buildings

Review

Digital Twin for Fault Detection and Diagnosis of Building Operations: A Systematic Review

Faeze Hodavand [1], Issa J. Ramaji [2,*] and Naimeh Sadeghi [1]

1 School of Engineering, Construction and Management, Khajeh Nasir Toosi University of Technology, Tehran 19697 64499, Iran; f.hodavand@email.kntu.ac.ir (F.H.); sadeghi@kntu.ac.ir (N.S.)
2 School of Engineering, Construction and Computing, Roger Williams University, SE 108, One Old Ferry Road, Bristol, RI 02809, USA
* Correspondence: iramaj@rwu.edu; Tel.: +1-(814)-441-2495

Abstract: Intelligence in Industry 4.0 has led to the development of smart buildings with various control systems for data collection, efficient optimization, and fault detection and diagnosis (FDD). However, buildings, especially with regard to heating, ventilation, and air conditioning (HVAC) systems, are responsible for significant global energy consumption. Digital Twin (DT) technology offers a sustainable solution for facility management. This study comprehensively reviews DT performance evaluation in building life cycle and predictive maintenance. 200 relevant papers were selected using a systematic methodology from Scopus, Web of Science, and Google Scholar, and various FDD methods were reviewed to identify their advantages and limitations. In conclusion, data-driven methods are gaining popularity due to their ability to handle large amounts of data and improve accuracy, flexibility, and adaptability. Unsupervised and semi-supervised learning as data-driven methods are important for FDD in building operations, such as with HVAC systems, as they can handle unlabeled data and identify complex patterns and anomalies. Future studies should focus on developing interpretable models to understand how the models made their predictions. Hybrid methods that combine different approaches show promise as reliable methods for further research. Additionally, deep learning methods can analyze large and complex datasets, indicating a promising area for further investigation.

Keywords: Digital Twin (DT); fault detection and diagnosis (FDD); building operation; predictive maintenance (PDM); heating, ventilation, and air conditioning (HVAC) system; data-driven methods

Citation: Hodavand, F.; Ramaji, I.J.; Sadeghi, N. Digital Twin for Fault Detection and Diagnosis of Building Operations: A Systematic Review. *Buildings* **2023**, *13*, 1426. https://doi.org/10.3390/buildings13061426

Academic Editor: Nuno M. M. Ramos

Received: 11 April 2023
Revised: 16 May 2023
Accepted: 26 May 2023
Published: 31 May 2023

1. Introduction

The Fourth Industrial Revolution, also known as Industry 4.0, is transforming how we work and live. Energy consumption, environmental impact, cost, and occupant comfort have become focal points in this industry. The construction industry is a significant component of industrial improvement and accounts for approximately 8–10% of various countries' economies [1,2]. Additionally, studies have shown that people in the US and UK spend most of their time inside buildings [3,4]. One of the major challenges in modern cities is energy consumption. In Europe, buildings are responsible for 40% of total energy generation and approximately 28% of CO_2 emissions [5,6]. A significant portion of this energy is consumed by heating, ventilation, and air conditioning (HVAC) systems, which serve as the heart of any structure to maintain a comfortable indoor climate for occupants.

Operational faults in building facilities can cause an increase of 20% to 30% in total energy consumption [7]. Early detection and identification of faults can reduce this range, and effectively monitoring building components is critical to achieving this goal. Facility management teams require real-time, accurate, and comprehensive data to perform daily maintenance activities and provide accurate information [8]. However, facility inspection, maintenance assessment, and data collection activities are time-consuming, and the budget and resources devoted to building maintenance often need to be increased. As a result, maintenance personnel are dissatisfied with inadequate funding [9–12], and the quality of maintenance inspections can lead to poor maintenance and management of facilities. Building maintenance and management approaches need to be optimized to address these challenges.

Previous reviews have focused on classifying FDD methods and describing their characteristics. In addition, studies have explored the potential of Digital Twin (DT) technology in facilitating maintenance activities. However, there needs to be more research on how DT can aid the FDD process in buildings and which FDD method is most effective when integrated with DT technology.

Therefore, the research question that this paper aims to address is: How can DT aid the FDD process in buildings, and which FDD method is most effective when integrated with Digital Twin technology? To answer this question, the paper will comprehensively review FDD methods for building facilities, focus on using Digital Twin technology. It will classify and compare various FDD methods, analyze their suitability for use with DT technology, and provide recommendations for future research.

The remaining sections of this review are organized as follows. Section 2 introduces the background of maintenance activities and the usage of DT in facility management. Section 3 presents the search strategy used to identify relevant studies. Section 4 categorizes FDD methods and discusses their advantages and limitations. The results from Section 4 are identified and visualized in Section 5. Finally, Section 6 comprises a summary of the findings and concluding remarks from this systematic review.

2. Background

2.1. Maintenance

The operation and maintenance phase accounts for over 60% of the building life cycle, making it one of the most expensive activities in the construction industry [13–15]. Maintenance activities can generally be divided into three categories [16–18]:

1. Reactive maintenance is also known as run-to-failure or unplanned maintenance.
2. Preventive maintenance (PVM) involves a scheduled plan at specific times, but does not account for the system's condition or health. In some cases, routine inspection is more expensive than replacement.
3. Predictive maintenance is a data-based model that predicts a system's failure using statistical or machine learning models to improve decision-making processes and avoid downtime.

PDM focuses on scheduled maintenance activities and early fault detection and diagnosis. Early detection of faults before the system breaks down or becomes noticeable to home occupants can reduce maintenance costs and energy usage [19]. At the same time, facility managers can use historical data to improve maintenance plans and actions. Faults in building facilities refer to deviations from specified thresholds or predicted values with a certain accuracy that can cause systems to malfunction or result in process changes [20,21].

PDM uses building information modeling (BIM) as a visualization tool to monitor operational activities and enable facility managers to make accurate decisions to increase component service life. BIM has revolutionized our perspective on building function and information transmission [22]. According to the National BIM Standard—United States (NBIMS-US) standard [23], BIM creates a digital representation of physical and functional operation characteristics that can be considered a shared knowledge resource. BIM enables stakeholders to visualize and monitor construction activities [24,25] as a crucial component of information management to decrease risks (time, cost), conflict, and waste material over the building life cycle [26–28].

However, BIM as a virtual process and shared model cannot support real-time changes [29–31] and needs other resources to integrate data, a critical challenge of smart building management [32]. Moreover, managing data collected from BIM and other system devices, such as computerized maintenance management systems (CMMS), building automation systems (BAS), and building management systems (BMS), requires time and expert knowledge [33–36]. The adoption of DT technology is required to make BIM more compatible with both the micro (construction site) and macro (city districts) stages of construction [37].

Digital Twin technology enhances BIM's capabilities (3D information model) by incorporating real-time data from the Internet of Things (IoT) devices and cloud-based software to create an active replica of the entire facility, complete with entities, networks, and infrastructures [38,39]. The emergence of DT technology has provided a promising solution for tackling the challenges of complexity and integration in smart cities. The digital transformation has played a vital role in facilitating the adoption of Digital Twin technology in smart cities, along with other disruptive technologies, such as IoT, cloud computing, big data, cyber-physical systems, and AI [40]. By leveraging these technologies, smart cities can enhance their capacity to offer optimized services and products that address various societal challenges while prioritizing sustainability and intelligence [41,42].

The use of Digital Twin offers a new approach to facility management. Couprey et al. [14] identified four categories of challenges to BIM-based DT creation and usage. These include a need for more data standardization, a lack of which can hinder communication between parties and the prediction of the next phase. A lack of up-to-date data is a significant concern in this field [43–48]. In a review of DT advantages in different stages of smart construction [49], Boje et al. [34] categorized its abilities into physical, data, and virtual phases.

PDM is at the forefront of utilizing Digital Twin. By selecting an advanced model based on data acquisition, PDM can achieve highly accurate predictions and reduce maintenance costs by up to 30% [15]. The demand for smart buildings, sensor networks, low-latency communications, and specialized automation technology is increasing. Developing Digital Twin to continuously and cost-effectively monitor building facilities to meet energy demands is important.

Hosamo et al. [36] proposed a Digital Twin predictive maintenance framework for detecting automated faults and monitoring the air-handling unit (AHU) performance of an educational building in Norway. The system can identify faults and predict AHU components' future state.

The rapid development of intelligent control systems has highlighted the need to improve visualization tools to better understand the real world, which is essential for future facility management. Extended reality (XR) technology, which encompasses augmented reality (AR), mixed reality (MR), and virtual reality (VR), presents new opportunities for enhancing maintenance activities. XR devices enable users to interact with information in the real world in a more user-friendly way, offering different experiences [50]. Recent studies have examined the benefits of combining XR devices with Digital Twin technology to enhance maintenance procedures and task location [51,52] for facility managers and stakeholders [53–56]. Casini [50] has concluded that using XR technology in building management results in improved human performance and better understanding and optimization of maintenance tasks, which supports more informed decision-making.

2.2. Digital Twin Definition

The Digital Twin (DT) concept dates back to the NASA Apollo project in the 1960s [57], but Grieves first introduced the term in 2002 at the University of Michigan. He developed a model for product life cycle management that mirrored real space in virtual space and linked the two to transfer data [58]. Over the last two decades, the growth of DT technology across various industrial sectors has given rise to multiple definitions of this concept. Grieves and Vickers [59] updated their definition of DT to include a set of virtual information constructs representing physical products from the micro to the macro geometrical level. Physical information can be accessed through its DT. Van Der Hon et al. [60] reviewed 46 DT definitions from review papers and proposed a comprehensive definition as "a virtual representation of a physical system (and its associated environment and processes) that is updated through the exchange of information between the physical and virtual systems".

The application of DT started in the field of aeronautics and has expanded to other areas, such as airframes [61,62], industry [63,64], product service systems [65,66], systems engineering [67], health care [68,69], structural health monitoring [70], product design [71], product life cycle management [72], construction [73], urban planning and transportation [74,75], disaster management [76], infrastructure management [77], agriculture [72,78], and more. While most use cases of DT are in the field of manufacturing [79], the development of DT in the construction industry was not achieved until 2018 [73,80].

The application of DT in buildings has been focused on design [81], maintenance phase [39], and construction in the last phase [82,83]. However, the complex process of the construction industry in management challenges and collecting high-dimension data has caused difficulties in convincing stakeholders of DT's feasibility and investment in time and money [84,85]. Although DT in the building requires high initial investment and expert knowledge [86], it greatly benefits the construction life cycle.

These benefits include design optimization [87], as-built models [88] in the design phase, materials management [89], structure design optimization [90,91], resource management [92], quality management [93], workforce safety [94,95] in the construction phases, energy management [96,97], safety monitoring [98], facility and asset management [99], sustainable development [100], and real-time monitoring in the maintenance phase [73,84].

The concept of the DT in construction is expanding into new phases, such as the demolition and recovery phase, to overcome limited landfill challenges in megacities and waste management concerns [101,102]. Hu et al. [90] presented DT perspectives on construction life cycle aspects using a Six M methodology. They highlighted the advantages of DT in the building life cycle based on an Ishikawa diagram to categorize tools, techniques, and requirements in different building planning stages.

Despite the potential advantages, the lack of standards for adopting DT in construction remains a challenge. BIM facilitates collaboration among different professionals and enables the implementation of new technologies required for DT adoption [103]. However, integrating data from various sources, including BIM, IoT, BAS, and CMMS, can be challenging due to variations in the construction industry [104].

To address these challenges, construction companies must develop a clear digital strategy that includes a technology adoption roadmap, data governance policies, and employee training programs. Construction companies can focus on adopting a sociotechnical systems perspective, which considers both the technical and social factors that can affect the success of digital transformation initiatives. Successful implementation of digital transformation and DT in construction can lead to improved operational efficiency, reduced costs, and data-driven decision-making. However, the full advantages of DT can only be realized if organizations undergo a digital transformation and prioritize the enhancement of their digital capabilities [105].

2.3. Digital Twin for Facility Management

The use of dynamic or smart condition monitoring has become increasingly popular in recent times. However, many existing monitoring systems could be more efficient and provide accurate data in real time to react quickly to critical changes. These maintenance solutions collect historical data and monitor system conditions that reflect one step behind the current state. In contrast, the DT model continuously uses information collected at any moment from HVAC systems, lighting fixtures, indoor sensors, and other devices to provide an accurate and up-to-date representation of its physical twin.

A DT is not just a passive copy of the real system, but an active component that can continuously monitor the condition of its physical twin and provide experts with recommendations on process optimization, maintenance forecasting, planning, and design improvement. DT uses IoT, artificial intelligence (AI), and BIM to integrate a virtual object with a physical object throughout its entire life cycle. By combining DT with predictive maintenance, real-time data from every facility's physical asset can be obtained and continually compared with historical data.

Every change in the systems and facilities is updated in the 3D digital model, allowing the DT to recognize when a component is broken down or requires inspection. Facility managers receive alarms, allowing them to know how and when to react and predict the system's condition in different scenarios through simulation. The ability to analyze different maintenance approaches provides the opportunity to anticipate and evaluate potential outcomes prior to deploying a maintenance team to the site. As a result, they can rely on the DT to extend the overall asset life cycle, avoid overservicing of assets, and keep maintenance costs in check.

With the expanding concept of automated maintenance management and the growing interest in IoT, intelligent buildings generate a large amount of data. Xie et al. [4] proposed a DT data platform to tag real-time data to support dynamic building asset management.

3. Search Strategy

In this study, two literature databases, Scopus and Web of Science (WOS), and an academic search engine, Google Scholar, were selected because of their extensive coverage of scientific and engineering fields [106]. Scopus provides a wider range of journals and offers faster citation analysis than WOS. However, it is important to acknowledge that Scopus primarily includes recent articles [107]. Additionally, both databases exhibit disparities in coverage across countries and languages [108]. These databases provide access to high-quality peer-reviewed journals, conference proceedings, and books across various academic fields. In addition, advanced search features are available, allowing researchers to save search queries and export search results in various formats, including Excel, which are useful for quantitative analysis.

On the other hand, Google Scholar is a freely accessible web search engine that provides a comprehensive collection of citation data, surpassing Scopus and WOS in terms of coverage [109]. However, its citation information is often incomplete and infrequently updated [107]. Using all three resources can ensure comprehensive coverage of relevant papers across a range of disciplines and minimize the possibility of missing important studies.

To gain an overall insight and overview of the concept of DT, a broad search was conducted using the main keyword "Digital Twin" to identify the different fields where the concept is being used, as well as the countries and time periods where research has been conducted. The search was conducted using the Scopus database, which offers comprehensive coverage of scholarly literature [107] and provides advanced tools, including "Scival", which offers graphical representations of research activity and impact.

In the next step, a two-stage search strategy was conducted to refine the search to the subject of "Digital Twin for fault detection and diagnosis (FDD) of building operations" using three academic resources: Scopus, WOS and Google Scholar. The details of the search strategy and the results are provided in the following subsections.

According to the Scopus database, papers across different fields that consecutively used the keyword search "Digital Twin" in their title have been published since 1973, as illustrated in Figure 1. The number of publications has continuously grown in the last two years, reaching up to 1000 papers each year, indicating the increasing significance of DT in various subject areas. However, the adoption of DT in the building industry started relatively late, in 2018, compared to the overall usage of DT.

Around 30.8% of the documents related to DT are in the field of engineering, and China has the highest number of publications in this area, as shown in Figures 2 and 3.

This study aimed to review the DT concept in building facilities, classify FDD methods, and explore how DT can aid the FDD process. To achieve the objectives of this review, we retrieved papers related to the DT in smart facility management in buildings. A mixed review technique was employed to understand the domain knowledge comprehensively. Quantitative analyses using Scopus and VOSviewer software tools were conducted to comprehend relevant papers.

