



**DIGITAL TWIN AND AI ENABLED PREDICTIVE
MAINTENANCE IN BUILDING INDUSTRY**

HU WEI

**SCHOOL OF MECHANICAL AND AEROSPACE ENGINEERING
NANYANG TECHNOLOGICAL UNIVERSITY**

2024

DIGITAL TWIN AND AI ENABLED PREDICTIVE MAINTENANCE
IN BUILDING INDUSTRY

HU WEI

SCHOOL OF MECHANICAL AND AEROSPACE ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2024

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

05/October/2023

.....
Date

NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
.....
HU WEI

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

11 Jan 24

.....
Date

NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
.....
Cai Yiyu

Authorship Attribution Statement

This thesis contains material from 4 papers published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 3 is published as **Wei Hu**, Kendrik Yan Hong Lim, and Yiyu Cai*2022. "Digital Twin and Industry 4.0 Enablers in Building and Construction: A Survey" Buildings 12, no. 11: 2004. <https://doi.org/10.3390/buildings12112004>

The contributions of the co-authors are as follows:

- I was the lead author. I prepared the manuscript and conducted the analyses and surveys.
- Kendrik Yan Hong Lim participated in the discussion of the paper framework and revised the manuscript.
- Prof Cai Yiyu provided the initial project direction and revised the manuscript drafts.

Chapter 4 is based on **Wei Hu**, Zhuoheng Xie, and Yiyu Cai*. "A survey of predictive maintenance in the building and construction", which has been submitted to "Journal of Building Engineering" and is under the 2nd round review.

The contributions of the co-authors are as follows:

- I was the lead author. I prepared the manuscript and conducted the analyses and surveys.
- Xie Zhuoheng reviewed the manuscript.
- Prof Cai Yiyu guided the ideas and revised the paper structure.

Chapter 5 is published as **Wei Hu**, Yiyu Cai*.2024. "A semi-supervised method for digital twin-enabled predictive maintenance in the building industry" Neural Comput & Applic. <https://doi.org/10.1007/s00521-024-09926-1>

The contributions of the co-authors are as follows:

- I was the lead author. I constructed the ideas, formulated the deep learning models, wrote the manuscript, and conducted the experiment analysis.
- Prof Cai Yiyu discussed the research ideas and revised the manuscript drafts.

Chapter 6 is published as **Wei Hu**, Xin Wang, Khery Tan, and Yiyu Cai*. 2023. "Digital Twin-Enhanced Predictive Maintenance for Indoor Climate: A Parallel LSTM-Autoencoder Failure Prediction Approach." Energy & Buildings 301 (November): 113738. <https://doi.org/10.1016/j.enbuild.2023.113738>.

The contributions of the co-authors are as follows:

- I was the lead author. I prepared the manuscript and conducted the analyses and surveys.
- Kendrik Yan Hong Lim participated in the discussion of the paper framework and revised the manuscript.
- Prof Cai Yiyu provided the initial project direction and revised the manuscript drafts.

05/October/2023

.....
Date

NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
.....
HU WEI

Acknowledgement

First and foremost, I would like to express my deep and sincere appreciation to my supervisor Professor Cai Yiyu for taking care of me during challenging moments, guiding me through confusing periods, and always supporting me. His expertise in critical thinking, problem-solving, and academic writing has been tremendously beneficial to me. Also, I wish to express my thanks to Professor Frank Guan Yunqing from the Singapore Institute of Technology (SIT) for his professional advice and mentorship for the literature review and research direction.

I would also like to extend my gratitude to the members of the thesis advisory committee: Professor Moon Seung Ki and Professor Zheng Jianmin. Their insightful comments and support have significantly contributed to the enhancement and broadening of my research.

Last but not least, I acknowledge my family and friends for their support during my overseas study life, especially during the challenging period of the COVID-19 pandemic.

Content

List of Figures	XI
List of Tables.....	XIII
Chapter 1 Introduction	16
1.1 Background.....	17
1.1.1 Building Industry and Construction 4.0	17
1.1.2 Digitalisation Development.....	18
1.1.3 Evolution of Maintenance.....	19
1.1.4 Six M Analysis Method	20
1.2 Motivation	21
1.3 Objective and Scope	23
1.4 Contribution of Thesis.....	25
1.5 Outline of Thesis	27
Chapter 2 Literature Review	29
2.1 DT in the Building Industry	29
2.1.1 Data Acquisition	29
2.1.2 Data Processing	31
2.1.3 Modelling and Simulation	32
2.1.4 Decision Support Enablers	34
2.1.5 Computation Enabler.....	40
2.2 PdM in Building Industry	41
2.2.1 PdM Framework.....	41
2.2.2 PdM Methodology	42
2.2.3 PdM Application	43
2.3 Research Gaps	45
2.4 Research Opportunities and Proposed Model.....	47
Chapter 3 The Survey of DT in Building Industry	48
3.1 Integration of Technologies using a DT-adapted Framework.....	49
3.2 DT Perspectives on Construction Lifecycle Aspects Based on a Six M Methodology	52
3.2.1 Machine	52
3.2.2 Manpower.....	53
3.2.3 Material.....	53
3.2.4 Measurement.....	54
3.2.5 Milieu	54
3.2.6 Method.....	55
3.3 Future Directions for DT in Construction 4.0	56
3.4 Leveraging Future Directions for DT-enabled PdM in the Building Industry	58

3.5 Remark.....	60
Chapter 4 The Survey of PdM in Building Industry	65
4.1 Research Requirement and Summary.....	65
4.2 Research Methodology.....	66
4.3 Keywords Network Analyses.....	69
4.4 Three-by-Three M-based Literature Analyses	74
4.4.1 Hard Ms.....	82
4.4.2 Soft Ms.....	83
4.5 Discussion	86
4.5.1 Existing Challenges and Gaps.....	86
4.5.2 Implications for Future Research and Practice	87
4.5.3 Remark.....	90
Chapter 5 A Semi-GAN-based Failure Prediction Approach for Building Facilities	92
5.1 Machine-Perspective PdM Implementation in Building Industry	92
5.2 Research Requirement and Summary.....	93
5.3 DT-enabled PdM Framework in Building Industry	93
5.4 GAN	95
5.5 Semi-supervised GAN.....	96
5.6 Proposed Failure Prediction Method	98
5.6.1 Data Processing	98
5.6.2 Details of Training.....	100
5.6.3 The Implementation of the Failure Prediction Framework	104
5.7 Experiment and Validation.....	106
5.7.1 Dataset Description	106
5.7.2 Experiment Design.....	107
5.7.3 Online Implementation.....	109
5.8 Result and Discussion	110
5.8.1 Evaluation Metrics.....	110
5.8.2 Comparison Under Different Random Seeds.....	111
5.8.3 Effect of the Labelled Data Size	112
5.8.4 Effect of the Data Imbalance Ratio.....	113
5.8.5 Discussion on Adaptability and Scalability	114
5.9 Discussion on Practical Implementation.....	115
5.10 Remark.....	117
Chapter 6 A Parallel LSTM-AE Failure Prediction Approach for Indoor Climate.....	119
6.1 Milieu-Perspective PdM Implementation in Building Industry	119
6.2 Research Requirement and Summary.....	120
6.3 DT-enabled PdM Framework for Indoor Climate.....	121

