Cyber-Physical Systems: Digital Technologies and Applications*; Kravets, A.G., Bolshakov, A.A., Shcherbakov, M.V., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 83–96.

40. Foldager, F.F.; Thule, C.; Balling, O.; Larsen, P. Towards a Digital Twin—Modelling an Agricultural Vehicle. In *Leveraging Applications of Formal Methods, Verification and Validation: Tools and Trends*; Springer International Publishing: Cham, Switzerland, 2021; pp. 109–123.

41. Skobelev, P.O.; Mayorov, I.V.; Simonova, E.V.; Goryanin, O.I.; Zhilyaev, A.A.; Tabachinskiy, A.S.; Yalovenko, V.V. Development of models and methods for creating a digital twin of plant within the cyber-physical system for precision farming management. *J. Physics: Conf. Ser.* **2020**, *1703*, 012022. [CrossRef]

42. Skobelev, P.; Laryukhin, V.; Lakhin, O.; Grachev, S.; Yalovenko, V.; Yalovenko, O. Multi-agent approach for developing a digital twin of wheat. In Proceedings of the 2020 IEEE International Conference on Smart Computing (SMARTCOMP), Bologna, Italy, 14–17 September 2020; pp. 268–273.
43. Neethirajan, S.; Kemp, B. Digital Twins in Livestock Farming. *Animals* **2021**, *11*, 1008. [CrossRef]
44. Jo, S.K.; Park, D.H.; Park, H.; Kwak, Y.; Kim, S.H. Energy Planning of Pigsty Using Digital Twin. In Proceedings of the 2019 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, 16–18 October 2019; pp. 723–725.
45. Howard, D.A.; Ma, Z.; Aaslyng, J.M.; Jørgensen, B.N. Data Architecture for Digital Twin of Commercial Greenhouse Production. In Proceedings of the 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), Ho Chi Minh City, Vietnam, 14–15 October 2020; pp. 1–7.