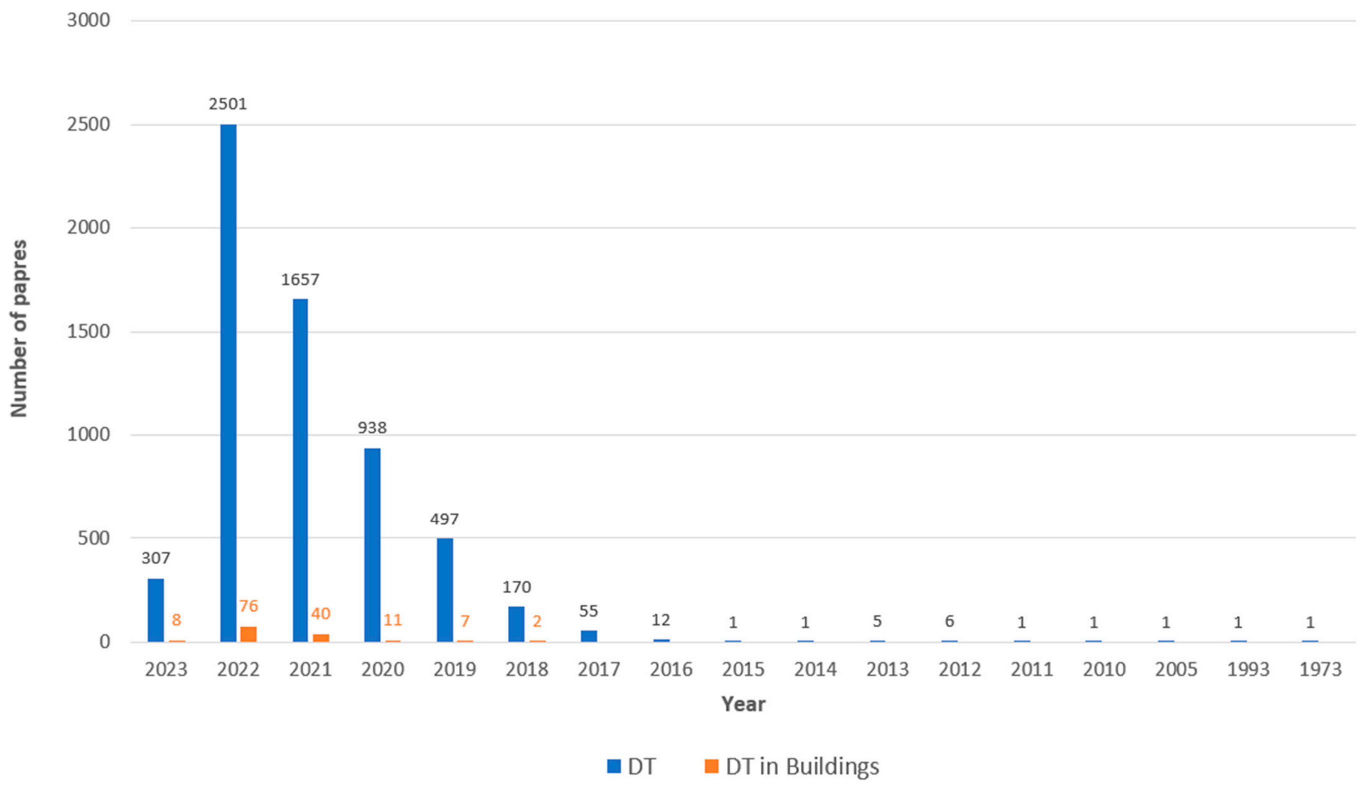

Figure 1. Documents by publication year.

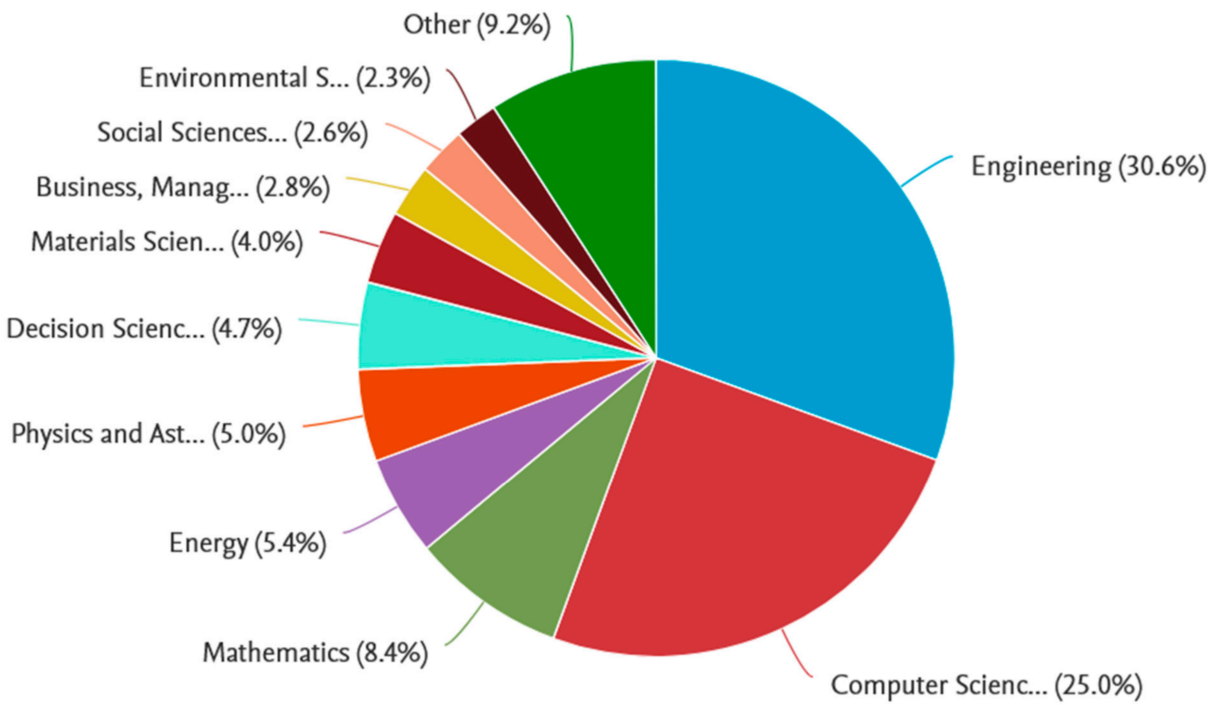

Figure 2. Documents by subject area.

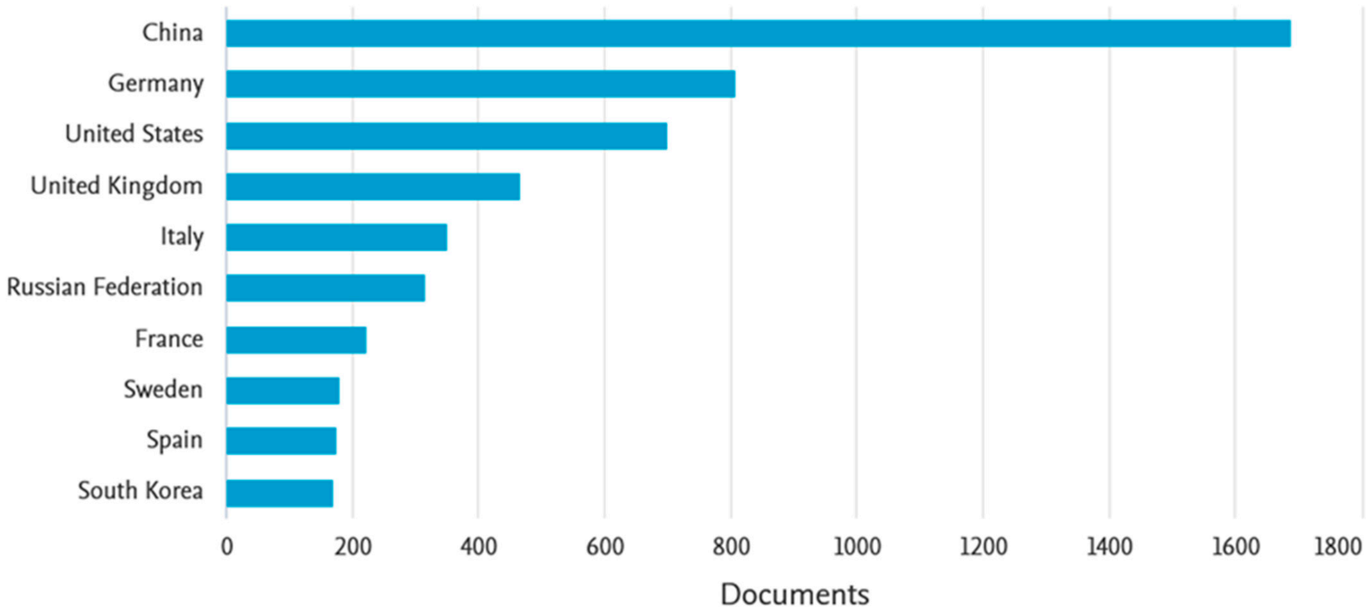

Figure 3. Documents by country.

3.1. Keyword Research

Following the preferred reporting items for systematic reviews (PRISMA) guidelines [110], this study explored the literature in two steps. Firstly, various keywords were examined to define the concept of DTs in buildings. Secondly, FDD methods in building operations were reviewed and classified to determine the most appropriate method for DT for smart maintenance.

The exploratory search was conducted in February 2023 using the keywords "Digital Twin" AND "building". The search was limited to article and review papers and focused on the subject area and source titles related to building on Scopus, WOS, and Google Scholar.

In the second step, keyword research was divided into two categories. In the first category, the concept of smart facility management in maintenance was explored using "predictive maintenance" and "machine learning" from 2019 to 2023. Then, "predictive maintenance" combined with "HVAC systems" and "building" papers published between 2018 and 2023 were extracted to specify research in building facilities.

In the second category, papers on "early fault detection", "fault detection method", and "fault diagnosis method" were reviewed from 2018 to 2023 to discover trend methods in this field. Furthermore, keywords from publications of the previous decade were investigated to enhance our understanding of the methodologies employed in the field of FDD. As a result, the keyword "fault detection and diagnosis in HVAC systems" completed the research. This yielded methods keywords including "data driven", "knowledge-based", "black-box", "gray-box", "first principle model", "expert system", "fuzzy logic", "Bayesian network", "ANN", "LSTM", "KNN, "decision tree", "random forest", "clustering", "unsupervised", "supervised", "SVDD", "SAE", and "PCA."

The PRISMA flowchart can be found in Figure 4, providing a visual representation of the paper selection process. The flowchart outlines the different stages of the search strategy, including the identification, screening, eligibility, and inclusion of papers. Initially, a total of 3340 papers were retrieved from the database search, with 28 additional papers deemed relevant from other citation sources. After the removal of duplicates and unrelated papers, 719 papers were screened by title and abstract, and 219 studies were selected for full-text review. Finally, 200 studies were included in the final analysis after removing duplicates and unrelated papers.

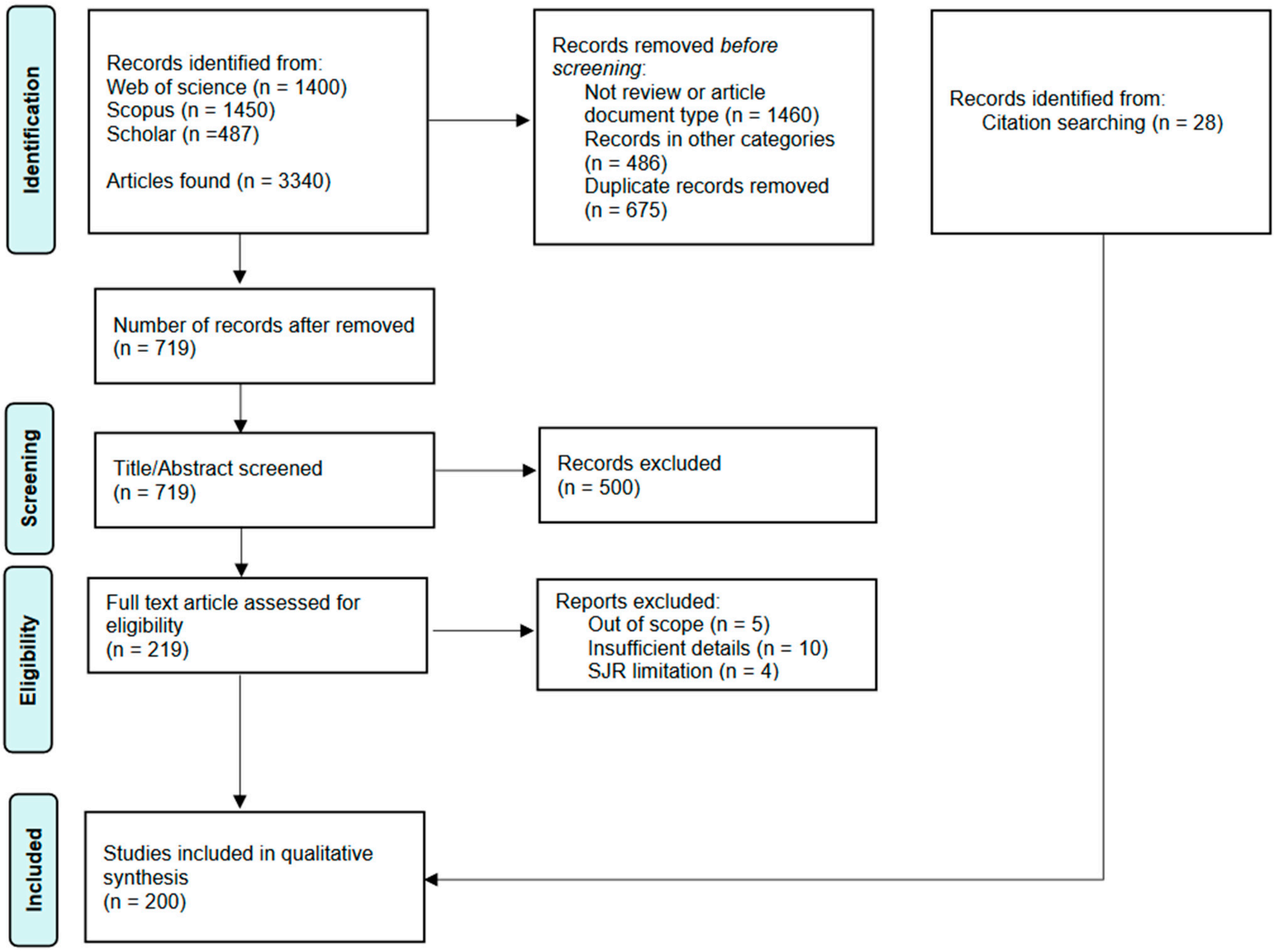

Figure 4. Flowchart of systematic literature review (n = number of papers).

3.2. Quantitative Analysis

VOSviewer, a free tool for analyzing and visualizing research networks [111], was used to show the co-occurrence of keywords from papers in both Scopus and Web of Science, as depicted in Figure 5. VOSviewer employs clustering algorithms to group similar keywords into clusters, enabling the identification of important research themes. In the analysis, each keyword is represented as a node, and the strength of association between keywords is indicated by line thickness and node distance [112].

A total of 833 keywords were extracted from both datasets, with minimum keyword occurrence set at 5. Table 1 presents a summary of the analysis, showcasing the top 29 keywords with their occurrence and total link strength values. The occurrence column denotes the number of times each keyword appears in the articles, while the total link strength column indicates the strength of the connections between the keywords in the network. VOSviewer calculates the total link strength based on the number and strength of co-occurrences between keywords. A higher total link strength value indicates a stronger connection between the keyword and other keywords in the network. BIM and IoT emerged as the most prominent keywords for the topic of DT.

VOSviewer color-codes the keywords based on their appearance in different time periods. As shown in Figure 5, the yellow-colored keywords represent the most analyzed topics in the recent period, from 2020 to 2022. The overlay visualizations of keywords presented in Figure 5 indicate that predictive maintenance, energy efficiency, blockchain, extended reality, metaverse, deep learning, and historic building information modeling have emerged as trended keywords since 2020, as shown.

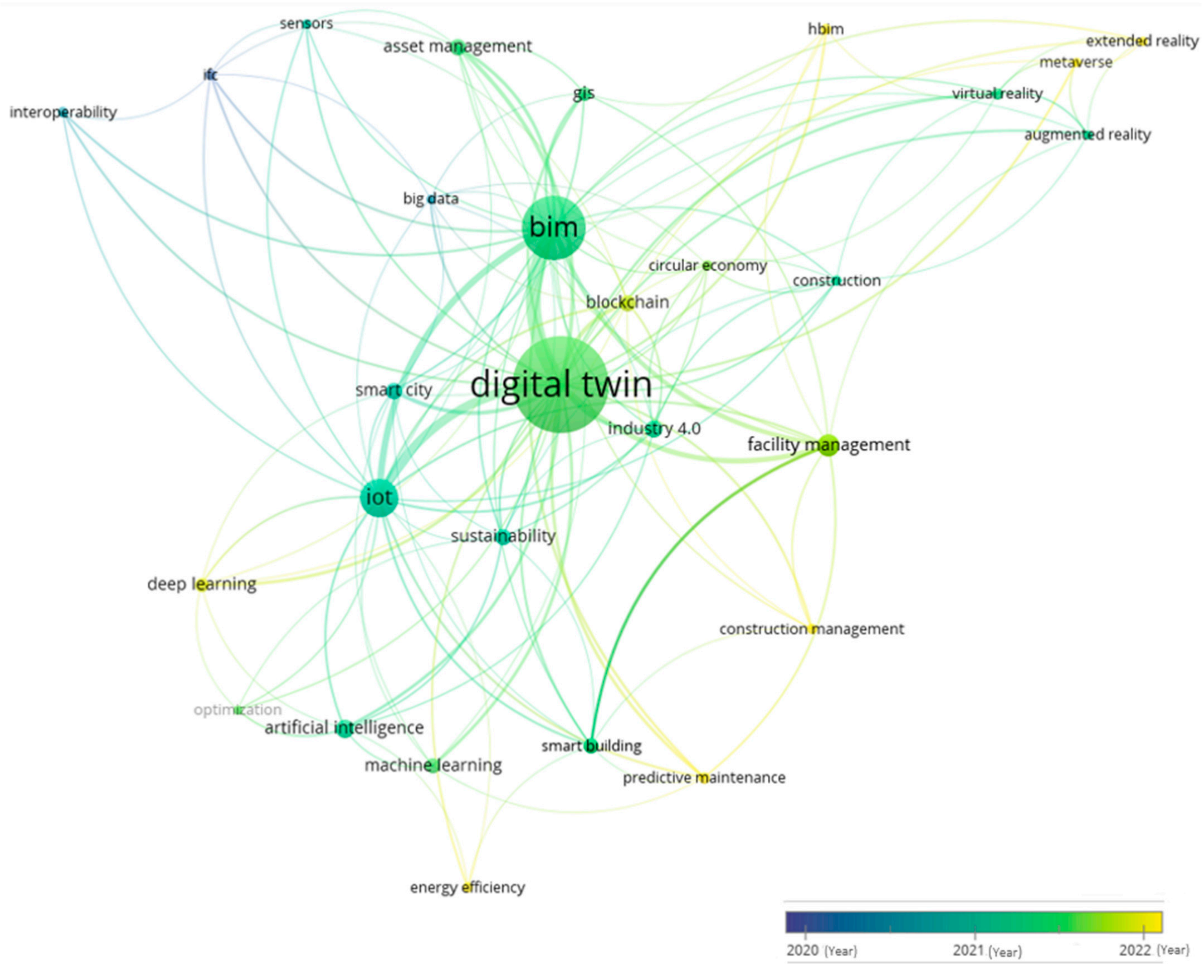

Figure 5. Overlay visualization of frequent terms between 2020 to 2022.