6.4 LSTM and AE.....	123
6.4.1 LSTM	123
6.4.2 AE	125
6.5 Parallel LSTM-AE Failure Prediction Method.....	127
6.5.1 Problem Formulation.....	127
6.5.2 Data Preprocessing.....	129
6.5.3 Proposed Model.....	130
6.5.4 Model Analysis	132
6.6 Experiment and Validation.....	135
6.6.1 Dataset Description	135
6.6.2 Experiment Design.....	137
6.7 Result and Discussion	141
6.7.1 AE for Failure Prediction.....	141
6.7.2 Effect of Warning Time and Severity.....	143
6.7.3 Comparison with Different Models	145
6.7.4 Effect of Dataset.....	147
6.7.5 Online Deployment.....	150
6.7.6 Discussion on DT-enhanced PdM Implementation.....	152
6.8 Remark.....	153
Chapter 7 Conclusion and Future Work	155
7.1 Summary.....	155
7.2 Contributions.....	156
7.3 Limitation Discussion.....	158
7.4 Future Direction	159
Bibliography	162

List of Figures

Figure 1. Schematic diagram of DT	16
Figure 2. The evolution of the construction industry.....	18
Figure 3. State-of-art digitalisation development.....	19
Figure 4. The evolution of the maintenance approach	20
Figure 5. Research framework and contributions of the thesis	25
Figure 6. A unified technological architecture featuring co-construction elements for Construction 4.0	51
Figure 7. Distribution of existing applications on Six M and BLM aspects.....	52
Figure 8. Identification of Studies.....	67
Figure 9. An overview of the research methodology.....	68
Figure 10. Publication counts by country.....	69
Figure 11. Network visualisation of keywords co-occurrence map about PdM in the B&C industry	70
Figure 12.The Three-by-Three M approach for PdM in B&C industry	75
Figure 13. Distribution of Hard MS Applications & Soft MS methodology	76
Figure 14. A unified framework of DT-enabled PdM in the building industry	94
Figure 15. Schematic of the GAN approach.....	96
Figure 16. Schematic of Semi-supervised GAN approach.....	97
Figure 17. Sample of Semi-supervised GAN training.....	100
Figure 18. Proposed failure prediction framework based on Semi-supervised GAN	104
Figure 19. Data sample & Corresponding sources	105
Figure 20. Online platform for PdM.....	110
Figure 21. Comparison under different random seeds	112
Figure 22. Effect of the labelled data size	113
Figure 23. Effect of the data imbalance ratio.....	114
Figure 24. Data-driven method.....	114
Figure 25. Sensor Deployment Blueprint & Final Plan	117
Figure 26.The framework of DT-enabled PdM for IAQ (up left), Prototype of hardware system (up right), IScan2BIM pipeline (down).....	123
Figure 27. LSTM structure	124
Figure 28. AE	126
Figure 29. USAQI values and categories.....	127
Figure 31. RUL	129
Figure 30. TS	129
Figure 32. Feature correlation analysis	129
Figure 33. The schematic diagram of the proposed model.....	134

Figure 34. LSTM-AE model.....	134
Figure 35. The location distribution of IoT devices.....	136
Figure 36. Predictors and targets of the test set.....	137
Figure 37. Distribution of reconstruction loss (a), Benchmark line in overall index (b).....	142
Figure 38. Failure distribution on the test set by AE component, (a) severity = 100, (b) severity =150, (c) severity = 200	143
Figure 39. Model performance for different severity levels under different warning times.....	145
Figure 40. Failure prediction result of the proposed method (warning time = 30 m, severity = low) ..	147
Figure 41. Comparative study of different benchmark models (warning time = 30 m)	147
Figure 42. Comparative study of datasets with different locations (warning time = 30 m, severity = low)	149
Figure 43. BIM reconstruction from IScan2BIM.....	150
Figure 45. IoT platform (Web & APP)	151
Figure 44. The online platform of PdM for indoor climate.....	151
Figure 46. Data imbalance problem.....	159
Figure 47. Prediction of multiclass faults & concurrent faults	160
Figure 48. BIM model for building facilities (lift)	160
Figure 49. Multi-functional DT platform	161

List of Tables

Table 1. Data acquisition	36
Table 2. Data processing	37
Table 3. Modelling & Simulation	38
Table 4. Decision support enablers.....	39
Table 5. Computation enabler.....	40
Table 6. ML methods in PdM in building facilities	43
Table 7. Benefits of DT applications categorised by Six M aspects based on B&C lifecycle stages	62
Table 8. Future directions for DT in Construction 4.0	64
Table 9. Scientometric analysis of keywords about PdM in B&C industry	71
Table 10. M definition in the Three-by-Three M method.....	76
Table 11. The Three-by-Three M matrix for literature analysis	77
Table 12. Detail analysis of Soft Ms	79
Table 13. Loss function and parameter update.....	97
Table 14. Example of data sample classification	98
Table 15. Hyperparameters tuning.....	102
Table 16. Schematic of self-training.....	108
Table 17. True value & Predictive value	111
Table 18. System configuration.....	115
Table 19. Training time comparison (label data size = 42 days, data imbalance ratio = 1:1)	115
Table 20. Equations explanation in LSTM.....	125
Table 21. The distribution metrics of the test set	137
Table 22. Implementation environment.....	138
Table 23. Hyperparameters tuning	138
Table 24. Definition of TP, TN, FP, and FN.....	139

Abstract

The rapid advancement of information and communication technologies (ICT) and artificial intelligence (AI) has catalysed a significant shift in maintenance practices within the building industry, paving the way for a data-driven paradigm. Predictive maintenance (PdM) has emerged as a critical approach to anticipating failures and reducing unscheduled maintenance tasks. However, the surge in ICT implementations within building-related infrastructure presents several challenges for PdM research and development. Current frameworks are often constrained to specific facilities and lack scalability and generality. Additionally, existing studies focus on condition monitoring and fault detection, with insufficient attention to failure prediction. The effectiveness of PdM has also been hindered by the dependence on labelled datasets, which are expensive and time-consuming to generate. Furthermore, indoor climate management is crucial to building performance. However, it has received less attention than facilities maintenance in PdM research despite its integral role in occupant comfort and environmental sustainability.

In response to these challenges, this thesis introduces a unified framework that integrates Industry 4.0 technologies within a digital twin (DT) structure, grounded in the innovative Six M methodology—Machine, Manpower, Material, Measurement, Milieu, and Method. This approach emphasises the entire building lifecycle, enabling stakeholders to optimise operations, resource allocation, and decision-making across multiple facets of building management. The 6M methodology is crucial for transforming PdM by providing a structured and holistic approach to integrating diverse building assets and operational processes within the DT environment, thereby enhancing system scalability and operational efficiency. The research also employs a Three-by-Three M analysis methodology and a keywords network analysis to identify key research clusters and critical factors in existing DT-enabled PdM-related studies. This analysis underscores the transformative potential of DTs in revolutionising PdM applications across the building industry.

These fundamental studies pave the way for future PdM applications in the building and construction (B&C) industry, significantly advancing maintenance strategies. Following a comprehensive investigation, this thesis introduces a pioneering failure

prediction methodology that addresses the challenge of limited labelled datasets by leveraging semi-supervised generative adversarial networks (GANs). This innovative approach enables the model to utilise labelled and unlabelled data, reducing the reliance on costly manual labelling while improving prediction accuracy. Based on publicly available datasets from building facilities, empirical results demonstrate the model's superior performance in predicting failures, enhancing the system's proactive maintenance capabilities. An online platform was also developed to integrate real-time monitoring with predictive alarms, allowing for efficient, data-driven decision-making in building maintenance.