Table 1. Prominent keywords with network parameters.

No.	Keyword	Occurrences	Total Link Strength
1	Artificial Intelligence	12	22
2	Asset Management	10	24
3	Augmented Reality (AR)	4	13
4	Big Data	4	12
5	Building Information Modeling (BIM)	83	152
6	Blockchain	10	30
7	Circular Economy	5	15
8	Construction	5	16
9	Construction Management	5	12
10	Deep Learning	8	14
11	Digital Twin	158	241
12	Energy Efficiency	5	6
13	Extended Reality (XR)	4	7
14	Facility Management	13	41
15	Geographic Information System (GIS)	7	16
16	Historic Building Information Modelling (HBIM)	5	8
17	Industry 4.0	11	28
18	Industry Foundation Classes (IFC)	4	9
19	Interoperability	5	12
20	Internet Of Things (IoT)	37	99
21	Machine Learning	9	13
22	Metaverse	4	10
23	Optimization	4	9
24	Predictive Maintenance	6	20
25	Sensors	4	9
26	Smart Building	7	16
27	Smart City	10	20
28	Sustainability	10	19
29	Virtual Reality (VR)	6	17

4. Fault Detection and Diagnosis

According to keyword analysis, predictive maintenance has become one of the most prominent keywords in recent years. Early fault detection and diagnosis (EFDD) is a critical subset of predictive maintenance that is gaining increasing attention in facility management. EFDD involves using data analysis tools and algorithms to detect anomalies in the performance of building systems and components before they result in significant failures. By identifying faults early on, facility managers can address them proactively and avoid costly repairs or replacements. As such, there is a growing need to investigate further EFDD methods used in recent papers, classify them, and identify the most effective approaches for implementing them in smart building systems. In addition, it is crucial to determine which EFDD method is most capable of integrating with DT to maximize its potential in smart facility management.

FDD methods can be classified into three categories—analytical-based, knowledge-based, and data-driven methods [113–117]—as shown in Figure 6. Analytical-based methods rely on mathematical models and physical laws to detect faults and anomalies in building systems. Knowledge-based methods, on the other hand, use expert knowledge and rules to detect faults and make decisions. Finally, data-driven methods utilize statistical analysis and machine learning algorithms to identify patterns and anomalies in data to detect faults. These three categorized methods can contain different fault detection and diagnosis techniques, enabling us to compare them effectively.

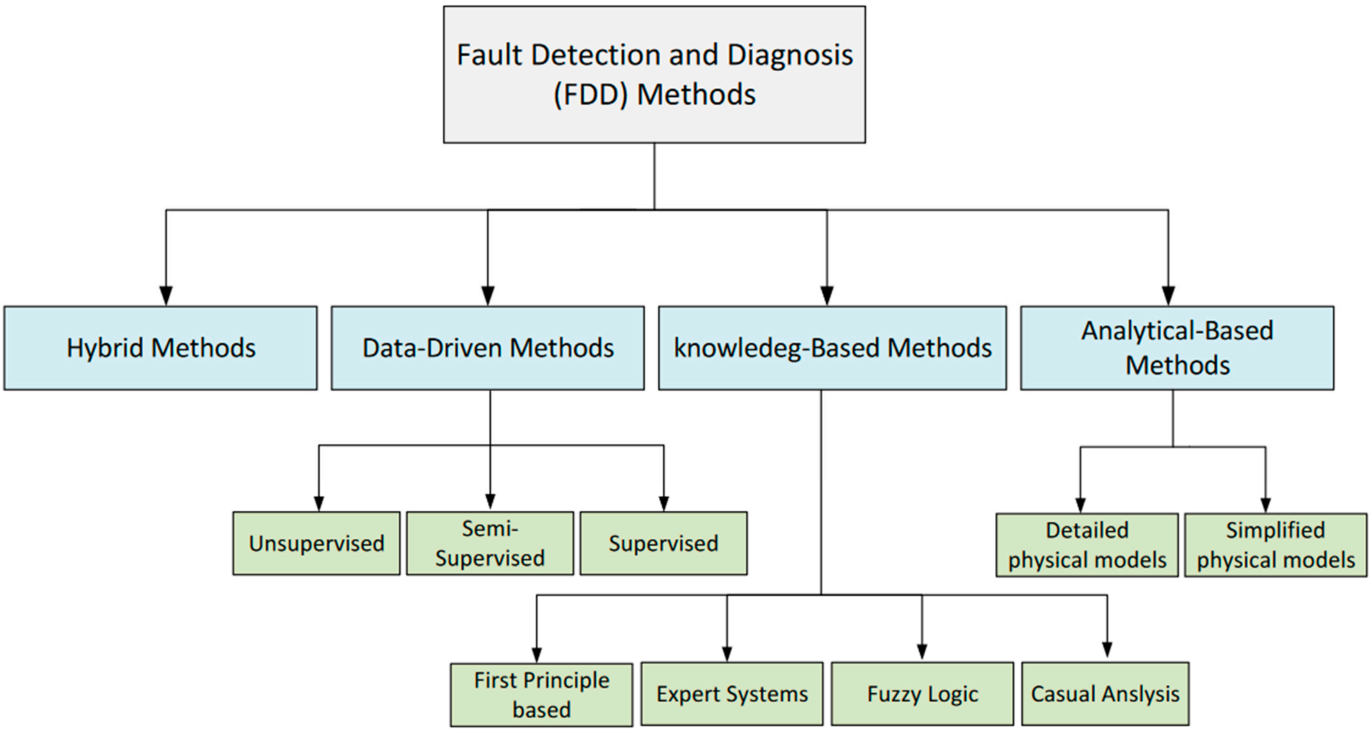

Figure 6. Classification of FDD methods in building operations based on [102–105,107–109].

4.1. Analytical-Based Methods

Analytical-based methods for FDD are based on first principles and a physical understanding of the system, which involves building a mathematical model to detect faults by comparing residuals and measured data [113,115,118]. These methods are useful for routine operations and maintenance in smart buildings to reduce errors [19], and can be categorized into simplified and detailed physical models [119].

Gray-box methods are a well-known subset of analytical-based methods that use empirical data and mathematical models to identify faults. Gray-box methods are a hybrid approach between purely data-driven and model-based methods. They use limited knowledge about the system and combine it with data analysis techniques to improve FDD performance. Analytical-based methods involve mathematical models and physical laws to identify faults, while data-driven methods rely on data analysis and machine learning algorithms. Compared with white-box models, gray-box models are faster and more accurate, require fewer data, and have better reliability, although they require expert knowledge. Because most building processes are similar, gray-box models can be easy to model [120–122].

Ahamed et al. [123] proposed two gray-box models for predicting the supply air temperature for air-handling units. The models were designed to be used as virtual sensors for automated HVAC system fault detection, which can be a cost-effective alternative to physical sensors. The models' inputs include mixed air temperature, cooling coil valve signal, and chilled water inlet temperature to predict the supply air temperature, and they were tested in two buildings. The authors concluded that model B had better accuracy and prediction performance.

Ranade et al. [124] proposed a computationally efficient method for fault diagnosis of fan-coil unit terminals in building HVAC systems. The proposed method combines physical and statistical models in a gray-box approach to detect faults in the HVAC system. The method is tested on an actual building, and the results show that it can detect faults with high accuracy and computational efficiency. The proposed method can be used for proactive maintenance of HVAC systems and energy savings.

4.1.1. Detailed Physical Models

Detailed physical models, such as feedforward and autoregressive exogenous (ARX) models [125], require a comprehensive understanding of the physical relationships between all system components and can simulate both normal and faulty states of the system [119,126]. They typically use equations and mathematical models to describe the physical relationships between various system components and how they interact.

Andriamamonjy et al. [127] proposed an automated model-based approach for FDD in air-handling units (AHUs) using building information modeling (BIM) and Modelica. The authors developed a detailed physics-based model of the AHU and used it to generate residuals, which were compared to threshold values for fault detection. A diagnostic algorithm was then used to identify the root cause of the fault. The approach was tested on a real AHU system, and the results showed that the method can detect and diagnose faults accurately and quickly. The proposed approach can be used for proactive maintenance of AHU systems, reducing energy consumption, and improving indoor air quality.

4.1.2. Simplified Physical Models

Simplified physical models use simplified assumptions and lumped parameter approaches to transform coupled space partial differential equations into ordinary differential equations [119,126], in contrast to detailed physical models. They can use performance metrics and reduced-order models [125] and require less training information than black-box models or data-driven methods, making them more practical for applications [128].

Zhao et al. [128] developed a simplified physical model-based FDD strategy for centrifugal chillers based on Cui and Wang's [129] model. The approach utilizes a simplified physical model that describes the chiller's dynamic behavior and then develops a residual generator to detect and isolate faults. The authors also proposed a customized tool that integrates the simplified physical model and the residual generator to enable real-time FDD. The proposed model can be used without fault samples and can be applied to model other chillers based on basic principles. The results show that the approach effectively detects and isolates several common faults in the chiller. The customized tool was also demonstrated to provide real-time FDD capabilities for the chiller. The authors suggest that the proposed approach can be used for other types of HVAC systems and can provide such benefits as improved system reliability, reduced energy consumption, and decreased maintenance costs.

4.2. Knowledge-Based Method

Knowledge-based methods are commonly utilized when a system's physical or mathematical modeling is costly, computationally complex, or when there is a lack of data available. They are also suitable for systems with a small number of inputs, outputs, and states or when specialized knowledge is required for modeling. As a result, they are usually applied to small-scale systems with limited operating modes. However, developing rules based on the system's physical properties can take time and effort, making it difficult to understand the system comprehensively [113,130]. Knowledge-based methods can be broadly categorized into several categories, including causal analysis, expert systems, fuzzy logic, and first-principle-based methods.

4.2.1. Causal Analysis

Several methods are used for causal analysis in FDD, including signed directed graphs (SDGs), structural graphs, and fault trees [113,116,117]. With minimal information, digraph-based methods can establish relationships between processes and diagnose faults [131]. Process variables are represented as nodes in an SDG, and their relationships are depicted as direct or simple arcs.

Data-driven methods have gained significant attention in FDD, thanks to the availability of data from building management and other control systems. However, these methods often need more interpretability. Conversely, knowledge-based methods require expert knowledge to interpret causal relationships between symptoms and faults. To address this challenge, Zhang et al. [132] proposed a causal discovery- and inference-based method inspired by expert reasoning. The framework establishes a causal structure graph that connects fault and symptom nodes based on individual average causal effect estimation. Backward structural causal model-based causal inference is then used to detect and diagnose faults based on symptoms. The proposed method was compared to five traditional data-driven methods (KNN, SVM, CART, DNN, and CNN) and was found to have diagnostic accuracy comparable to these methods. Furthermore, the proposed method demonstrated less training and hyperparameter optimization time than DNN and CNN.

4.2.2. Fuzzy Logic

The study of fuzzy logic began in 1920, but the term "fuzzy logic" was coined by Professor Lotfi Asgarzadeh in 1965 at Berkeley University. Fuzzy logic is a subset of Boolean logic that introduces the concept of relative correctness and computes it based on a statement's degree of certainty or correctness, expressed by a number between 0 and 1. It can describe the condition of a system and how it works, making it useful for fault classification and diagnosis [133–135].

Lauro et al. [136] used fuzzy logic to detect abnormal power consumption of building fan coil units. The proposed method involves analyzing fault detection using statistical methods, neighborhood and average value comparisons without errors, and employing a clustering technique. The analysis was implemented after the fault detection stage based on fuzzy sets and fuzzy logic. TThe fan coil was tested on a one-day monitoring dataset in December 2013. The results demonstrated the effectiveness of the proposed method in automating the detection and diagnosis of abnormal power consumption.

Similarly, Sulaiman et al. [137] developed a fuzzy logic controller to settle room temperature and a fuzzy fault detection method for a centralized chilled water system. Two fuzzy logic fault detection methods were used to identify any fault in this simulation work using MATLAB/SIMULINK. The results showed that the controller worked well at the desired temperature, and faults were detected by comparing the actual and normal supply air-flow rates.

In another study, Lo et al. [138] proposed a fuzzy genetic algorithm for automatic fault detection in HVAC systems. A fuzzy system monitors the HVAC system continuously and classifies different fault levels based on the fuzzy rule table optimized by a genetic algorithm. The simulated results demonstrated that this system could distinguish different fault levels even under a limited sensor signal environment.

Zhou et al. [139] proposed an FDD method for chillers based on fuzzy modeling and artificial neural networks using data from the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Research Project RP-1043. The data are first prepared and trained with a fuzzy model. Then, classifications are made using an artificial neural network model, which is effective in pattern recognition and used as a tool for fault detection.

4.2.3. First-Principle-Based Method

First-principle models are rule-based methods built on a fundamental understanding of physical phenomena in systems such as mass and energy. In HVAC systems, these methods involve the development of mathematical models that describe the system's physical components, such as heat exchangers, fans, and pumps, and the interactions between these components. These models are then used to simulate the system's behavior under normal and fault conditions, and the resulting outputs are compared to actual system measurements to detect and diagnose faults. They can accurately estimate the system's dynamic behavior if the model response is faster than the system performance and sudden faults cannot be detected [140].

Norford et al. [141] developed FDD frameworks for AHUs based on the ASHRAE 1020 project, comparing first-principle-model-based and electrical power correlation methods. The first-principle-based method predicts the temperature output of the monitored system to assess its performance. Faults can be detected as a degradation in the expected system performance, and fault diagnosis is processed by expert knowledge. Both methods detected almost all the faults in the two matched AHUs. Still, the first method misdiagnosed several faults, could not recognize the sudden action of the system, and needed more sensors to perform accurately.

4.3. Data-Driven Methods

Data-driven methods refer to approaches that use data to build models or make predictions without relying on explicit physical or mathematical models of the system. Instead, these methods use statistical or machine learning techniques to identify patterns and relationships in the data and make predictions based on those patterns. The goal of data-driven methods is to provide accurate predictions or classifications based on the available data without requiring a detailed understanding of the system's underlying physical or mathematical models. It can be considered a dimensionality reduction technique that learns the pattern classification and error function based on input data. Therefore, this approach is suitable for modern engineering systems with large domains. Data-driven methods can be classified into supervised, semi-supervised, and unsupervised learning.

4.3.1. Supervised Methods

Supervised machine learning methods involve training datasets with labeled inputs and outputs to adjust weights until the model can accurately understand the patterns in the data. There are two types of supervised learning: classification and regression. Classification algorithms include support vector machines (SVM), decision trees (DTs), k-nearest neighbors (KNN), random forest (RF), and supervised neural networks. Regression algorithms include linear, logistical, and polynomial regression.

ANNs are a commonly used type of supervised machine learning for FDD in building operations [16,142–146]. They simulate the structure and function of the human brain's neurons to process and transmit data. ANNs can be trained using a variety of learning algorithms, such as back-propagation, to adjust the weights of the connections between neurons and optimize the network's performance on a specific task. One commonly used ANN model is the multilayer perceptron (MLP), which consists of multiple neurons that process input, transmit the results to other neurons, and continue until a decision is reached.

SVM is another popular supervised machine learning algorithm for classification and regression problems [147]. It can be used for both binary and multiclass classification problems. The basic idea behind SVM is to find the best hyperplane that separates the data into different classes. The hyperplane is chosen to maximize the distance between the nearest data points of the two classes, known as the margin. SVM works by transforming the input data into a higher dimensional space and finding the optimal separating hyperplane to achieve high accuracy and precision [145,148,149]. One of the critical advantages of SVM is its ability to handle high-dimensional data and its robustness to noise.

Cheng et al. [150] developed a data-driven predictive maintenance framework for chillers in three academic buildings in Hong Kong, using 15 variables collected from IoT sensors, a BIM model, and facility management systems. They used ANN and SVM for condition prediction of these chillers over three months and one year and found that ANN was more accurate, but had a longer processing time than SVM.

RF is a widely used supervised learning algorithm due to its simplicity, usability, and ability to produce excellent results without adjusting meta-parameters [16,145,151]. This algorithm is suitable for both classification and regression tasks. The name "random forest" comes from the fact that it creates a forest of decision trees. These trees are constructed using the "bagging" method, where each tree is trained on a random subset of the training data. The final result of the algorithm is obtained by aggregating the outputs of all the trees in the forest.

Wang et al. [152] proposed a method for detecting and diagnosing multiple faults in variable air volume (VAV) terminals using a self-adaptive zone air-temperature model based on historical data. The method addresses the challenges of dealing with complex and multiple faults, which can be difficult to identify using traditional FDD methods. The proposed method first uses a self-adaptive model to extract features from the data collected from the VAV terminals. Then, a layered random forest algorithm is applied to classify the data and identify any faults present. The authors concluded that the proposed method detected 19 known faults with 96% accuracy.

DTs are supervised learning techniques for classification and regression problems. They utilize a tree structure to analyze decision-making processes and predict responses based on learned decision rules derived from features. Hosamo et al. [36] developed a DT predictive maintenance framework for fault detection of AHUs in educational buildings, using 16 factors collected from BIM, IoT sensors, and facility manager (FM) systems. Three machine learning models—including ANNs, support vector machines (SVMs), and decision trees (DTs)—were applied to predict the system's condition. The results showed that the DTs, followed by SVMs, had better accuracy. However, based on the DTs' prediction accuracy and error indices, the SVM performed better than the other two methods, albeit at the cost of increased computational time.

KNN is a nonparametric supervised learning method [153] used widely for classification and regression because of its effectiveness and simplicity. The KNN method is based on finding the best nearest neighbor (the k value indicates the count of the nearest neighbors), and the quality of the predictions depends on the distance measured. Dey et al. [154] proposed an automated FDD method for the HVAC terminal unit (TU). The framework consists of supervised and unsupervised methods to detect faults remotely. Data from TU was collected using the building energy management system (BEMS) of a seventeen-floor building in London. They applied clustering to similar group data for fault diagnosis and then trained the data using multiclass support vector machine (MC-SVM) to predict faults. In the end, an automation notification is sent to the maintenance expert. MC-SVM is compared with a KNN, a nonparametric method for classification (k as 1 and 3). The result showed that MC-SVM had better accuracy even with less data training.

Liu et al. [155] proposed another supervised method that presented a systematic fault diagnosis approach for chillers. They utilized a classification approach based on the association (CBA) algorithm. They used ASHRAE RP-1043 chiller operation data, filtered using the interquartile range rule algorithm and moving slope-based method. The continuous data were separated to equal distances by the binning method, and the CBA method was developed to diagnose seven faults.

Deep learning methods simplify training large neural network layers and can achieve high accuracy even with limited labeled data [156]. In FDD processes, deep learning can eliminate the need for data preprocessing, feature extraction, and selection, and models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GAN), autoencoders, and long short-term memory (LSTM) are used [157].

Lee et al. [158] developed a real-time FDD method using deep learning for a 10-floor building simulation in Energy Plus to reduce the complexity of data. The method used six variables to train the deep neural network (DNN) for fault diagnosis, including room temperature and the water leaving the cooling coil. The sensors installed on the AHUs were used for this purpose, and the real-time diagnosis showed high accuracy.

Cheng et al. [159] proposed a new FDD method for the AHU in HVAC systems using multiscale convolutional neural networks (MCNNs). The method used offline/online model training data from eight sensors in the system, building management system data, and real-time data for training/updating the MCNN model. The trained model was then used to detect faults using multiscale monitoring signals. The method identified five states of the air-conditioning system, including one normal state and four faults: duct air leakage, reduced fan efficiency, stuck cooling coil valve, and outdoor damper excess. MCNNs were compared with five other common methods, such as neural networks (NNs), PCA, SVM, the wavelet transforms of neural networks, and extreme learning machines, and they showed better performance for all types of errors.

In Yun et al. [160], a supervised autoencoder (SAE) was used for data-driven fault detection to reconstruct the input from ASHRAE 1312-RP data. SAE was used to predict input labels and reconstruct them with better performance than autoencoders [161]. The model diagnosed faults if the reconstruction error of the SAE was within the error limit; otherwise, it could not make a precise judgment. The model was compared with ANN and SVM, and the results showed that the SAE had better performance and was more sensitive in an undefined state.

Taheri et al. [162] studied deep recurrent neural networks (DRNNs) for FDD in HVAC systems. DRNNs of different depths were tested, and the best numbers of hidden layers and layer sizes were chosen. The performance of DRNNs was compared with RF and gradient boosting (GB), and the results indicated that DRNNs had higher precision than the other methods.

In another paper, Yan et al. [163] proposed a back-propagation multidimensional Taylor network (BP-MTN) classifier for diagnosing faults in HVAC air-conditioning units. The BP-MTN classifier uses a fully connected layer, a softmax layer, and a ReLU function to classify faults while reducing model complexity accurately. The paper explored the selection of polynomial orders and activation functions through extensive experiments, showing that the BP-MTN performs better than other methods in accurately classifying all 13 faults in ASHRAE 1312-RP data. The higher the order of the polynomial, the higher the fault diagnosis accuracy, but this increases model complexity so that the ReLU activation function can be used for high accuracy at a lower order.

4.3.2. Semi-Supervised Methods

Semi-supervised methods are useful when the labeled training set is limited, and faulty training data are scarce for FDD in HVAC systems. Such methods involve training a model to detect and diagnose faults using a small amount of labeled data, which comprises both normal and faulty system behavior and a considerable amount of unlabeled data. Techniques such as clustering, active learning, and generative models are typically used to effectively utilize labeled and unlabeled data. These methods classify and update the training data by comparing each input with previous non-faulty data [114,164] and adding new faulty data when necessary.

Li et al. [165] proposed a semi-supervised FDD approach using modified GANs. Modified GANs use labeled and unlabeled data from the building HVAC system to learn the pattern distribution of unlabeled data and combine this information with labeled data. A self-training scheme is developed to modify the imbalanced data. The results showed the effectiveness of the modified GAN in FDD of imbalanced data.

Similarly, Fan et al. [166] addressed the limitation of collecting faulty label data by investigating semi-supervised methods. They proposed a semi-supervised FDD method based on neural networks for AHUs collected from the ASHRAE research project RP-1312. Learning rate, threshold, and iteration were conducted to identify key semi-supervised learning parameters. The results showed that the framework could perform well in detecting unseen faults and diagnosing limited label data.

In another study, Albayati et al. [167] developed two semi-supervised methods and one supervised method for a single packaged rooftop unit (RTU) FDD. SVM was chosen as the supervised method due to its simplicity and high accuracy compared to other methods. The two semi-supervised methods utilized SVM, but they differed from unsupervised methods by utilizing KNN and clustering for the two models. The semi-supervised methods yielded reliable and promising results.

4.3.3. Unsupervised Methods

Unsupervised machine learning is a type of machine learning that can analyze and classify unlabeled datasets. This is particularly useful in complex systems such as HVAC systems, where obtaining accurately labeled data for training may be difficult or expensive. Common unsupervised algorithms include clustering, autoencoders, GANs, principal component analysis (PCA), and association rule mining (ARM).

Cluster analysis, or clustering, is a popular unsupervised learning algorithm that groups similar data into clusters. It is typically used when there are no assumptions about the likely relationships within the data. Examples of clustering algorithms include k-means, c-means, hierarchical, and mean-shift clustering.

PCA is a widely used statistical model for exploratory data analysis and FDD [168–171]. It is a linear dimensionality reduction technique [172] that does not require previous knowledge and can quickly detect faults [173]. However, because the relationships between data are not always transparent, PCA is more suitable for fault diagnosis than fault detection [114,126].

Amruthnath et al. [174] aimed to develop an early fault detection framework for an exhaust fan using unsupervised machine learning algorithms. The framework is based on vibration signals collected from sensors over 12 days. Various features were extracted from the data, and unsupervised machine learning algorithms, including PCA T2 statistic, hierarchical clustering, k-means, fuzzy c-means clustering, and model-based clustering, were applied to train the data. The T2 statistic in PCA had excellent performance for fault detection with minimal prior knowledge, but detecting faults at different levels became more challenging. In situations where cost is critical, clustering could be a better choice, and system conditions can be monitored until faults occur.

Mao et al. [175] proposed a method for chiller sensor fault detection based on empirical mode decomposition (EMD) threshold denoising and PCA. The EMD decomposes the raw data into several intrinsic mode functions (IMFs) and a residual component. Then, the IMFs are filtered by threshold denoising to eliminate the noise. The PCA is then applied to the denoised data to reduce the dimensionality of the feature space and identify the most significant features for fault detection. The authors evaluated the proposed method on a real chiller dataset. This new method resulted in higher data quality and efficiency in fault detection compared to traditional PCA methods.

Another study by Day et al. [176] proposed a machine learning-based multilevel automatic fault detection system for remote HVAC fan coil unit (FCU) behavior. The study was conducted on a commercial building in London, England to detect three FCU faults. The algorithm was implemented in two steps. First-level clustering was performed to separate the patterns of faulty and non-faulty FCUs and analyze each group. Second-level clustering was implemented to detect more fault groups. Three clustering methods, k-means, average linkage hierarchical clustering, and Gaussian mixture model (GMM), were used. Results showed that the clustering of the GMM performed better than the other two clustering methods in the first-level clustering.

LSTM is an unsupervised deep learning method designed to learn long-term dependence and overcome vanishing data over time [177]. Unlike standard feedforward neural networks, LSTM has feedback connections, such as an RNN, and can process single data points and sequences of data, text, and speech data [178]. It is used in forecasting time series [179] and fault detection.

Boabdalai et al. [180] proposed a model for predicting maintenance requirements in building facilities using an LSTM-based approach. They conducted a case study on two air-conditioning systems, three boilers, and three pumps in a 15,000-square-meter sports center in Paris. The data was trained using LSTM, and if the data exceeded a defined threshold, it was considered an error and recorded in the management system. The facility management was notified via SMS. The model detected three types of errors: correct errors confirmed by the technician, false warnings, and undetected errors. Feedback on the results was collected by the management team and stored in the computerized management and maintenance system to update the model.

GAN is an unsupervised machine learning algorithm designed by Goodfellow and colleagues in June 2014 [181]. The GAN learns to generate artificial data with the same statistics as the training data, which can help rebalance faulty and non-faulty datasets, significantly improving model accuracy and performance. Yan et al. [182] developed an unsupervised learning technique, the generative adversarial network method. They used real-world chiller data from ASHRAE project 1043-RP and fault samples generated by rebalancing with the conditional Wasserstein generative adversarial network (CWGAN) algorithm. A multiclass SVM evaluated the classification performance using the customized CWGAN framework, and seven typical faults were investigated.

Choi and Yoon [183] proposed an autoencoder-based FDD for BASs. The autoencoder (AE)-based model is used for the fault detection phase, and two types of information-generated structures were developed for the fault diagnosis process: recurrent error model (REM) and latent space model (LSM). The LSMs performed better than the REMs in the diagnosis process.

4.4. Hybrid Methods

Hybrid methods in the context of FDD combine analytical, knowledge-based, and data-driven techniques to achieve more accurate and reliable FDD results. These hybrid methods can leverage the strengths of each technique to overcome their limitations and improve overall FDD performance. For example, analytical methods can provide a physical understanding of the system, knowledge-based methods can incorporate expert knowledge and heuristics, and data-driven methods can handle complex nonlinear relationships and patterns in data. By combining these approaches, hybrid methods can achieve better FDD accuracy, robustness, and efficiency.

Zhu et al. [184] proposed a fault detection method for HVAC systems using a combination of LSTM and SVDD algorithms. The LSTM captures the system's dynamic behavior by modeling the time series data. The output of the LSTM model is fed into the SVDD algorithm to detect anomalies in the system. The method is evaluated on a real dataset and outperforms state-of-the-art methods, showing robustness to data imbalance and noise. The approach can reduce maintenance costs and improve energy efficiency by detecting faults early.

Tun et al. [185] developed a hybrid FDD method that uses both random forest (RF) and SVM. The model uses random forest to prescreen the data and identify potential faults, then uses SVM to diagnose and classify the faults. The hybrid method achieved a 98% accuracy rate for detecting 13 types of HVAC faults, including cooling coil valve faults, exhaust air damper faults, outside air damper faults, duct leakage, and return air fan failures.

A Bayesian network (BN) is a combination of graph and probability theory that shows how a set of random variables are independently related. The graph is used to formulate the probability distribution along with network variables. When the graph's structure is known, probabilistic models can be used to predict the variables. If the structure needs to be clarified, expert knowledge can be used to learn the structure of the model and predict the variables. BNs can be built manually based on expert prior knowledge or used the entire dataset to develop a machine learning model [21,186–188].

Li et al. [21] presented a fault diagnosis framework for HVAC systems using a combination of data-driven methods and expert knowledge. They collected data from three sensors (temperature, pressure, and flow rate) on an AHU in a commercial building. They used a diagnostic Bayesian network (DBN) as a machine learning method to understand the relationship between 9 types of faults and 17 symptoms. The expert identified the fault mechanism using a causal graph and directions when a fault occurred. The results showed that most AHU faults could be recognized, and the framework had good performance and reliability.

In another study, Dong et al. [189] proposed a hybrid FDD framework based on a first-principle-based model and a Bayesian network. They simulated the energy performance of the entire building by utilizing a case study in MATLAB and incorporating information from the BIM model. The output data of energy performance were used as training data for the graphical model, and an anomaly score was calculated based on network learning. The facility team verified the diagnostic results, and AHU and chiller faults were detected using this framework.

Rafati et al. [190] reviewed nonintrusive load monitoring (NILM), which has recently gained significant attention. NILM regulates electricity consumption by measuring the power input of the house. Unlike other methods that require sensors to be installed in the system, NILM does not require separate measurements, which reduces costs. It can identify the wasted energy consumption of HVAC systems by their on/off modes. Fault detection using NILM can be performed by detecting abnormalities using previous patterns in past consumption data, analyzing the act amplitude-frequency spectrum, identifying anomalous transients, and evaluating usage cycles [191]. Machine learning methods used with NILM include SVM, KNN [192], and a rule-based algorithm called UNUM [193,194]. This review also revealed that NILM could identify more faults in rooftop cooling units (RTUs) and reduce remote monitoring costs.

Yan et al. [195] proposed using a generative adversarial network algorithm to address the imbalanced data problem in fault diagnosis for air-conditioning systems. Real-world fault data from ASHRA 1312-RP were used in this study. In the detection phase, the conditional Wasserstein generative adversarial network (CWGAN) was employed to balance the number of normal training data and faulty training data samples. In the fault diagnosis phase, the model increased the number of samples in each type of fault to enhance diagnostic capability. Next, various models, such as RF, SVM, multilayer perceptron, KNN, and DTs, were used to complete the diagnosis and troubleshooting algorithm. The proposed model outperformed supervised learning models and demonstrated better performance.

Lin et al. [196] created a fault correction algorithm to automate control hunting using lambda tuning open-loop rules. The algorithm was incorporated into commercial FDD software. Control hunting is a common fault in commercial buildings that can cause suboptimal performance and early failure of HVAC equipment. The algorithm was successfully tested in an office building, where it automatically detected and corrected hunting faults in nine variable air volume boxes.

5. Discussion

The classification of FDD methods presented in this paper is organized in Figure 7. Recent studies from 2018 to 2023 have focused mainly on data-driven methods. Among these, ANN and SVM have been the most commonly applied supervised methods, while PCA has been widely used for unsupervised methods in recent years, as shown in Figure 8.

Semi-supervised and unsupervised methods are emerging trends in FDD due to their ability to handle large amounts of data and their effectiveness in scenarios with limited labeled data. Additionally, hybrid methods that combine different approaches are gaining popularity in FDD, as they can achieve superior results compared to individual methods.