In addition to traditional facilities maintenance, this thesis extends PdM applications to include indoor climate management, addressing the gap in existing research. Indoor climate, particularly air quality, directly impacts occupant well-being, comfort, and productivity and can be an essential indicator of system failures or inefficiencies. Proper PdM ensures optimal air quality by predicting and preventing system malfunctions that could lead to poor ventilation, temperature control, or humidity levels. The proposed framework incorporates remaining useful life (RUL) and time shift (TS) methods, dividing the prediction tasks into supervised and unsupervised subtasks. A parallel prediction model, combining long short-term memory (LSTM) networks and autoencoders (AE), is developed to handle this complex task. An innovative DT-enabled PdM framework for indoor climates is validated through an online platform that reconstructs 3D building models and provides real-time monitoring and alerts. Experimental results demonstrate the framework's ability to accurately predict faults at varying warning times and severities using practical datasets from buildings in Singapore.

In conclusion, this thesis explores the integration of DT and PdM in the building industry through a comprehensive analysis of four academic papers, leading to significant contributions in system scalability, efficiency, and sustainability. The proposed methodologies enhance building maintenance practices and extend PdM's impact to encompass critical indoor climate factors, paving the way for more resilient, cost-effective, and sustainable building operations.

Chapter 1 Introduction

With increasing interest in automatic and intelligent systems to enhance the B&C industry, DT is gaining popularity as a cost-effective solution to meet stakeholder requirements. As a prevailing Industry 4.0 manufacturing technology, DT is often regarded as a "high fidelity digital model of the physical asset with bi-directional real-time data exchange for simulation and decisions support functions" [1], as shown in Figure 1. Often considered to be a versatile and scalable solution [2] [3] [4], DT offers a cost-effective approach towards resource tracking, scenario simulation, and solution generation. In recent years, there has been a growing interest from both academia and industry alike towards the use of DT systems in the B&C industry, which is often regarded as being behind the curve in terms of intelligence and advancement, with commonly cited improvement areas during design and engineering processes, on-site execution, and facility management (FM).

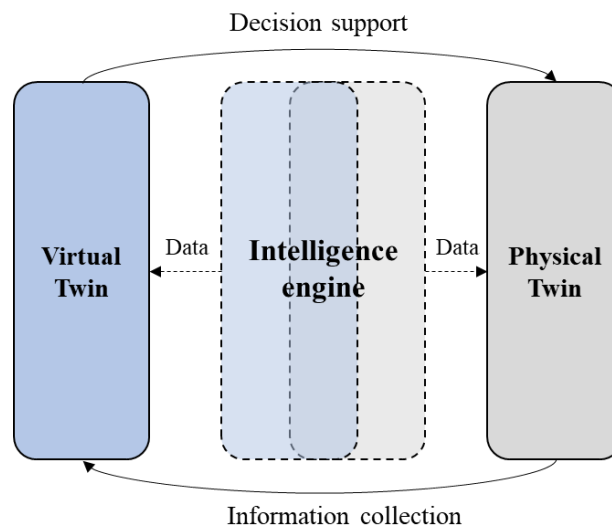


Figure 1. Schematic diagram of DT

Building assets maintenance is widely recognised as the most important integral part of FM because it is considered the most labour-intensive, time-consuming, and influential process. Effective maintenance strategies can contribute to the service lifecycle extension and maintenance cost reduction [5]. New knowledge and methodologies like DT from Industry 4.0 enabling tools are utilised in Maintenance 4.0 [6] to implement more advanced and proactive methods, such as failure prediction, fault classification,

and maintenance schedule optimisation. These new approaches can be assimilated into the PdM paradigm to address the disadvantages of the current maintenance method. PdM is a complex data-driven strategy aiming to monitor the condition states and forecast the failure events of the target asset. On the one hand, its prerequisites are real-time condition monitoring of components on the specific equipment level. On the other hand, decision-making in the PdM process is highly reliant on integrating various types of data and information with AI algorithms. Digital tools, especially the advent of DT, bring promising opportunities to create effective maintenance strategies and PdM applications because they significantly enhance data interoperability and reusability.

1.1 Background

DT is considered one effective solution of PdM in the building industry because it significantly enhances data interoperability and reusability [7]. This section endeavours to elucidate the historical trajectory and intrinsic attributes of these successive phases of the building industry, digitalisation technology, and the maintenance industry. By elucidating the evolution and intersections of these three domains, this section offers valuable insights into the integration of DT for PdM in the evolving landscape of the building industry.

1.1.1 Building Industry and Construction 4.0

The building or construction industry is often perceived as a conventional domain with limited adoption of advanced technologies [8]. Figure 2 describes the division of evolutionary phases, highlighting the transformative impact of technology on its trajectory. Beginning with Construction 1.0, characterised by rudimentary tools such as shovels and diggers, the industry relied on manual labour and basic machinery. Construction 2.0 ushered in the era of mechanisation, featuring the advent of backhoes, loaders, and cranes as pivotal technologies, greatly enhancing productivity. The transition to Construction 3.0 saw software integration, with Autodesk CAD as a representative technology, revolutionising design and project management. Currently, the industry stands at the cusp of Construction 4.0 [7], an era characterised by technology fusion. This phase incorporates ICT, high-fidelity modelling, and AI as

central components, promising unprecedented efficiency, sustainability, and innovation in construction practices.

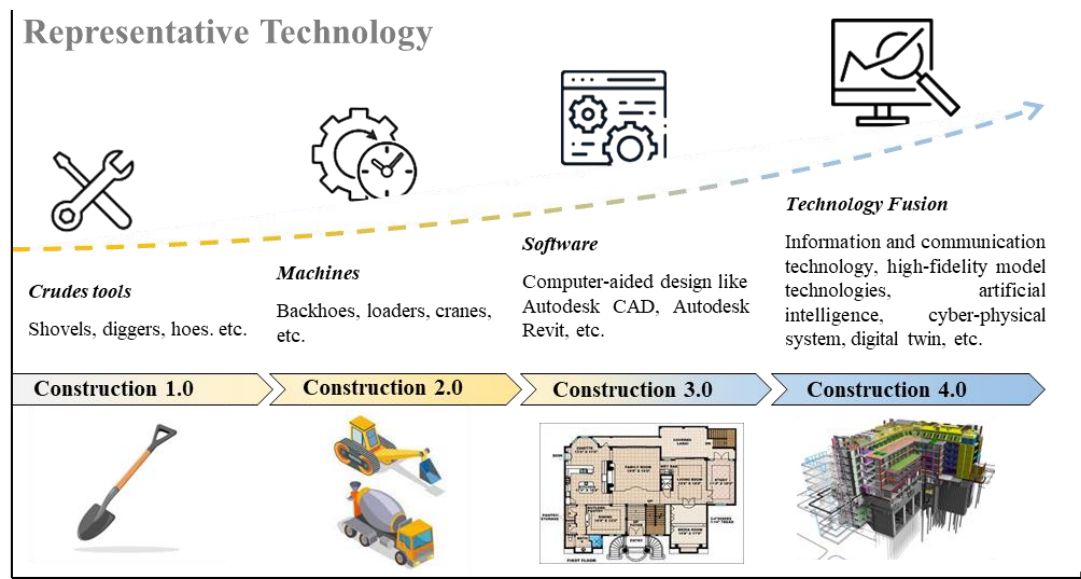


Figure 2. The evolution of the construction industry

1.1.2 Digitalisation Development

The importance of digitisation lies in its transformative capacity to enhance efficiency, productivity, and decision-making across various industries, ushering in a new era of streamlined operations and innovation [9]. The evolution of digitalisation technology is depicted in Figure 3, and each phase is marked by remarkable advancements in its core attributes and capabilities. Commencing with wireless sensor networks (WSN), digitalisation primarily focused on sensing data collection and status monitoring, enabling data collection and real-time tracking. Subsequently, the emergence of the Internet of Things (IoT) introduced a paradigm shift characterised by interconnected devices and seamless integration with cloud computing, fostering enhanced communication and data processing capabilities. The progression into the cyber-physical system (CPS) phase brought a novel dimension, emphasising the integration of cyber and physical elements while harnessing data analytics to optimise system performance and decision-making. Presently, the digitalisation landscape is anchored in the DT stage. It is notable for its high-fidelity modelling and real-time bidirectional communications, which are pivotal in the journey. Looking ahead, the Metaverse

promises immersive virtual environments and cross-platform integration, further reshaping the contours of digitalisation in multifaceted dimensions [10].

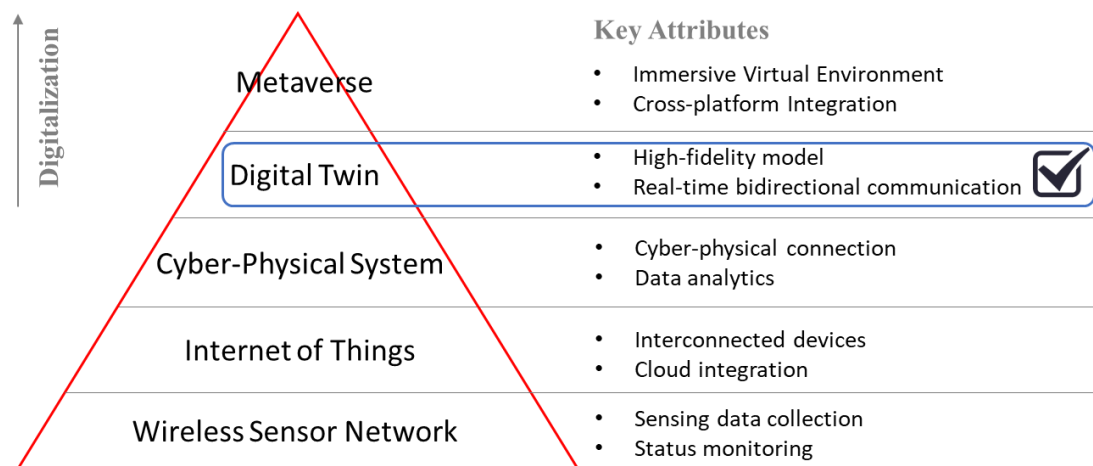


Figure 3. State-of-art digitalisation development

1.1.3 Evolution of Maintenance

Maintenance is the combination of all technical, administrative, and managerial actions during the product lifecycle to retain or restore it to a state in which it performs the required function [11], while effective maintenance plays a vital role in delivering the functionality and serviceability of assets and facilities in B&C industry. The evolution of maintenance has witnessed a transformative progression from traditional corrective maintenance practices to a proactive and predictive paradigm, and its evolution history is outlined in Figure 4. Maintenance was initially developed using the corrective approach, which addresses failure events after their occurrence and often results in operational disruptions and costly downtime. Subsequently, scheduled inspections and repairs are utilised as preventive methods to solve failures and optimise equipment lifespan. Condition-based maintenance ushered in a more refined strategy, relying on real-time monitoring systems to detect anomalies and initiate maintenance actions when necessary. Maintenance is evolving towards the PdM strategy, aiming to optimise asset reliability and operational efficiency by pre-emptively addressing issues before they manifest into critical failures.

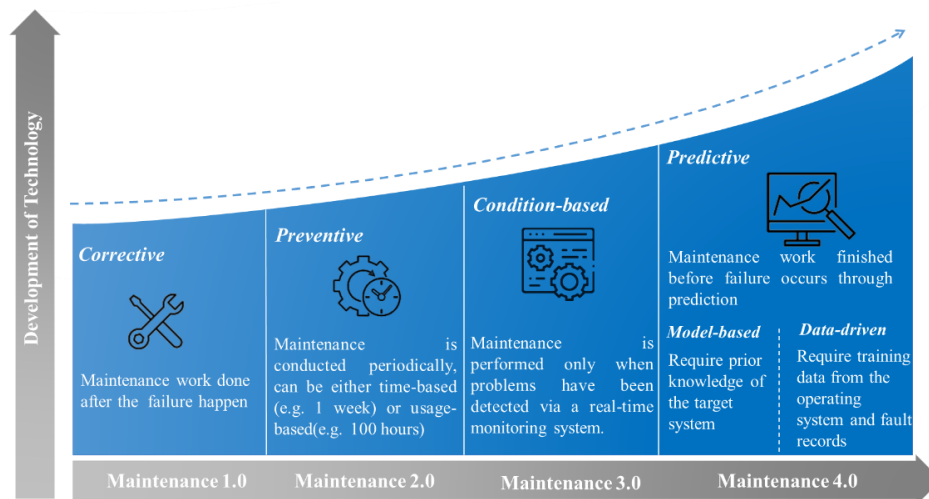


Figure 4. The evolution of the maintenance approach

1.1.4 Six M Analysis Method

The refined Six M approach, originally developed from Ishikawa’ s diagram [263] and adapted from [259], was chosen to analyse the critical determinants of successful PdM projects in the B&C sector. This methodology was selected due to its structured and comprehensive nature, which allowed for a detailed examination of the key factors—namely, Man, Machine, Material, Method, Measurement, and Management—across the project lifecycle. The Six M approach offers a systematic framework that captures the complexity of PdM projects by highlighting the interaction of technical and managerial components.

Compared to other available approaches, such as traditional fault tree analysis or root cause analysis, the Six M method provides a more holistic perspective, making it particularly suited for the multifaceted nature of PdM in the B&C industry. This adaptability is especially important in PdM contexts, where the integration of various technological and human factors plays a decisive role in project success. By using the refined Six M approach, we ensure that the analysis is not only comprehensive but also aligns with the complexities and unique challenges of the B&C sector, making it an ideal choice for this study.

1.2 Motivation

Although the building industry has developed fast in the last few decades, it is often touted as inefficient and unproductive due to the lack of cyber-physical interconnectivity [8]. As advanced B&C solutions require advanced representation and computational models to provide valuable insights and wisdom, DT has emerged as a core enabler for Construction 4.0 developments, defined as “a means of finding a coherent complementarity between the main emerging technological approaches in the construction industry” [7]. Comprising real-time multi-asset connectivity, simulation, and decision support functionalities, many recent studies have utilised Industry 4.0 technologies with DT systems to fulfil construction-specific applications. However, DT in construction differs from other industries in that it utilises Building Information Modelling (BIM) and other domain-specific protocols to meet unique stakeholder requirements. Hence, there is a need to explore the role of DT as a platform to integrate other industry 4.0 techniques holistically to form a technology foundation for Construction 4.0 and examine the benefits of holistically using DT as a platform from technical, project management, and building lifecycle perspectives.

Maintenance is widely acknowledged as the most considerable part of the lifecycle of building facilities and assets, with approximately one-third of management expenses being attributed to the maintenance process [12]. According to Manyika et al. [13], promoting predictive strategies can save up to 630 billion USD in the maintenance process during 2025. Recently, there has been a rising interest from both academic and industrial sectors towards utilising PdM in the B&C industry. The advent of digital tools and intelligent algorithms presents a promising prospect for developing strategies and applications for PdM in the B&C industry, where efficiency is frequently criticised due to the underutilisation of advanced technologies and intelligent algorithms [8]. With growing awareness of PdM implementation in the B&C area, there is a need to gain a deeper understanding of its overview of the application, methodology, research trends and gaps.