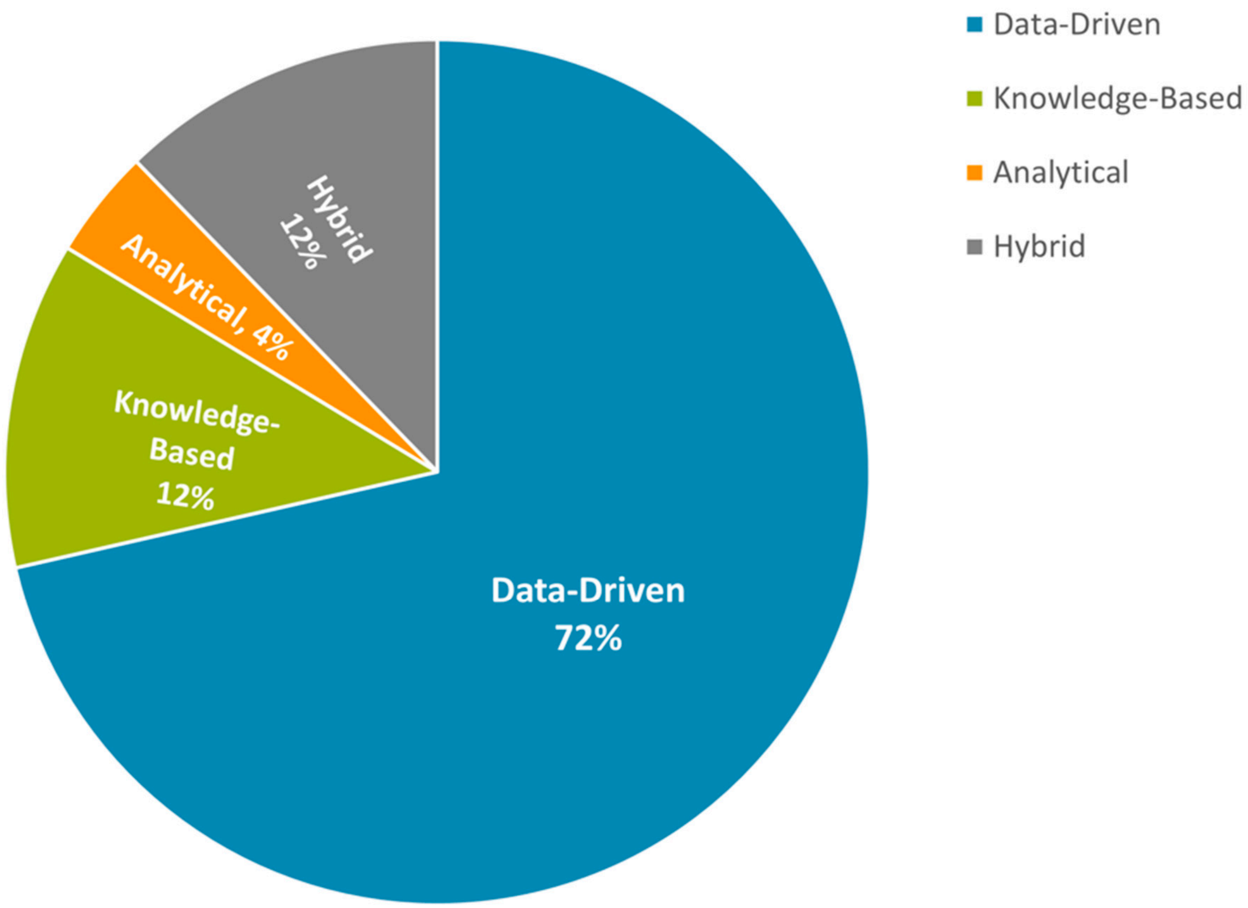

Figure 7. FDD method classification.

Based on the sensors used for data collection and system monitoring in the reviewed papers, Figure 9 presents a classification of the systems studied for FDD methods. The figure shows that AHU systems are the most commonly examined for FDD methods, with frequently installed temperature and flow rate sensors.

Table 2 summarizes the papers reviewed in the field of FDD in buildings.

FDD is a complex process in buildings due to the intricacy of facilities, including HVAC systems, various occupant behaviors, and building types. Therefore, developing a method that can reduce human intervention and interpret the relationship between faults and symptoms functionally is essential. Analytical or physical models can achieve high accuracy when well formulated, and detailed physical models based on first principles can model fault and normal operations. However, they require sufficient sensors and are more suitable for industrial systems than building systems [119,126]. Additionally, they require significant information to apply them to the same systems to simulate some faults [197].

In contrast, simplified physical models need less knowledge to calibrate. They require analysis to find key features, reduce return data by focusing on specific areas [125], and are more capable of future use [119]. Manually capturing information to develop physical models is time-consuming and increases the error-prone. Moreover, comparing the phase with the model output may cause a loss of accurate data at a specific time, because the sample time of building management systems is about 5 min or more [189].

Knowledge-based methods utilize past experiences, such as abnormal patterns, faulty symptoms, and performance criteria, to detect and diagnose faults. These methods can achieve high accuracy in small systems with low complexity or in situations where historical data are not available or the mathematical model of the system is complex and time-consuming [118]. An advantage of this approach is that experts can interpret the FDD process as more physically meaningful based on established models [198]. However, knowledge-based methods may adapt poorly to system changes compared to data-driven methods, potentially leading to lost faults [113,199,200]. Additionally, the development cost of these methods is high and requires specialized knowledge of the system [113,118,130,199,200].

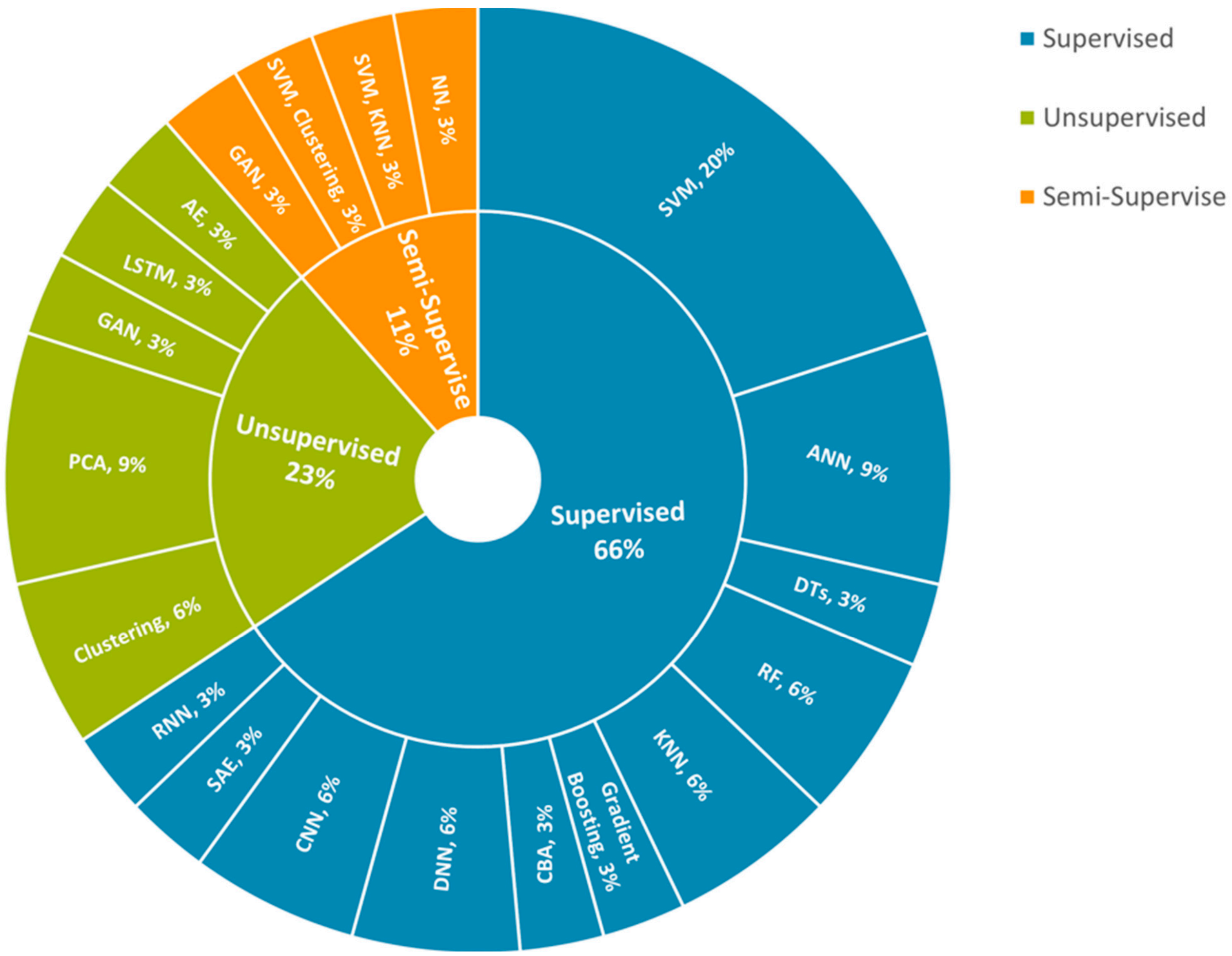

Figure 8. Classification of data-driven methods.

The data-driven methods rely on supervised classification, regression, and unsupervised learning techniques to detect and diagnose faults. However, supervised classification requires significant data and features to achieve good prediction, often limited to experimental or simulation data in studies [198]. On the other hand, unsupervised learning requires massive training data and can produce redundant results without a clear definition of output and data sorting. To address this issue, semi-supervised learning combines both labeled and unlabeled data. However, this method requires higher computational costs than supervised learning. In the diagnosis field, interpretable methods such as decision trees, fuzzy logic, and Bayesian networks are preferred over black-box models with no knowledge of their internal workings.

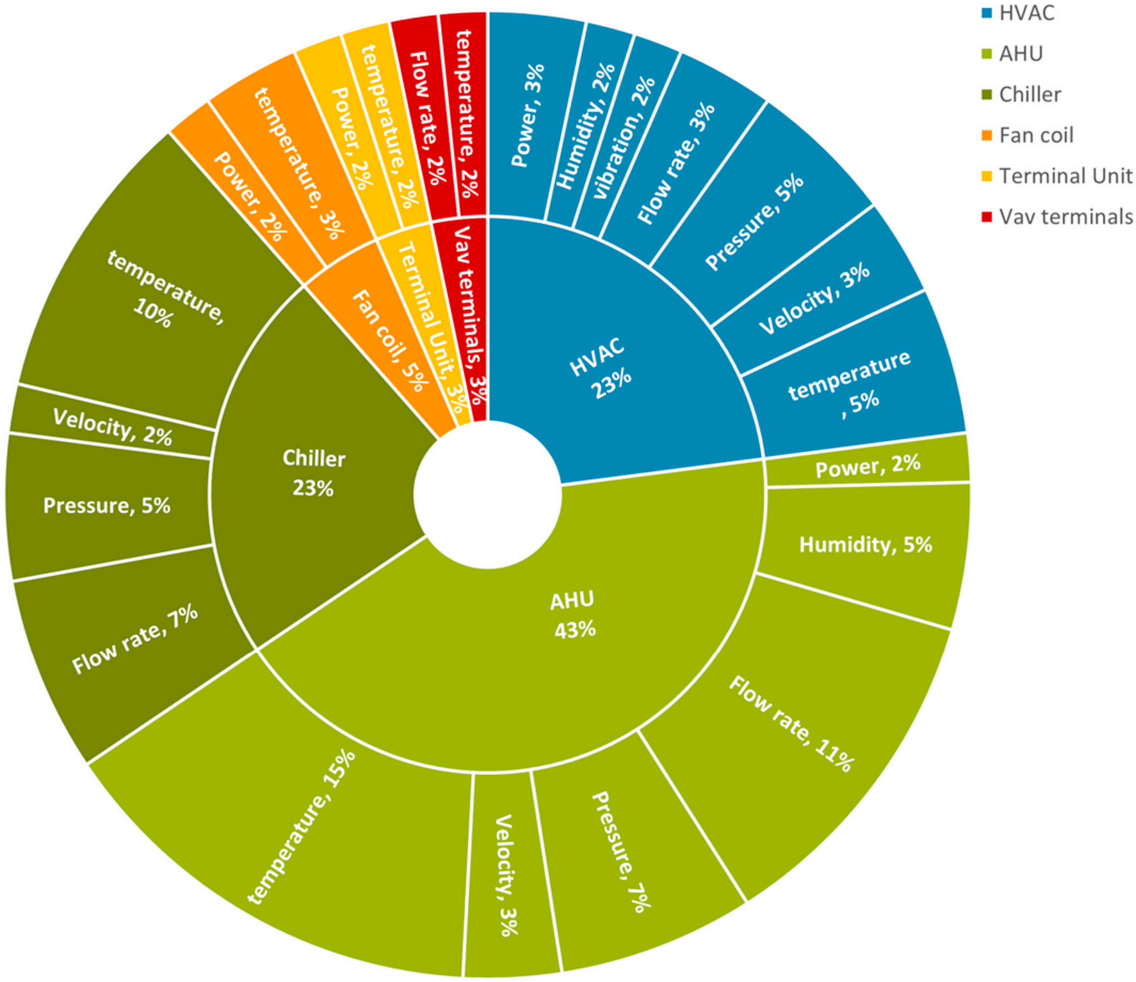

Figure 9. Classification of sensors.

Compared to analytical-based methods and knowledge-based methods, data-driven methods are more commonly used in FDD for smart buildings. These approaches do not require expert knowledge or a physical understanding of system operation. However, the performance and accuracy of these models are highly dependent on the quality and sufficient amount of data. Data-driven approaches have been considered the most reliable method for large-scale systems in recent years. It has been concluded that the accuracy of supervised learning methods relies on the accuracy of training labeled data in a complex system, and the results of unsupervised learning are difficult to interpret. Therefore, the combination of these two methods can evaluate the system at high speed and beyond the knowledge of engineers [114]. With increasing data collected in buildings, deep learning methods have become a trend in FDD methods. They can predict faults in less time and with limited labeled data. However, more studies are needed to apply deep learning methods in actual buildings in the future.

Table 2. Summary of FDD methods for building operation.

Reference	Target	Method	Result
[189]	HVAC FDD	Combination of first-principle-based and Bayesian network (knowledge-based)	Real-time building energy FDD.
[141]	AHU FDD	First-principle model and electrical power correlation method	The first method needs more sensors to perform accurately.
[127]	AHU FDD	Detailed physical model-based (analytical)	BIM can facilitate model-based FDD methods.
[128]	Chiller FDD	simplified physical model-based (analytical)	Does not require fault data.
[132]	HVAC FDD	Casual graph (knowledge driven-based methods)	Compared with five other data-driven methods, proposed method had high accuracy and decreased model training time.
[139]	Chiller FDD	Fuzzy modeling (knowledge-based) and ANN (supervised)	
[136]	FCU FDD	Fuzzy logic (knowledge-based)	Automates abnormal consumption detection and diagnosis.
[180]	AC, boiler, and pump fault detection	LSTM (unsupervised)	
[183]	BASs FDD	Autoencoder (unsupervised, deep learning)	LSMs had better performance than REMs in fault diagnosis.
[150]	MEP components fault alarm	ANN and SVM (supervised)	ANN had better performance than SVM.
[158]	AHU FDD	DNN (supervised)	Good accuracy in real-time diagnosis.
[162]	HVAC FDD	DRNN and RF and GB (supervised)	DRNN had reliable results.
[159]	AHU FFD	MCNNs, NN, SVM (supervised) and PCA (unsupervised)	MCNNs had better performance than mentioned methods.
[36]	AHU fault detection	ANN, SVM, DTs (supervised)	DTs then SVM had better accuracy, but the support vector machine performed better than the other two methods.
[175]	Chiller fault detection	PCA (unsupervised)	Higher data efficiency in fault detection than PCA's traditional method.
[160]	AHU fault detection	SAE, SVM, and ANN (supervised)	The SAE had better performance and was sensitively in an undefined state.
[165]	HVAC FDD	Semi-supervised (based on modified GAN)	High accuracy under imbalanced data.
[166]	AHU FDD	Semi-supervised (based on NN)	Good performance for FDD of limited label data and unseen faults.
[167]	RTU FDD	Semi-supervised (based on SVM, KNN, Clustering) and SVM (supervised)	
[174]	Exhaust fan fault detection	PCA T2 statistic, hierarchical clustering, k-means, fuzzy c-means clustering, and model-based clustering (unsupervised)	T2 statistic had an excellent performance for fault detection with minimum knowledge. Clustering will be a better choice if the cost is essential.
[154]	TU AFFD	MC-SVM, KNN (supervised)	MC-SVM had good accuracy.
[176]	FCU fault detection	k-means, average linkage hierarchical clustering, and GMM clustering (unsupervised)	The GMM model performs better in the first-level clustering than the other two clustering methods.
[155]	Chiller fault diagnosis	CBA algorithm (supervised)	
[184]	HVAC fault detection	Combination of LSTM and support vector data description (unsupervised, data-driven)	
[182]	Chiller FDD	CWGAN (unsupervised)	
[21]	AHU fault diagnosis	DBN (knowledge-based, data-driven)	It had good performance and reliability.
[163]	AHU fault diagnosis	BP-MTN classifier (supervised)	
[195]	AHU FDD	CWGAN (unsupervised), RF, SVM, multilayer perceptron, KNN and decision tree (supervised)	CWGAN had better performance than supervised learning.
[185]	HVAC FDD	Combination of RF and SVM (supervised)	
[152]	vav terminals FDD	RF (supervised)	High accuracy and performance.
[137]	Centralized chilled water system fault detection	Fuzzy logic (knowledge-based)	
[138]	HVAC fault detection	Fuzzy genetic algorithm (knowledge-based)	distinguish different fault levels.
[196]	HVAC FDD	Automated control hunting fault correction algorithm based on lambda tuning open-loop rules (hybrid method)	

Digital Twin technology can facilitate maintenance activities in several ways. Firstly, it can enable real-time monitoring of building components and systems, allowing for early detection of faults and anomalies. This can help to prevent major failures and reduce downtime. Secondly, it can allow for predictive maintenance, where algorithms analyze data from the DT to predict when maintenance activities will be needed, thus optimizing maintenance schedules and reducing costs. Thirdly, it can improve communication and collaboration between stakeholders such as maintenance personnel, building operators, and occupants, leading to more efficient and effective maintenance activities. This paper comprehensively reviews FDD methods for building facilities, discussing their advantages and limitations. The most effective FDD method for using DT technology depends on several factors, such as the type of building, the system or component being monitored, and the available data. However, the papers generally indicate that data-driven methods are becoming more popular due to their ability to handle large amounts of data and identify complex patterns and anomalies. These methods rely on historical data and are well-suited to the context of smart buildings where BAS and CMMS enable system monitoring and data collection.