Fault detection and prediction are two critical functions for implementing PdM solutions. Detection refers to the real-time recognition of abnormal behaviour during

routine operations, while prediction aims to forecast the likelihood of failure events in a specific timeframe. Detection contributes to immediate alerts for investigation or action, whereas prediction helps in maintenance activities planning and resource allocation optimisation. Typically, higher prerequisites of data, algorithm, and virtualisation are required by prediction than detection, resulting in less failure prediction-related research. Overall, failure prediction aims to monitor the condition states and forecast the failure events of target assets, which relies on the massive real-time data and prerequisites for the meticulous and comprehensive data labelling processes. The progression of real-time data acquisition technologies and FM systems has greatly enriched the data required by intelligent approach, while data labelling requests domain expertise and a deep understanding of target assets. Therefore, diminishing the reliance on labelled datasets of PdM solutions can significantly reduce maintenance costs and improve algorithm efficiency.

Moreover, the emergency of PdM in the building industry has expanded from facilities to indoor climates, as air quality is highly relevant to residential health, comfort, and work efficiency. Modern humans' lifestyles spend more than 80% of their activities and lives within indoor environments [14], and the indoor climate is closely related to life quality and work efficiency [15]. However, most existing pollution policies and legislation are focused on the outdoor environment, while the negative impacts of interior pollution are frequently overlooked [16]. Thus, evaluating and improving indoor air quality (IAQ) has recently become an emerging problem that is raising attention in academic areas. In the IAQ field, many existing solutions have been developed for status monitoring and anomaly detection, among which IoT and AI algorithms are demonstrated as utility and efficiency solutions [17][18][19]. Therefore, leveraging the collected environmental data to achieve failure prediction with advanced algorithms and improved interaction platforms holds great promise and value.

The identified research gaps can be distilled into the following key motivations, driving the development of a DT-enabled PdM system aimed at achieving comprehensive, intelligent, and efficient maintenance and management operations in the building industry.

- **Inefficiency and underutilisation of advanced technologies in the B&C industry:** Despite fast development, the B&C industry is still regarded as inefficient due to its lack of cyber-physical interconnectivity.
- **Need for a holistic DT-enabled platform:** While DT has demonstrated value in other industries, its potential to act as a core platform that integrates Industry 4.0 techniques in the B&C industry requires further exploration.
- **Rising significance of PdM:** As maintenance costs comprise a large portion of building lifecycle expenses, there is significant potential to reduce these costs through the adoption of PdM strategies, which are currently underexplored in this industry.
- **Emerging importance of failure prediction technologies:** While fault detection has been more widely researched, the development of failure prediction methods, which are highly dependent on advanced data and algorithms, presents a critical research gap.
- **Increasing focus on indoor climate and air quality:** Given that most people spend over 80% of their time indoors, there is a growing interest in leveraging PdM for improving IAQ, which is linked to health and productivity but often overlooked in current research and policy.
- **Advancement of real-time data acquisition and AI-driven solutions:** The availability of enriched real-time data and intelligent algorithms presents an opportunity to improve PdM systems, particularly through reducing reliance on labelled datasets and improving efficiency.

1.3 Objective and Scope

The operation and maintenance (O&M) phase represents a significant portion of the lifecycle of building and civil infrastructure assets, with maintenance procedures accounting for over 30% of total management costs [12]. The primary objective of this

research is to develop a methodology for the implementation of DT-enabled PdM systems in the building industry. This thesis aims to address the challenges and research gaps through the following specific objectives:

- 1) Integrate Industry 4.0 technologies within the DT framework, emphasising the advantages of DT as a foundational platform for the B&C industry.
- 2) Investigate the role of PdM in the B&C industry, identifying key challenges and exploring future research directions.
- 3) Develop a failure prediction methodology for building facilities, utilising both labelled and unlabelled data to improve the effectiveness of PdM solutions.
- 4) Explore the potential of integrating DT with DL models to address PdM challenges related to indoor climate management.
- 5) Demonstrate how the DT framework can accelerate the adoption of PdM strategies within the built environment through the proposed framework and an applied online application.

Logical Relationship Between Technical Chapters:

Chapter 3 explores the technologies and formulates a comprehensive framework for DT integration in the B&C industry, providing the technological foundation to support the objectives outlined in Objective 1.

Chapter 4 investigates the role of PdM in the B&C industry through network analysis and highlights key research challenges, directly addressing Objective 2 and providing a basis for the development of the failure prediction methodology.

Chapter 5 proposes a Semi-GAN failure prediction model, addressing Objective 3 by focusing on predictive analytics with labelled and unlabeled data, enhancing the PdM solutions for building facilities.

Chapter 6 extends the work by integrating DL models, including LSTM and AE, for failure prediction in indoor climate management, aligning with Objectives 4 and 5. This chapter also presents a practical demonstration of the DT-enabled PdM system applied to real-world scenarios.

The chapters are structured to progressively build towards achieving the research objectives, ensuring a logical flow from conceptual development to practical implementation.

1.4 Contribution of Thesis

This thesis studies the implementation of DT-enabled PdM solutions in the B&C industry. The critical research contributions of this thesis are highlighted in Figure 5 and presented as follows:

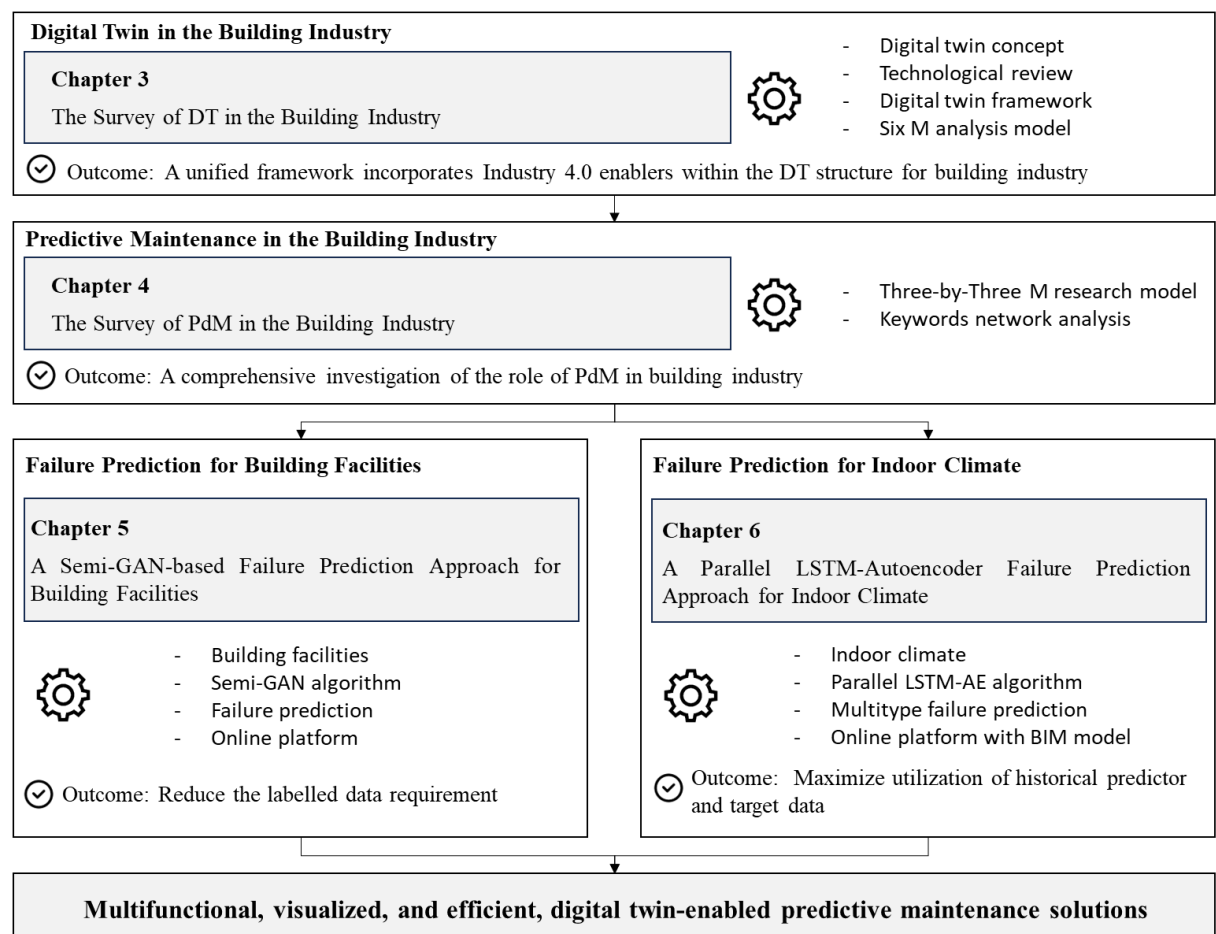


Figure 5. Research framework and contributions of the thesis

1) A unified framework for B&C has been modelled to incorporate Industry 4.0 technologies within the DT structure through a systematic literature review, which highlights how Industry 4.0 technologies can be incorporated and provides a practical guide toward deploying feasible DT-enabled solutions in the B&C industry. Besides, the Six M methodology is highlighted to ensure the scalability of construction DT systems and investigate DT's role in handling multi-asset integration and complex scenarios.

2) Keyword mapping and analyses of recent PdM studies in the building industry have been conducted, and clusters are identified as Physical-numerical PdM, Physical-intelligent PdM, Virtual-intelligent PdM, and Virtual-physical PdM according to their embodiment in function and method. Additionally, to further investigate the role of PdM in the B&C industry as well as provide more valuable insights into the current state, Ishikawa's diagram invented for quality engineering is adapted into the Three-by-Three M paradigm for in-depth analysis.

3) A novel failure prediction approach based on Semi-GAN that can incorporate unlabelled and labelled samples into the training procedure has been proposed. It can obtain the distribution and structural information inherent in unlabeled datasets, which can contribute significantly to PdM development. Less labelled data are needed for the failure prediction approach to achieve satisfactory accuracy with the information extracted from unlabelled data. Besides, the proposed approach can be embedded in the proposed 5-layer DT framework as the PdM service provided in the functional layer. The proposed method is tested and validated through a comparative analysis, including common supervised and unsupervised learning methods. Experiments have been carried out by adjusting the size of the labelled dataset and the data imbalance ratio for the training set.

4) The RUL and TS methods are employed to split the prediction problem into the combination of a supervised and an unsupervised subtask, followed by the development of a parallel prediction model integrating the LSTM and AE methods. Besides, a DT-enabled PdM framework has been proposed for indoor climates, validated through the establishment of an online platform designed to reconstruct the 3D building model and demonstrate real-time monitoring and alert information of indoor climates.

Experiments show the effectiveness of the proposed model under different warning times and fault severity through a comparison study with other benchmark models based on a practical dataset collected from different buildings in Singapore. Moreover, the practical online platform has been developed as an illustrative case for future DT-enhanced PdM solutions.

1.5 Outline of Thesis

This section provided an outline of this thesis:

Chapter 2 provides a systematic review of DT and PdM in the B&C industry from perspectives of (1) the Evolution background of the maintenance and building industry, (2) DT in the building industry, and (3) PdM in the building industry. Research gaps are identified and summarised at the end of this Chapter.

Chapter 3 formulates a comprehensive framework underpinning DT integration in the building industry and conducts an in-depth analysis employing the SixM methodology, also factoring in the construction lifecycle. Finally, it provides a forward-looking perspective on the evolution of DT in the construction landscape.

Chapter 4 extracts research clusters of PdM in the building industry based on keyword network analyses and a novel Three-by-Three M research paradigm. Then, the challenges and future directions are analysed. This Chapter aims to provide a fundamental reference for the PdM applications in the B&C industry and suggests future developments.

Chapter 5 proposes the Semi-GAN failure prediction approach for PdM of building facilities and demonstrates detailed information about model building and implementation steps. The effectiveness of the proposed method is verified through a comparison study with other machine learning (ML) algorithms. Moreover, it further explores its performance with different labelled data sizes and data imbalance ratios.

Chapter 6 presents a parallel prediction model integrating the LSTM and AE methods to predict multi-severity failure, which is implemented in real-world IAQ datasets sourced from diverse buildings in Singapore. Besides, a depiction of the proposed DT-enabled PdM framework for indoor climate is provided and verified by the online platform that showcases building models, air quality data, and alert information. This practical demonstration can provide a guiding example for future endeavours in DT-enhanced PdM solutions.

Chapter 7 concludes the key research findings of the thesis and discusses the possible directions of future works based on our study.

Chapter 2 Literature Review

The growing complexity and operational demands of the modern building industry have spurred the adoption of innovative technologies to enhance efficiency and reduce downtime. In this Chapter, we systematically review the convergence of DT technology and PdM strategies within the building industry. The existing academic publications are divided into two sections: DT in the building industry and PdM in the building industry. The Chapter is structured into three subsections: DT in the Building Industry, PdM in the Building Industry, and identifying research gaps. Finally, as PdM and DT for the building industry are emerging areas, more specific literature is provided for each research problem in the following chapters.

2.1 DT in the Building Industry

While DT solutions are increasingly adopted to boost efficiency and competitiveness, a diverse mix of Industry 4.0 technologies are used to fulfil applications specific to construction contexts [9] [20]. This section systematically reviews the existing publications concerning their application of Industry 4.0 technologies across four fundamental DT aspects: data acquisition, data processing, simulation and modelling, and decision support enablers.

2.1.1 Data Acquisition

The data acquisition process starts with raw data extraction and ends with the information being passed on to a cloud-based server or database. These technologies are highlighted in Table 1 with their corresponding construction applications and enabling tools.

WSN provides a constant input of raw data for tracking and monitoring applications, which is fundamental to the creation of DT systems in construction projects. *For structural health monitoring (SHM) applications*, Bhuiyan et al. [21] showcased a WSN-to-CPS design transition approach, highlighting the design requirements, deployment hurdles, and networking guidelines, while Loubert et al. [22] facilitated

long-distance communications within reinforced precast concrete. Yang et al. [23] developed mobile robot detection systems to perform data collection in dangerous environments. ***For building performance applications***, Zhang et al. [24] integrated heterogeneous data from different buildings into a CPS, while Lin and Cheung [25] established an environmental monitoring management system. ***For building cost efficiency applications***, Khajavi et al. [26] established an office building façade DT based on more than 25,000 sensor reading instances, while Grübel et al. [27] developed an indoor sensor network for building performance analysis. Abrol et al. [28] showcased an economical and noninvasive energy-efficient approach. ***For sustainability applications***, González et al. [29] proposed a calibration methodology to reduce the number of sensors required for building energy monitoring, while Keskin and Mengüç [30] implemented an adaptive vent system to localise and customise building thermal conditions. Liu et al. [31] [32] implemented a CPS for greenhouse gas emission monitoring to enhance sustainability efforts in prefabricated construction.