DT facility management software, which utilizes real-time data from connected sensors and smart IoT devices, presents a new paradigm for data mining. Integrating with the DT concept, appropriate fault detection methods could lead to a new understanding of smart facility management and building operations, focusing on improving occupant comfort and satisfaction. Future studies should address the issue of the transferability of methods, as most building subsystems have similar components with different structures, making it critical to transfer models from one system to another with different faults [132].

6. Summary and Concluding Remarks

Fault detection and diagnosis are essential for building facilities due to their complexity and the need for efficient and effective maintenance. This study systematically reviewed 200 papers on predictive maintenance and FDD methods in buildings, focusing on the potential of Digital Twin technology to improve facility management. This analysis reveals the advantages and limitations of various methods and provides insights into the potential of DT technology to optimize building operations. Analytical-based methods, knowledge-based methods, and data-driven methods are commonly used for FDD in buildings.

While physical models based on first principles can achieve high accuracy, they require sufficient sensors and are more suitable for industrial systems than building systems. On the other hand, knowledge-based methods rely on past experiences to detect and diagnose faults, but they may not adapt to changing conditions. Data-driven methods, which rely on supervised classification, regression, and unsupervised learning techniques, are becoming more popular due to their ability to handle large amounts of data and identify complex patterns and anomalies.

This review concludes that DT, in combination with data-driven methods, provides a practical approach to smart facility management, enabling real-time data collection and analysis to enhance occupant comfort and ensure sustainable operations. DT technology can enhance FDD by enabling real-time monitoring of building components and systems, predictive maintenance, and improved stakeholder communication and collaboration. Additionally, new simulation experiences, such as extended reality technology, offer opportunities to improve predictive capabilities with DT.

However, more research is needed to investigate unsupervised and deep learning methods when fault data from complex systems such as HVAC systems are scarce. Additionally, to ensure the effectiveness of DT in FDD, there is a need for a standardized solution for integrating data from various sources, including building information modeling, building automation systems, energy management systems, and occupancy data.

Construction companies require a clear digital strategy and focus on improving digital transformation to utilize the benefits of DT fully. In addition, to maximize the benefits of DT technology, it is important for stakeholder to carefully consider their goals and objectives before implementing a DT. They should also conduct a thorough analysis of the costs and benefits of the technology and ensure they have the necessary resources and expertise to utilize it effectively.

Future research should be conducted on new construction phases, such as demolition and recovery. Improving the interpretability of data-driven methods is also essential for building operators and maintenance personnel to understand and act upon the results of FDD. Furthermore, the transferability of methods should be addressed in future studies, as most building subsystems have similar components with different structures. Accurate and reliable DT models that can replicate the behavior of physical systems are needed.

With the increasing adoption of renewable energy systems in buildings, there is a need for FDD methods that can detect and diagnose faults in these systems. Finally, a comprehensive evaluation of the economic benefits of FDD is needed to justify its implementation in buildings. It is crucial to investigate how FDD can improve occupant comfort and satisfaction in buildings, as this is a critical factor in the overall success of building operations.

Funding: This research received no external funding.

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

DT	Digital Twin
FDD	fault detection and diagnosis
EFDD	early fault detection and diagnosis
PDM	predictive maintenance
HVAC	heating, ventilation, and air conditioning
PDM	predictive maintenance
BIM	building information modeling
NBIMS-US	National BIM Standard—United States
CMMS	computerized maintenance management systems
BAS	building automation systems
BMS	building management systems
BEMS	building energy management system
AHU	air-handling unit
XR	extended reality
AR	augmented reality
MR	mixed reality
VR	virtual reality
AI	artificial intelligence
IoT	Internet of Things
PRISMA	preferred reporting items for systematic reviews
ARX	autoregressive exogenous
VAV	variable air volume
WOS	Web of Science
GIS	geographic information system
HBIM	historic building information modeling
IFC	industry foundation class
SDG	signed directed graph
NN	neural network
ANN	artificial neural network
DNN	deep neural network
KNN	k-nearest neighbors algorithm
SVM	support vector machine
CNN	convolutional neural network
CART	classification and regression tree
MLP	multilayer perceptron

RF	random forest
GAN	generative adversarial network
FM	facility manager
DT	decision tree
CBA	classification approach based on association
TU	terminal unit
MCNN	multiscale convolutional neural network
SAE	supervised autoencoder
DRNN	deep recurrent neural network
GB	gradient boosting
BP-MTN	back-propagation multidimensional Taylor network
RF	random forest
RNN	recurrent neural networks
PCA	principal component analysis
ARM	association rule mining
FCU	fan coil unit
BN	Bayesian network
NILM	nonintrusive load monitoring
CWGAN	conditional Wasserstein generative adversarial network
AE	autoencoder
REM	recurrent error model
LSM	latent space model

References

1. Dixit, S.; Mandal, S.N.; Sawhney, A.; Singh, S. Relationship between skill development and productivity in construction sector: A literature review. *Int. J. Civ. Eng. Technol.* **2017**, *8*, 649–665.
2. Johnson, R.M.; Babu, R.I.I. Time and cost overruns in the UAE construction industry: A critical analysis. *Int. J. Constr. Manag.* **2020**, *20*, 402–411. [CrossRef]
3. Klepeis, N.E.; Nelson, W.C.; Ott, W.R.; Robinson, J.P.; Tsang, A.M.; Switzer, P.; Behar, J.V.; Hern, S.C.; Engelmann, W.H. The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants. *J. Expo. Sci. Environ. Epidemiol.* **2001**, *11*, 231–252. [CrossRef] [PubMed]
4. Xie, X.; Merino, J.; Moretti, N.; Pauwels, P.; Chang, J.Y.; Parlikad, A. Digital Twin enabled fault detection and diagnosis process for building HVAC systems. *Autom. Constr.* **2023**, *146*, 104695. [CrossRef]
5. Hong, T.; Koo, C.; Kim, J.; Lee, M.; Jeong, K. A review on sustainable construction management strategies for monitoring, diagnosing, and retrofitting the building's dynamic energy performance: Focused on the operation and maintenance phase. *Appl. Energy* **2015**, *155*, 671–707. [CrossRef]
6. Seyedzadeh, S.; Rahimian, F.P.; Glesk, I.; Roper, M. Machine learning for estimation of building energy consumption and performance: A review. *Vis. Eng.* **2018**, *6*, 5. [CrossRef]
7. Li, Y.; O'Neill, Z. A critical review of fault modeling of HVAC systems in buildings. *Build. Simul.* **2018**, *11*, 953–975. [CrossRef]
8. Mihai, S.; Davis, W.; Hung, D.V.; Trestian, R.; Karamanoglu, M.; Barn, B.; Prasad, R.; Venkataraman, H.; Nguyen, H.X. A Digital Twin framework for predictive maintenance in industry 4.0. In Proceedings of the Conference: 2020 International Conference on High Performance Computing & Simulation, Barcelona, Spain, 26–30 July 2021; pp. 1–9.
9. Lee, H.H.Y.; Scott, D. Overview of maintenance strategy, acceptable maintenance standard and resources from a building maintenance operation perspective. *J. Build. Apprais.* **2009**, *4*, 269–278. [CrossRef]
10. Peter, W.T. Maintenance practices in Hong Kong and the use of the intelligent scheduler. *J. Qual. Maint. Eng.* **2002**, *8*, 369–380.
11. Pitt, T.J. Data requirements for the prioritization of predictive building maintenance. *Facilities* **1997**, *15*, 97–104. [CrossRef]
12. Zhan, J.; Ge, X.J.; Huang, S.; Zhao, L.; Wong, J.K.W.; He, S.X. Improvement of the inspection-repair process with building information modelling and image classification. *Facilities* **2018**, *37*, 395–414. [CrossRef]
13. Becerik-Gerber, B.; Jazizadeh, F.; Li, N.; Calis, G. Application areas and data requirements for BIM-enabled facilities management. *J. Constr. Eng. Manag.* **2012**, *138*, 431–442. [CrossRef]
14. Coupry, C.; Noblecourt, S.; Richard, P.; Baudry, D.; Bigaud, D. BIM-Based Digital Twin and XR devices to improve maintenance procedures in smart buildings: A literature review. *Appl. Sci.* **2021**, *11*, 6810. [CrossRef]
15. Mourtzis, D.; Zogopoulos, V.; Vlachou, E. Augmented reality application to support remote maintenance as a service in the robotics industry. *Procedia Cirp* **2017**, *63*, 46–51. [CrossRef]
16. Çınar, Z.M.; Abdussalam Nuhu, A.; Zeeshan, Q.; Korhan, O.; Asmael, M.; Safaei, B. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability* **2020**, *12*, 8211. [CrossRef]
17. Sang, G.M.; Xu, L.; e Vrieze, P.; Bai, Y.; Pan, F. Predictive maintenance in Industry 4.0. In Proceedings of the 10th International Conference on Information Systems and Technologies, Valencia, Spain, 21–25 November 2020; pp. 1–11.

18. Theissler, A.; Pérez-Velázquez, J.; Kettelgerdes, M.; Elger, G. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107864. [CrossRef]
19. Rogers, A.; Guo, F.; Rasmussen, B. A review of fault detection and diagnosis methods for residential air conditioning systems. *Build. Environ.* **2019**, *161*, 106236. [CrossRef]
20. Himmelblau, D.M. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*; Elsevier Science Limited: Amsterdam, The Netherlands, 1978; Volume 8.
21. Li, T.; Zhao, Y.; Zhang, C.; Luo, J.; Zhang, X. A knowledge-guided and data-driven method for building HVAC systems fault diagnosis. *Build. Environ.* **2021**, *198*, 107850. [CrossRef]
22. Eastman, C.; Teicholz, P.; Sacks, R.; Liston, K.; Handbook, B. A guide to building information modeling for owners, managers, designers, engineers and contractors. *BIM Handb.* **2011**, *2*, 147–150.
23. Frequently Asked Questions about the National BIM Standard-United StatesTM. Available online: https://www.nationalbimstandard.org/faqs#faq1 (accessed on 27 January 2023).
24. Boddupalli, C.; Sadhu, A.; Rezazadeh Azar, E.; Pattyson, S. Improved visualization of infrastructure monitoring data using building information modeling. *Struct. Infrastruct. Eng.* **2019**, *15*, 1247–1263. [CrossRef]
25. Chong, H.-Y.; Wang, J.; Shou, W.; Wang, X.; Guo, J. Improving quality and performance of facility management using building information modelling. In Proceedings of the International Conference on Cooperative Design, Visualization and Engineering, Seattle, DC, USA, 14–17 September 2014; pp. 44–50.
26. Li, C.Z.; Guo, Z.; Su, D.; Xiao, B.; Tam, V.W. The Application of Advanced Information Technologies in Civil Infrastructure Construction and Maintenance. *Sustainability* **2022**, *14*, 7761. [CrossRef]
27. Baghalzadeh Shishehgarkhaneh, M.; Keivani, A.; Moehler, R.C.; Jelodari, N.; Roshdi Laleh, S. Internet of Things (IoT), Building Information Modeling (BIM), and Digital Twin (DT) in Construction Industry: A Review, Bibliometric, and Network Analysis. *Buildings* **2022**, *12*, 1503. [CrossRef]
28. Kang, K.-Y.; Wang, X.; Wang, J.; Xu, S.; Shou, W.; Sun, Y. Utility of BIM-CFD Integration in the Design and Performance Analysis for Buildings and Infrastructures of Architecture, Engineering and Construction Industry. *Buildings* **2022**, *12*, 651. [CrossRef]
29. Marocco, M.; Garofolo, I. Integrating disruptive technologies with facilities management: A literature review and future research directions. *Autom. Constr.* **2021**, *131*, 103917. [CrossRef]
30. Motamedi, A.; Hammad, A.; Asen, Y. Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management. *Autom. Constr.* **2014**, *43*, 73–83. [CrossRef]
31. Volk, R.; Stengel, J.; Schultmann, F. Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Autom. Constr.* **2014**, *38*, 109–127. [CrossRef]
32. Wetzel, E.M.; Thabet, W.Y. The use of a BIM-based framework to support safe facility management processes. *Autom. Constr.* **2015**, *60*, 12–24. [CrossRef]
33. Motamedi, A.; Soltani, M.M.; Setayeshgar, S.; Hammad, A. Extending IFC to incorporate information of RFID tags attached to building elements. *Adv. Eng. Inform.* **2016**, *30*, 39–53. [CrossRef]
34. Keller, C. CAFM/IWMS—Balancing Technology, Processes, and Objectives. In *Technology for Facility Managers: The Impact of Cutting-Edge Technology on Facility Management*; International Facility Management Association (IFMA), Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012; pp. 3–20.
35. Nojedehi, P.; O'Brien, W.; Gunay, H.B. A methodology to integrate maintenance management systems and BIM to improve building management. *Sci. Technol. Built Environ.* **2022**, *28*, 1097–1114. [CrossRef]
36. Hosamo, H.H.; Svennevig, P.R.; Svidt, K.; Han, D.; Nielsen, H.K. A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics. *Energy Build.* **2022**, *261*, 111988. [CrossRef]
37. Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a semantic Construction Digital Twin: Directions for future research. *Autom. Constr.* **2020**, *114*, 103179. [CrossRef]
38. Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital Twin: Enabling technologies, challenges and open research. *IEEE Access* **2020**, *8*, 108952–108971. [CrossRef]
39. Lu, Q.; Xie, X.; Parlikad, A.K.; Schooling, J.M. Digital Twin-enabled anomaly detection for built asset monitoring in operation and maintenance. *Autom. Constr.* **2020**, *118*, 103277. [CrossRef]
40. Feroz, A.K.; Zo, H.; Chiravuri, A. Digital transformation and environmental sustainability: A review and research agenda. *Sustainability* **2021**, *13*, 1530. [CrossRef]
41. Savastano, M.; Amendola, C.; Bellini, F.; D'Ascenzo, F. Contextual impacts on industrial processes brought by the digital transformation of manufacturing: A systematic review. *Sustainability* **2019**, *11*, 891. [CrossRef]
42. Reis, J.; Marques, P.A.; Marques, P.C. Where are smart cities heading? A meta-review and guidelines for future research. *Appl. Sci.* **2022**, *12*, 8328. [CrossRef]
43. Tchana, Y.; Ducellier, G.; Remy, S. Designing a unique Digital Twin for linear infrastructures lifecycle management. *Procedia CIRP* **2019**, *84*, 545–549. [CrossRef]
44. Bolshakov, N.; Badenko, V.; Yadykin, V.; Celani, A.; Fedotov, A. Digital Twins of complex technical systems for management of built environment. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Chennai, India, 16–17 September 2020; p. 062045.

45. Cureton, P.; Dunn, N. Digital Twins of cities and evasive futures. In *Shaping Smart for Better Cities*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 267–282.
46. Lu, Q.; Xie, X.; Heaton, J.; Parlikad, A.K.; Schooling, J. From BIM towards Digital Twin: Strategy and future development for smart asset management. In *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future: Proceedings of SOHOMA 2019*; Springer Nature: Berlin, Germany, 2020; pp. 392–404.
47. Peng, Y.; Zhang, M.; Yu, F.; Xu, J.; Gao, S. Digital Twin hospital buildings: An exemplary case study through continuous lifecycle integration. *Adv. Civ. Eng.* **2020**, *2020*, 8846667. [CrossRef]
48. Prušková, K. BIM Technology and Changes in Traditional Design Process, Reliability of Data from Related Registers. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Chennai, India, 16–17 September 2020; p. 032049.
49. Tao, F.; Cheng, J.; Qi, Q.; Zhang, M.; Zhang, H.; Sui, F. Digital Twin-driven product design, manufacturing and service with big data. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 3563–3576. [CrossRef]
50. Casini, M. Extended Reality for Smart Building Operation and Maintenance: A Review. *Energies* **2022**, *15*, 3785. [CrossRef]
51. Williams, R.; Erkoyuncu, J.A.; Masood, T. Augmented reality assisted calibration of Digital Twins of mobile robots. *IFAC-Pap* **2020**, *53*, 203–208. [CrossRef]
52. Prouzeau, A.; Wang, Y.; Ens, B.; Willett, W.; Dwyer, T. Corsican twin: Authoring in situ augmented reality visualisations in virtual reality. In Proceedings of the International Conference on Advanced Visual Interfaces, Ischia Island, Italy, 28 September–2 October 2020; pp. 1–9.
53. Palmarini, R.; Erkoyuncu, J.A.; Roy, R.; Torabmostaedi, H. A systematic review of augmented reality applications in maintenance. *Robot. Comput. Integr. Manuf.* **2018**, *49*, 215–228. [CrossRef]
54. Delgado, J.M.D.; Oyedele, L.; Demian, P.; Beach, T. A research agenda for augmented and virtual reality in architecture, engineering and construction. *Adv. Eng. Inform.* **2020**, *45*, 101122. [CrossRef]
55. Zhu, Y.; Li, N. Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review. *J. Saf. Sci. Resil.* **2021**, *2*, 1–10. [CrossRef]
56. Shafto, M.; Conroy, M.; Doyle, R.; Gleassgen, E.; Kemp, C.; LeMoigne, J.; Wang, L. Draft modelling, simulation, information technology & processing roadmap. *Technol. Area* **2010**, *11*.
57. Grieves, M.W. Product lifecycle management: The new paradigm for enterprises. *Int. J. Prod. Dev.* **2005**, *2*, 71–84. [CrossRef]
58. 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.
59. VanDerHorn, E.; Mahadevan, S. Digital Twin: Generalization, characterization and implementation. *Decis. Support Syst.* **2021**, *145*, 113524. [CrossRef]
60. Liao, M.; Renaud, G.; Bombardier, Y. Airframe Digital Twin technology adaptability assessment and technology demonstration. *Eng. Fract. Mech.* **2020**, *225*, 106793. [CrossRef]
61. Millwater, H.; Ocampo, J.; Crosby, N. Probabilistic methods for risk assessment of airframe Digital Twin structures. *Eng. Fract. Mech.* **2019**, *221*, 106674. [CrossRef]
62. Lattanzi, L.; Raffaeli, R.; Peruzzini, M.; Pellicciari, M. Digital Twin for smart manufacturing: A review of concepts towards a practical industrial implementation. *Int. J. Comput. Integr. Manuf.* **2021**, *34*, 567–597. [CrossRef]
63. Serugendo, G.D.M.; Cutting-Decelle, A.-F.; Guise, L.; Cormenier, T.; Khan, I.; Hossenlopp, L. Digital Twins: From Conceptual Views to Industrial Applications in the Electrical Domain. *Computer* **2022**, *55*, 16–25. [CrossRef]
64. Woitsch, R.; Sumereder, A.; Falcioni, D. Model-based data integration along the product & service life cycle supported by Digital Twinning. *Comput. Ind.* **2022**, *140*, 103648.
65. Leng, J.; Yan, D.; Liu, Q.; Zhang, H.; Zhao, G.; Wei, L.; Zhang, D.; Yu, A.; Chen, X. Digital Twin-driven joint optimisation of packing and storage assignment in large-scale automated high-rise warehouse product-service system. *Int. J. Comput. Integr. Manuf.* **2021**, *34*, 783–800. [CrossRef]
66. Li, C.; Huang, Z.; Huang, Z.; Wang, Y.; Jiang, H. Digital Twins in engineering dynamics: Variational equation identification, feedback control design and their rapid update. *Nonlinear Dyn.* **2022**, *111*, 4485–4500. [CrossRef]
67. Hassani, H.; Huang, X.; MacFeely, S. Impactful Digital Twin in the Healthcare Revolution. *Big Data Cogn. Comput.* **2022**, *6*, 83. [CrossRef]
68. Croatti, A.; Gabellini, M.; Montagna, S.; Ricci, A. On the integration of agents and Digital Twins in healthcare. *J. Med. Syst.* **2020**, *44*, 161. [CrossRef] [PubMed]
69. Dang, H.V.; Tatipamula, M.; Nguyen, H.X. Cloud-based Digital Twinning for structural health monitoring using deep learning. *IEEE Trans. Ind. Inform.* **2021**, *18*, 3820–3830. [CrossRef]
70. Liu, Q.; Zhang, H.; Leng, J.; Chen, X. Digital Twin-driven rapid individualised designing of automated flow-shop manufacturing system. *Int. J. Prod. Res.* **2019**, *57*, 3903–3919. [CrossRef]
71. Wang, Y.; Wang, S.; Yang, B.; Zhu, L.; Liu, F. Big data driven Hierarchical Digital Twin Predictive Remanufacturing paradigm: Architecture, control mechanism, application scenario and benefits. *J. Clean. Prod.* **2020**, *248*, 119299. [CrossRef]
72. Opoku, D.-G.J.; Perera, S.; Osei-Kyei, R.; Rashidi, M. Digital Twin application in the construction industry: A literature review. *J. Build. Eng.* **2021**, *40*, 102726. [CrossRef]
73. Nochta, T.; Wan, L.; Schooling, J.M.; Parlikad, A.K. A socio-technical perspective on urban analytics: The case of city-scale Digital Twins. *J. Urban Technol.* **2021**, *28*, 263–287. [CrossRef]

74. Schrotter, G.; Hürzeler, C. The Digital Twin of the city of Zurich for urban planning. *PFG J. Photogramm. Remote Sens. Geoinf. Sci.* **2020**, *88*, 99–112. [CrossRef]
75. Ford, D.N.; Wolf, C.M. Smart cities with Digital Twin systems for disaster management. *J. Manag. Eng.* **2020**, *36*, 04020027. [CrossRef]
76. Pedersen, A.N.; Borup, M.; Brink-Kjær, A.; Christiansen, L.E.; Mikkelsen, P.S. Living and prototyping Digital Twins for urban water systems: Towards multi-purpose value creation using models and sensors. *Water* **2021**, *13*, 592. [CrossRef]
77. Nasirahmadi, A.; Hensel, O. Toward the Next Generation of Digitalization in Agriculture Based on Digital Twin Paradigm. *Sensors* **2022**, *22*, 498. [CrossRef]
78. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **2020**, *29*, 36–52. [CrossRef]
79. Qi, Q.; Tao, F.; Hu, T.; Anwer, N.; Liu, A.; Wei, Y.; Wang, L.; Nee, A. Enabling technologies and tools for Digital Twin. *J. Manuf. Syst.* **2021**, *58*, 3–21. [CrossRef]
80. Guo, J.; Zhao, N.; Sun, L.; Zhang, S. Modular based flexible Digital Twin for factory design. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1189–1200. [CrossRef]
81. Ferré-Bigorra, J.; Casals, M.; Gangolells, M. The adoption of urban Digital Twins. *Cities* **2022**, *131*, 103905. [CrossRef]
82. Greif, T.; Stein, N.; Flath, C.M. Peeking into the void: Digital Twins for construction site logistics. *Comput. Ind.* **2020**, *121*, 103264. [CrossRef]
83. Brilakis, I.; Pan, Y.; Borrmann, A.; Mayer, H.-G.; Rhein, F.; Vos, C.; Pettinato, E.; Wagner, S. *Built Environment Digital Twining*; Technical University of Munich: Munich, Germany, 2019. Available online: https://publications.cms.bgu.tum.de/reports/2020_Brilakis_BuiltEnvDT.pdf (accessed on 20 November 2022).
84. Madubuike, O.C.; Anumba, C.J.; Khallaf, R. A review of Digital Twin applications in construction. *J. Inf. Technol. Constr. (ITcon)* **2022**, *27*, 145–172. [CrossRef]
85. Madni, A.M.; Madni, C.C.; Lucero, S.D. Leveraging Digital Twin technology in model-based systems engineering. *Systems* **2019**, *7*, 7. [CrossRef]
86. Rafsanjani, H.N.; Nabizadeh, A.H. Towards digital architecture, engineering, and construction (AEC) industry through virtual design and construction (VDC) and Digital Twin. *Energy Built Environ.* **2023**, *4*, 169–178. [CrossRef]
87. Ospina-Bohórquez, A.; López-Rebollo, J.; Muñoz-Sánchez, P.; González-Aguilera, D. A Digital Twin for monitoring the construction of a wind farm. *Eng. Proc.* **2022**, *17*, 3.
88. Liu, Z.; Shi, G.; Qin, J.; Wang, X.; Sun, J. Prestressed Steel Material-Allocation Path and Construction Using Intelligent Digital Twins. *Metals* **2022**, *12*, 631. [CrossRef]
89. Hu, W.; Lim, K.Y.H.; Cai, Y. Digital Twin and Industry 4.0 Enablers in Building and Construction: A Survey. *Buildings* **2022**, *12*, 2004. [CrossRef]
90. He, R.; Li, M.; Gan, V.J.; Ma, J. BIM-enabled computerized design and digital fabrication of industrialized buildings: A case study. *J. Clean. Prod.* **2021**, *278*, 123505. [CrossRef]
91. Münker, S.; Wildemann, P.R.; Göppert, A.; Brell-Cokcan, S.; Schmitt, R.H. Online capability-based resource allocation for on-site construction operations utilizing Digital Twin models. *Constr. Robot.* **2021**, *5*, 211–226. [CrossRef]
92. Wang, Y.; Tao, F.; Zuo, Y.; Zhang, M.; Qi, Q. Digital-Twin-Enhanced Quality Prediction for the Composite Materials. *Engineering* **2023**, *22*, 23–33. [CrossRef]
93. Xu, M.; Nie, X.; Li, H.; Cheng, J.C.; Mei, Z. Smart construction sites: A promising approach to improving on-site HSE management performance. *J. Build. Eng.* **2022**, *49*, 104007. [CrossRef]
94. Hou, L.; Wu, S.; Zhang, G.; Tan, Y.; Wang, X. Literature review of Digital Twins applications in construction workforce safety. *Appl. Sci.* **2020**, *11*, 339. [CrossRef]
95. Francisco, A.; Mohammadi, N.; Taylor, J.E. Smart city Digital Twin–enabled energy management: Toward real-time urban building energy benchmarking. *J. Manag. Eng.* **2020**, *36*, 04019045. [CrossRef]
96. Agostinelli, S.; Cumo, F.; Guidi, G.; Tomazzoli, C. Cyber-physical systems improving building energy management: Digital Twin and artificial intelligence. *Energies* **2021**, *14*, 2338. [CrossRef]
97. Meschini, S.; Pellegrini, L.; Locatelli, M.; Accardo, D.; Tagliabue, L.; Di Giuda, G.; Avena, M. Toward cognitive Digital Twins using a BIM-GIS asset management system for a diffused university. *Front. Built Environ.* **2022**, *8*, 959475. [CrossRef]
98. Zhao, J.; Feng, H.; Chen, Q.; de Soto, B.G. Developing a conceptual framework for the application of Digital Twin technologies to revamp building operation and maintenance processes. *J. Build. Eng.* **2022**, *49*, 104028. [CrossRef]
99. Xia, H.; Liu, Z.; Maria, E.; Liu, X.; Lin, C. Study on city Digital Twin technologies for sustainable smart city design: A review and bibliometric analysis of geographic information system and building information modeling integration. *Sustain. Cities Soc.* **2022**, *84*, 104009. [CrossRef]
100. Kang, K.; Besklubova, S.; Dai, Y.; Zhong, R.Y. Building demolition waste management through smart BIM: A case study in Hong Kong. *Waste Manag.* **2022**, *143*, 69–83. [CrossRef]
101. Akbarieh, A.; Jayasinghe, L.B.; Waldmann, D.; Teferle, F.N. BIM-based end-of-lifecycle decision making and digital deconstruction: Literature review. *Sustainability* **2020**, *12*, 2670. [CrossRef]

102. Stojanovska-Georgievska, L.; Sandeva, I.; Krleski, A.; Spasevska, H.; Ginovska, M.; Panchevski, I.; Ivanov, R.; Perez Arnal, I.; Cerovsek, T.; Funtik, T. BIM in the center of digital transformation of the construction sector—The status of BIM adoption in North Macedonia. *Buildings* **2022**, *12*, 218. [CrossRef]
103. Rebentisch, E.; Rhodes, D.H.; Soares, A.L.; Zimmerman, R.; Tavares, S. The Digital Twin as an enabler of digital transformation: A sociotechnical perspective. In Proceedings of the 2021 IEEE 19th International Conference on Industrial Informatics (INDIN), Palma de Mallorca, Spain, 21–23 July 2021; pp. 1–6.
104. Broo, D.G.; Schooling, J. Digital Twins in infrastructure: Definitions, current practices, challenges and strategies. *Int. J. Constr. Manag.* **2023**, *23*, 1254–1263. [CrossRef]
105. Harzing, A.-W.; Alakangas, S. Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics* **2016**, *106*, 787–804. [CrossRef]
106. Falagas, M.E.; Pitsouni, E.I.; Malietzis, G.A.; Pappas, G. Comparison of PubMed, Scopus, web of science, and Google scholar: Strengths and weaknesses. *FASEB J.* **2008**, *22*, 338–342. [CrossRef] [PubMed]
107. Mongeon, P.; Paul-Hus, A. The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics* **2016**, *106*, 213–228. [CrossRef]
108. Martín-Martín, A.; Orduna-Malea, E.; Thelwall, M.; López-Cózar, E.D. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *J. Informetr.* **2018**, *12*, 1160–1177. [CrossRef]
109. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Moher, D. Updating guidance for reporting systematic reviews: Development of the PRISMA 2020 statement. *J. Clin. Epidemiol.* **2021**, *134*, 103–112. [CrossRef]
110. Van Eck, N.J.; Waltman, L. VOSviewer Manual, Manual for VOSviewer Version. 2011. Available online: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.8.pdf (accessed on 30 January 2023).
111. Van Eck, N.J.; Waltman, L. Visualizing bibliometric networks. In *Measuring Scholarly Impact: Methods and Practice*; Ding, Y., Rousseau, R., Wolfram, D., Eds.; Springer: Berlin, Germany, 2014; Volume 47, pp. 285–320.
112. Alzghoul, A.; Backe, B.; Löfstrand, M.; Byström, A.; Liljedahl, B. Comparing a knowledge-based and a data-driven method in querying data streams for system fault detection: A hydraulic drive system application. *Comput. Ind.* **2014**, *65*, 1126–1135. [CrossRef]
113. Mirnaghi, M.S.; Haghighat, F. Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. *Energy Build.* **2020**, *229*, 110492. [CrossRef]
114. Yu, Y.; Woradechjumroen, D.; Yu, D. A review of fault detection and diagnosis methodologies on air-handling units. *Energy Build.* **2014**, *82*, 550–562. [CrossRef]
115. Zhang, Y.; Jiang, J. Bibliographical review on reconfigurable fault-tolerant control systems. *Annu. Rev. Control.* **2008**, *32*, 229–252. [CrossRef]
116. Leong, C.Y. Fault detection and diagnosis of air handling unit: A review. *MATEC Web Conf.* **2019**, *255*, 06001. [CrossRef]
117. Chiang, L.H.; Russell, E.L.; Braatz, R.D. *Fault Detection and Diagnosis in Industrial Systems*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2000.
118. Katipamula, S.; Brambley, M.R. Methods for fault detection, diagnostics, and prognostics for building systems—A review, part I. *HVAC&R Res.* **2005**, *11*, 3–25.
119. Aguilar, J.; Ardila, D.; Avendaño, A.; Macias, F.; White, C.; Gomez-Pulido, J.; Gutiérrez de Mesa, J.; Garces-Jimenez, A. An autonomic cycle of data analysis tasks for the supervision of HVAC systems of smart building. *Energies* **2020**, *13*, 3103. [CrossRef]
120. Duun-Henriksen, A.K.; Schmidt, S.; Røge, R.M.; Møller, J.B.; Nørgaard, K.; Jørgensen, J.B.; Madsen, H. Model identification using stochastic differential equation grey-box models in diabetes. *J. Diabetes Sci. Technol.* **2013**, *7*, 431–440. [CrossRef] [PubMed]
121. Shi, Z.; O'Brien, W. Development and implementation of automated fault detection and diagnostics for building systems: A review. *Autom. Constr.* **2019**, *104*, 215–229. [CrossRef]
122. Ahamed, M.S.; Zmeureanu, R.; Cortrufo, N.; Candanedo, J. Gray-box virtual sensor of the supply air temperature of air handling units. *Sci. Technol. Built Environ.* **2020**, *26*, 1151–1162. [CrossRef]
123. Ranade, A.; Provan, G.; Mady, A.E.-D.; O'Sullivan, D. A computationally efficient method for fault diagnosis of fan-coil unit terminals in building Heating Ventilation and Air Conditioning systems. *J. Build. Eng.* **2020**, *27*, 100955. [CrossRef]
124. Bruton, K.; Raftery, P.; Kennedy, B.; Keane, M.M.; O'sullivan, D. Review of automated fault detection and diagnostic tools in air handling units. *Energy Effic.* **2014**, *7*, 335–351. [CrossRef]
125. Kim, W.; Katipamula, S. A review of fault detection and diagnostics methods for building systems. *Sci. Technol. Built Environ.* **2018**, *24*, 3–21. [CrossRef]
126. Andriamamonjy, A.; Saelens, D.; Klein, R. An auto-deployed model-based fault detection and diagnosis approach for Air Handling Units using BIM and Modelica. *Autom. Constr.* **2018**, *96*, 508–526. [CrossRef]
127. Zhao, Y.; Wang, S.; Xiao, F.; Ma, Z. A simplified physical model-based fault detection and diagnosis strategy and its customized tool for centrifugal chillers. *HVAC&R Res.* **2013**, *19*, 283–294.
128. Cui, J.; Wang, S. A model-based online fault detection and diagnosis strategy for centrifugal chiller systems. *Int. J. Therm. Sci.* **2005**, *44*, 986–999. [CrossRef]
129. Angeli, C. Diagnostic expert systems: From expert's knowledge to real-time systems. *Adv. Knowl. Based Syst. Model Appl. Res.* **2010**, *1*, 50–73.