As IoT is often utilised to establish cloud-enabled systems, the critical difference between IoT and WSN lies in the use of IP-enabled connectivity (aka internet connectivity) such as IPv6-based low-power wireless personal area network (6LoWPAN) [33]. ***For building performance optimisation***, Tagliabue et al. [34] and Liu et al. [35] proposed BIM-IoT-DT integrated frameworks to support building comfort and indoor safety management. ***For project management***, Niu et al. [36] proposed a deployment framework to improve the synergies between construction entities based on a knowledge-based taxonomy. ***For SHM applications***, Zonzini et al. [37] utilised an IoT-based architecture to improve vibration engineering techniques, emphasising damage detection and task prediction.

Social media technology offers a unique approach towards enhancing data acquisition and communication techniques while incorporating social science theorems as part of socio-technical systems. Turk and Kline [38] proposed a three-tier information service framework that encompasses physical structure, DT, and social networks to facilitate construction activities.

2.1.2 Data Processing

With the huge amount of real-time heterogeneous data collected, there is a need to facilitate raw data conversion and treatment to derive meaningful information for modelling and analysis. Table 2 highlights existing enabling technologies and tools used in existing studies to tackle industry-related challenges.

Semantic modelling allows intricate relations between construction entities to be mapped and represented before storage. As such, the collected data can be processed into useful information essential to aid decision-making and transparency for value chain enhancement [39]. *For equipment design and optimisation*, Wei and Akinci [40] introduced a novel image-based registration method for panoramic images by minimising semantic segmentation errors, while Haoyu et al. [41] proposed a linear segmentation method for tunnel representation to support asset dynamic updates.

Blockchain is a distributed ledger or decentralised database of transactions recorded by a network of computers [42], which increases user confidence and trust in information reliability in construction projects. *For project management*, Lee et al. [43] utilised DT to provide secure and reliable data communications between stakeholders, while Hunhevicz et al. [44] proposed the performance-based smart contract through the integration of digital building twin and blockchain. Jiang et al. [45] developed a blockchain-enabled platform to facilitate cross-enterprise information sharing during modular integrated construction. *For sustainable practices*, Li et al. [46] developed an intelligent service platform to incorporate ICT to achieve sustainable prefabricated housing construction through a smart product-service system (Smart PSS) approach.

Data mining serves as an extraction and pattern discovery process within large datasets and involves ML approaches, statistics, and database systems [47]. *For building performance optimisation*, Schmidt et al. [48] integrated CPS and cross-industry standard processes for data mining through a generic model-based design methodology. *For project management applications*, Pan and Zhang et al. [49] deployed a data mining-driven DT system using fuzzy miner and ARIMAX to identify potential bottlenecks and reallocate resources dynamically.

2.1.3 Modelling and Simulation

DT technologies rely on 3D high-fidelity models and simulations to provide comprehensive visualisation for evaluating specific scenarios and verifying automatically computed solutions while complimenting other construction-related enabling technologies highlighted in Table 3.

Utilising BIM technology, stakeholders can model building designs with dynamic optimisation and lifecycle consideration based on the various parameters setting [50]. This portion maps various BIM-DT solutions to core construction aspects and highlights enabling techniques used. *For FM*, Desogus et al. [51] and Wernerová et al. [52] utilised cloud-based BIM to enable building management capabilities. Adibfar and Costin. [53] developed a dynamic DT for the bridge through integrating real-time traffic data. Emphasising as-built models, Nicola Moretti et al. [54], Rausch and Haas [55] automated asset management processes to support cognitive buildings throughout various lifecycle aspects, while Huynh and Nguyen-Ky [56] proposed a cross-platform system to visualise data and manage comfort levels. Wang et al. [57] achieved the interaction and virtualisation of various processes during building construction. Torrecilla-García et al. [58] proposed a BIM-enhanced decision support approach for safety management in the building industry. *For SHM*, Lei et al. [59] developed a disaster prevention platform which identifies the building's structural state, while Yuan et al. [60] evaluated the integrity of temporary structures. Taraben and Morgenthal [61] used voxel-based methods to discretise acquired 3D geometries for building damage propagation. *For asset design and optimisation applications*, Al-Saeed et al. [62] and Schimanski et al. [63] implemented automation solutions within the construction manufacturing domain based on lean manufacturing paradigms and configure-to-order services. Kosse et al. [64] developed a DT framework to optimise the modularised construction of precast concrete. *For sustainability applications*, Agostinelli et al. [65], Kaewunruen et al. [66], Zhao et al. [67], and Banfi et al. [68] achieved the visualisation and assessment of Net Zero Energy Building (NZEB) solutions. Xing et al. [69] applied the product-service relationship to manage stakeholders involved in the various lifecycle stages to derive higher eco-efficiency with reduced material consumption and waste generation. Kaewunruen and Lian [70] established a 6D BIM for the railway

turnout system, which enables the assessment of schedule, cost, and sustainability and achieves a balance.

Simulation provides core functionalities, including but not limited to scenario visualisation and solution verification. *For structure design optimisation*, Lydon et al. [71] presented a coupled simulation approach to optimise the thermal design of a lightweight roof, while Kyvelou et al. [72] described the numerical simulation method for the verification and assessment of bridge structure. *For building performance optimisation*, Lilis et al. [73] proposed a discrete event simulation (DES)--based system to virtualise intelligent buildings via a scalable architecture.

Following that, the point cloud offers an efficient approach to map these virtual models via 3D scanning and photogrammetry software. *For SHM*, Omer et al. [74] digitised a typical masonry bridge in VR space as an alternative to traditional inspection methods, while Maroc et al. [75] proposed a novel method to transfer point cloud into parametric models for the historic masonry buildings detection. *For asset design and visualisation*, Xue et al. [76] processed urban LiDAR point clouds based on the object cross-sections, while Pantoja-Rosero et al. [77] automatically reconstructed the LOD3 models for existing buildings through an ML-based segmentation method. To model assets and structures, Jiang et al. [78] established the DT model of existing highway assets from map data, while Münzinger et al. [79] reconstructed tree models in the 3D city view.

Virtual/ Augmented reality (VR/AR) offers an immersive and interactive approach to engage with new tools and explore high-risk environments. *For human-robot collaborative work*, Wang et al. [80] established a remote collaborative system with an intuitive VR interface, which enables real-time bidirectional communication and supervision between workers and construction robots. *For urban planning and design*, Kikuchi et al. [81] integrated AR and drones into the detailed 3D model to achieve the city landscape visualisation, which allows non-expert users to understand and participate in the construction project.

2.1.4 Decision Support Enablers

To enable disruption management capabilities and facilitate lifecycle transition, construction systems rely on decision support functionalities such as semantic solution generation, which are established using tools and techniques highlighted in Table 4.

Computer vision (CV) enables the derivation of meaningful information from visual inputs to facilitate solution generation. *For the bridge maintenance system*, Shim et al. [82] combined both maintenance information and digital inspection systems to generate reliable decision-making to enhance the bridge maintenance process. *For FM*, Antonino et al. [83] utilised an image recognition module to detect user's movements. Lu et al. [84] proposed an image-processing approach to reconstruct 3D models from CAD drawings, while Pang and Biljecki [85] achieved the 3D reconstruction through the proposed image-to-mesh approach and street view images.