130. Kramer, M.A.; Palowitch Jr, B. A rule-based approach to fault diagnosis using the signed directed graph. *AIChE J.* **1987**, *33*, 1067–1078. [CrossRef]
131. Zhang, C.; Zhao, Y.; Zhao, Y.; Li, T.; Zhang, X. Causal discovery and inference-based fault detection and diagnosis method for heating, ventilation and air conditioning systems. *Build. Environ.* **2022**, *212*, 108760. [CrossRef]
132. Ali, Y.H. Artificial intelligence application in machine condition monitoring and fault diagnosis. *Artif. Intell. Emerg. Trends Appl.* **2018**. [CrossRef]
133. Dexter, A.L.; Ngo, D. Fault diagnosis in air-conditioning systems: A multi-step fuzzy model-based approach. *HVAC&R Res.* **2001**, *7*, 83–102.
134. Hellmann, M. Fuzzy Logic Introduction" a Laboratoire Antennes Radar Telecom. *FRE CNRS.* **2001**, *9*, 2272.
135. Lauro, F.; Moretti, F.; Capozzoli, A.; Khan, I.; Pizzuti, S.; Macas, M.; Panzieri, S. Building fan coil electric consumption analysis with fuzzy approaches for fault detection and diagnosis. *Energy Procedia* **2014**, *62*, 411–420. [CrossRef]
136. Sulaiman, N.A.; Othman, M.F.; Abdullah, H. Fuzzy logic control and fault detection in centralized chilled water system. In Proceedings of the 2015 IEEE Symposium Series on Computational Intelligence, Cape Town, South Africa, 7–10 December 2015; pp. 8–13.
137. Lo, C.; Chan, P.; Wong, Y.-K.; Rad, A.B.; Cheung, K. Fuzzy-genetic algorithm for automatic fault detection in HVAC systems. *Appl. Soft Comput.* **2007**, *7*, 554–560. [CrossRef]
138. Zhou, Q.; Wang, S.; Xiao, F. A novel strategy for the fault detection and diagnosis of centrifugal chiller systems. *HVAC&R Res.* **2009**, *15*, 57–75.
139. Cho, S.-H.; Yang, H.-C.; Zaheer-Uddin, M.; Ahn, B.-C. Transient pattern analysis for fault detection and diagnosis of HVAC systems. *Energy Convers. Manag.* **2005**, *46*, 3103–3116. [CrossRef]
140. Norford, L.K.; Wright, J.A.; Buswell, R.A.; Luo, D.; Klaassen, C.J.; Suby, A. Demonstration of fault detection and diagnosis methods for air-handling units. *Hvac&R Res.* **2002**, *8*, 41–71.
141. Bonvini, M.; Sohn, M.D.; Granderson, J.; Wetter, M.; Piette, M.A. Robust on-line fault detection diagnosis for HVAC components based on nonlinear state estimation techniques. *Appl. Energy* **2014**, *124*, 156–166. [CrossRef]
142. Hou, Z.; Lian, Z.; Yao, Y.; Yuan, X. Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Convers. Manag.* **2006**, *47*, 2479–2490. [CrossRef]
143. Magoulès, F.; Zhao, H.-X.; Elizondo, D. Development of an RDP neural network for building energy consumption fault detection and diagnosis. *Energy Build.* **2013**, *62*, 133–138. [CrossRef]
144. Carvalho, T.P.; Soares, F.A.; Vita, R.; Francisco, R.d.P.; Basto, J.P.; Alcalá, S.G. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024. [CrossRef]
145. Malekian, A.; Chitsaz, N. Concepts, procedures, and applications of artificial neural network models in streamflow forecasting. In *Advances in Streamflow Forecasting*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 115–147.
146. Nelson, W.; Culp, C. Machine Learning Methods for Automated Fault Detection and Diagnostics in Building Systems—A Review. *Energies* **2022**, *15*, 5534. [CrossRef]
147. Sexton, T.; Brundage, M.P.; Hoffman, M.; Morris, K.C. Hybrid datafication of maintenance logs from ai-assisted human tags. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1769–1777.
148. Susto, G.A.; McLoone, S.; Pagano, D.; Schirru, A.; Pampuri, S.; Beghi, A. Prediction of integral type failures in semiconductor manufacturing through classification methods. In Proceedings of the 2013 IEEE 18th Conference on Emerging Technologies & Factory Automation (ETFA), Cagliari, Italy, 10–13 September 2013; pp. 1–4.
149. Cheng, J.C.; Chen, W.; Chen, K.; Wang, Q. Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. *Autom. Constr.* **2020**, *112*, 103087. [CrossRef]
150. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
151. Wang, H.; Feng, D.; Liu, K. Fault detection and diagnosis for multiple faults of VAV terminals using self-adaptive model and layered random forest. *Build. Environ.* **2021**, *193*, 107667. [CrossRef]
152. Jung, M.; Niculita, O.; Skaf, Z. Comparison of different classification algorithms for fault detection and fault isolation in complex systems. *Procedia Manuf.* **2018**, *19*, 111–118. [CrossRef]
153. Dey, M.; Rana, S.P.; Dudley, S. Smart building creation in large scale HVAC environments through automated fault detection and diagnosis. *Future Gener. Comput. Syst.* **2020**, *108*, 950–966. [CrossRef]
154. Liu, J.; Shi, D.; Li, G.; Xie, Y.; Li, K.; Liu, B.; Ru, Z. Data-driven and association rule mining-based fault diagnosis and action mechanism analysis for building chillers. *Energy Build.* **2020**, *216*, 109957. [CrossRef]
155. Jin, B.; Li, D.; Srinivasan, S.; Ng, S.-K.; Poolla, K.; Sangiovanni-Vincentelli, A. Detecting and diagnosing incipient building faults using uncertainty information from deep neural networks. In Proceedings of the 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), San Francisco, CA, USA, 17–20 June 2019; pp. 1–8.
156. Hallaji, S.M.; Fang, Y.; Winfrey, B.K. Predictive maintenance of pumps in civil infrastructure: State-of-the-art, challenges and future directions. *Autom. Constr.* **2021**, *134*, 104049. [CrossRef]
157. Lee, K.-P.; Wu, B.-H.; Peng, S.-L. Deep-learning-based fault detection and diagnosis of air-handling units. *Build. Environ.* **2019**, *157*, 24–33. [CrossRef]

158. Cheng, F.; Cai, W.; Zhang, X.; Liao, H.; Cui, C. Fault detection and diagnosis for Air Handling Unit based on multiscale convolutional neural networks. *Energy Build.* **2021**, *236*, 110795. [CrossRef]
159. Yun, W.-S.; Hong, W.-H.; Seo, H. A data-driven fault detection and diagnosis scheme for air handling units in building HVAC systems considering undefined states. *J. Build. Eng.* **2021**, *35*, 102111. [CrossRef]
160. Le, L.; Patterson, A.; White, M. Supervised autoencoders: Improving generalization performance with unsupervised regularizers. *Adv. Neural Inf. Process. Syst.* **2018**, *31*.
161. Taheri, S.; Ahmadi, A.; Mohammadi-Ivatloo, B.; Asadi, S. Fault detection diagnostic for HVAC systems via deep learning algorithms. *Energy Build.* **2021**, *250*, 111275. [CrossRef]
162. Yan, Y.; Cai, J.; Tang, Y.; Chen, L. Fault diagnosis of HVAC AHUs based on a BP-MTN classifier. *Build. Environ.* **2023**, *227*, 109779. [CrossRef]
163. Beghi, A.; Brignoli, R.; Cecchinato, L.; Menegazzo, G.; Rampazzo, M.; Simmini, F. Data-driven fault detection and diagnosis for HVAC water chillers. *Control. Eng. Pract.* **2016**, *53*, 79–91.
164. Li, B.; Cheng, F.; Cai, H.; Zhang, X.; Cai, W. A semi-supervised approach to fault detection and diagnosis for building HVAC systems based on the modified generative adversarial network. *Energy Build.* **2021**, *246*, 111044.
165. Fan, C.; Liu, X.; Xue, P.; Wang, J. Statistical characterization of semi-supervised neural networks for fault detection and diagnosis of air handling units. *Energy Build.* **2021**, *234*, 110733.
166. Albayati, M.G.; Faraj, J.; Thompson, A.; Patil, P.; Gorthala, R.; Rajasekaran, S. Semi-supervised machine learning for fault detection and diagnosis of a rooftop unit. *Big Data Min. Anal.* **2023**, *6*, 170–184.
167. Zhao, Y.; Li, T.; Zhang, X.; Zhang, C. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renew. Sustain. Energy Rev.* **2019**, *109*, 85–101. [CrossRef]
168. Chen, Y.; Lan, L. Fault detection, diagnosis and data recovery for a real building heating/cooling billing system. *Energy Convers. Manag.* **2010**, *51*, 1015–1024. [CrossRef]
169. Du, Z.; Jin, X.; Wu, L. Fault detection and diagnosis based on improved PCA with JAA method in VAV systems. *Build. Environ.* **2007**, *42*, 3221–3232. [CrossRef]
170. Zhao, Y.; Xiao, F.; Wen, J.; Lu, Y.; Wang, S. A robust pattern recognition-based fault detection and diagnosis (FDD) method for chillers. *HVAC&R Res.* **2014**, *20*, 798–809.
171. Wang, S.; Xiao, F. Detection and diagnosis of AHU sensor faults using principal component analysis method. *Energy Convers. Manag.* **2004**, *45*, 2667–2686.
172. Mnassri, B.; Ananou, B.; Ouladsine, M. Fault detection and diagnosis based on PCA and a new contribution plot. *IFAC Proc. Vol.* **2009**, *42*, 834–839.
173. Amruthnath, N.; Gupta, T. A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. In Proceedings of the 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), Singapore, 26–28 April 2018; pp. 355–361.
174. Mao, Q.; Fang, X.; Hu, Y.; Li, G. Chiller sensor fault detection based on empirical mode decomposition threshold denoising and principal component analysis. *Appl. Therm. Eng.* **2018**, *144*, 21–30. [CrossRef]
175. Dey, M.; Rana, S.P.; Dudley, S. A case study based approach for remote fault detection using multi-level machine learning in a smart building. *Smart Cities* **2020**, *3*, 401–419. [CrossRef]
176. Hochreiter, S.; Schmidhuber, J. LSTM can solve hard long time lag problems. *Adv. Neural Inf. Process. Syst.* **1996**, *9*.
177. Bouabdallaoui, Y.; Lafhaj, Z.; Yim, P.; Ducoulombier, L.; Bennadji, B. Natural Language Processing Model for Managing Maintenance Requests in Buildings. *Buildings* **2020**, *10*, 160. [CrossRef]
178. Gasparin, A.; Lukovic, S.; Alippi, C. Deep learning for time series forecasting: The electric load case. *CAAI Trans. Intell. Technol.* **2022**, *7*, 1–25. [CrossRef]
179. Bouabdallaoui, Y.; Lafhaj, Z.; Yim, P.; Ducoulombier, L.; Bennadji, B. Predictive maintenance in building facilities: A machine learning-based approach. *Sensors* **2021**, *21*, 1044. [CrossRef]
180. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*.
181. Yan, K.; Chong, A.; Mo, Y. Generative adversarial network for fault detection diagnosis of chillers. *Build. Environ.* **2020**, *172*, 106698. [CrossRef]
182. Choi, Y.; Yoon, S. Autoencoder-driven fault detection and diagnosis in building automation systems: Residual-based and latent space-based approaches. *Build. Environ.* **2021**, *203*, 108066. [CrossRef]
183. Zhu, H.; Yang, W.; Li, S.; Pang, A. An Effective Fault Detection Method for HVAC Systems Using the LSTM-SVDD Algorithm. *Buildings* **2022**, *12*, 246. [CrossRef]
184. Tun, W.; Wong, J.K.-W.; Ling, S.-H. Hybrid random forest and support vector machine modeling for HVAC fault detection and diagnosis. *Sensors* **2021**, *21*, 8163. [CrossRef] [PubMed]
185. Zhao, Y.; Xiao, F.; Wang, S. An intelligent chiller fault detection and diagnosis methodology using Bayesian belief network. *Energy Build.* **2013**, *57*, 278–288. [CrossRef]
186. Xiao, F.; Zhao, Y.; Wen, J.; Wang, S. Bayesian network based FDD strategy for variable air volume terminals. *Autom. Constr.* **2014**, *41*, 106–118. [CrossRef]

187. He, S.; Wang, Z.; Wang, Z.; Gu, X.; Yan, Z. Fault detection and diagnosis of chiller using Bayesian network classifier with probabilistic boundary. *Appl. Therm. Eng.* **2016**, *107*, 37–47. [CrossRef]
188. Dong, B.; O'Neill, Z.; Li, Z. A BIM-enabled information infrastructure for building energy Fault Detection and Diagnostics. *Autom. Constr.* **2014**, *44*, 197–211. [CrossRef]
189. Rafati, A.; Shaker, H.R.; Ghahghahzadeh, S. Fault Detection and Efficiency Assessment for HVAC Systems Using Non-Intrusive Load Monitoring: A Review. *Energies* **2022**, *15*, 341. [CrossRef]
190. Armstrong, P.R.; Laughman, C.R.; Leeb, S.B.; Norford, L.K. Detection of rooftop cooling unit faults based on electrical measurements. *HVAC&R Res.* **2006**, *12*, 151–175.
191. Lai, Y.H.; Tsai, I.J.; Chiu, C.Y.; Lai, C.F. Non-Intrusive Load Monitoring applied in energy efficiency of the smart manufacturing industry: A case of air-conditioner. In Proceedings of the 2014 IEEE International Conference on Automation Science and Engineering (CASE), New Taipei, Taiwan, 18–22 August 2014; pp. 1127–1132.
192. Rashid, H.; Singh, P. Energy disaggregation for identifying anomalous appliance. In Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments, Delft, The Netherlands, 8–9 November 2017; pp. 1–2.
193. Rashid, H.; Singh, P.; Stankovic, V.; Stankovic, L. Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour? *Appl. Energy* **2019**, *238*, 796–805. [CrossRef]
194. Yan, K.; Huang, J.; Shen, W.; Ji, Z. Unsupervised learning for fault detection and diagnosis of air handling units. *Energy Build.* **2020**, *210*, 109689. [CrossRef]
195. Lin, G.; Pritoni, M.; Chen, Y.; Vitti, R.; Weyandt, C.; Granderson, J. Implementation and Test of an Automated Control Hunting Fault Correction Algorithm in a Fault Detection and Diagnostics Tool. *Energy Build.* **2023**, *283*, 112796. [CrossRef]
196. Bendapudi, S.; Braun, J.E.; Groll, E.A. Dynamic Model of a Centrifugal Chiller System—Model Development, Numerical Study, and Validation. *ASHRAE Trans.* **2005**, *111*, 132–148.
197. Chen, J.; Zhang, L.; Li, Y.; Shi, Y.; Gao, X.; Hu, Y. A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112395. [CrossRef]
198. Fan, C.-M.; Lu, Y.-P. A Bayesian framework to integrate knowledge-based and data-driven inference tools for reliable yield diagnoses. In Proceedings of the 2008 Winter Simulation Conference, Miami, FL, USA, 7–10 December 2008; pp. 2323–2329.
199. Luo, J.; Ghoshal, S.; Mathur, A.; Pattipati, K.R. Adaptive maintenance knowledge bases for field service. In Proceedings of the 2007 IEEE Aerospace Conference, Big Sky, MT, USA, 3–10 March 2007; pp. 1–11.
200. Hongm, W.; Tian-You, C.; Jin-Liang, D.; Brown, M. Data driven fault diagnosis and fault tolerant control: Some advances and possible new directions. *Acta Autom. Sin.* **2009**, *35*, 739–747.