ML involves the use of algorithms that utilise historical data and experience input to perform predictions and solution optimisations for decision support systems [86]. *For urban management*, Döllner [87] developed an ML/DL Geospatial analytics engine to derive domain or application-specific semantics. *For improving energy efficiency*, Alanne and Sierla [88] concluded that ML-equipped building-integrated energy systems with adaptability for unpredicted changes and Konstantakopoulos et al. [89] encouraged energy-efficient behaviour by facilitating the human-building interaction. Austin et al. [90] proposed the DT-based energy-saving architecture for smart cities. *For safety management*, Kamari and Ham [91] and Liu et al. [92] presented risk assessment frameworks for disaster preparedness and risk control through a DT system, while Gichane et al. [93] developed an elevator security DT system utilising a YOLOV3 algorithm. Pan et al. [94] proposed an AI-based segmentation method to capture and recognise important electrical and fire-safety entities. *For construction equipment monitoring*, Zhang et al. [95] developed a construction equipment recognition algorithm to facilitate asset performance evaluation. *For building performance optimisation*, Lv et al. [96] designed an AI-driven CPS to support indoor environment management via temperature response and control *for on-site construction optimisation*. Saini et al. [97] proposed an action planning system to optimise and automate the operation for the well construction, while Tariq et al. [98]

optimised the design of solar chimneys, considering energy efficiency and environmental factors. *For structure design optimisation*, Fernández-Cabán et al. [99] utilised a stochastic optimisation algorithm to support sustainable tall building design based on occupant comfort and building drift, and Abdelaziz and Hobeck [100] developed an optimal controller to reduce the vibration caused by wind.

Table 1. Data acquisition

Technology	References	Construction applications	Tools/ Techniques
WSN	[21][22][23]	Enhance SHM via CPS.	Connectivity: Zigbee; LoRa WAN; Bluetooth Controller: Raspberry Pi; Arm; Arduino Sensors:
	[24] [25]	Building performance evaluation. Include evacuation planning, monitoring energy usage, emissions (CO, radiological), and temperature.	1) SHM: Acceleration; Piezoelectric; Ultrasonic; Radar; Laser; Strain gages; Optical fibre; Vision; Corrosion; Linear voltage displacement transducers; Inclinometers; Footprint accelerometer; Inertial measurement unit (IMU); Meteorological
	[26][27][28]	Improve building cost efficiency through lifecycle management and energy conservation.	2) Building performance: Gas concentration; Temperature; Humidity; Hazard detecting; Cloud-Oriented Radiation; Radiological measurement device; CORSAIR; Occupancy 3) Building cost efficiency: Temperature; Passive infrared (PIR)
IoT	[29][30][31][32]	Develop sustainability practices. Include applications in HVAC systems, reduced energy usage, carbon emission monitoring, equipment, and raw material tracking.	4) Sustainability: Electrical power; Airspeed; Heat flux; Humidity; Wind speed; Wind direction; Radiation; Temperature; Acceleration; Barometric; GPS; Inductive displacement; Soil moisture; Asphalt strain; Horizontal inclinometer; RFID
	[34] [35]	Building performance optimisation. Include energy efficiency, sustainability assessment, indoor safety management, and enhanced FM system in the building lifecycle management (BLM) process.	Connectivity: Cellular networks (GSM/3G/4G); Wi-Fi; Universal mobile telecommunications system; Low-power wide-area network (LPWAB) Controller: Arduino; Programmable logic controller (PLC) IoT Sensor:
	[36]	Project management. Integrate CPS/ DT technologies to enhance efficiency and synergy.	1) Building performance: Temperature; Humidity; Smoke; Oxygen Concentration; Carbon monoxide concentration; Gate magnetic; Infrared
Social Media	[37]	Enhance SHM. Include PdM solution of infrastructure.	2) Project management: Ultrasonic, RFID tag, GPS, Load cells, Switch, Barometers, Accelerometer 3) SHM: Accelerometer, IMU Auxiliary tools: NodeJS; TICK stack; Grafana; Blockchain
	[38]	Enhance construction lifecycle management. Include plan, design, build, and usage aspects.	Auxiliary milieu: Log files, Emails, Social media messages, Building models

Table 2. Data processing

Technology	References	Construction applications	Tools/ Techniques
Semantic Modelling	[40] [41]	Asset design and optimisation. Enable equipment re-/configurations for disruption management. Localise a panorama with a sub-meter localisation error. Improve asset representation.	Software: Apache Jena, Protégé, Revit, Unreal Engine 4, DataSmith Language: XML, OWL, SPARQL, C++ Library: OpenCascade, OpenVDB Algorithm: Convolutional neural network (CNN), ResNet101, PSPNet
Blockchain	[43] [44] [45] [46]	Project management. Improve efficiency via contract implementation and stakeholder collaboration with increased reliability and service. Enhance information sharing and continuity for Modular Integrated Construction. Develop sustainable practices. Develop an intelligent platform integrating with blockchain to improve the sustainability of prefabricated housing construction.	Cloud platform: Microsoft Azure Database: Distributed ledger Mechanism: Consensus mechanism, Encryption mechanism Platform: Ethereum blockchain
Data Mining	[48] [49]	Building performance optimisation. Improve the energy efficiency of both legacy and modern buildings. Project management. Achieve a higher degree of intelligence and automation.	Algorithm: Inductive miner, fuzzy miner, ARIMAX mode Modelling languages: Petri net, business process modelling notation (BPMN); Standard: Cross-industry standard process for data mining (CRISP-DM), Model: CRISP-DM reference model

Table 3. Modelling & Simulation

Technology	References	Construction applications	Tools/ Techniques
BIM	[51] [52] [53] [54] [55] [56] [57] [58]	FM. Improve comfort, energy efficiency, and BLM. Include anomaly detection, maintenance work, and decision support systems.	BIM authoring tools: Autodesk Revit, ArchiCAD, Allplan, AECOSim, Tekla structures BIM auxiliary tools: BIMserver, Autodesk Navisworks, Revit DB Link, Dynamo BEM authoring tools: Green Building Studio, EnergyPlus, Design Builder, Open Studio, CYPETHERM HE;
	[59] [60] [61]	Enhance SHM. Include disaster planning and damage inspection.	
	[62] [63] [64]	Asset design and optimisation. Incorporate lean manufacturing and configure-to-order business approaches to automate construction-related productions. Optimise precast elements production.	
	[65] [66] [67] [69] [70]	Develop sustainability practices. To realise NZEB solutions, circular economy, carbon cost estimation, and other green initiatives via product-service paradigms, lifecycle considerations (BEM), and 6D BIM adoption.	
Simulation	[71] [72]	Structure design optimisation. To reduce prototype development time and cost through high-resolution analysis and parametric geometric modelling.	Coupled simulation: ANSYS Fluent, TRNSYS, MATLAB Numerical simulation: ABAQUS DES simulation: coroutines, open BMS, ZeroMQ library
	[73]	Building performance optimisation. Enable infrastructure visualisations for power and environment monitoring.	
Point cloud	[74] [75]	Structure health monitoring. Inspection services for digitised structures in a VR environment, Future damage validation for historic masonry structures.	Software: Cyclone register 360, Cloud Compare, Civil 3D Hardware: Stationary/Airborne/ terrestrial Laser scanner, Leica ScanStation P40, Leica ScanStation P20 Library: Point cloud library, ODAS library
	[76] [77] [78] [79]	Asset design and visualisation. Generates building and city models using LiDAR, gestalt design principles, and as-built reconstruction approaches. Include ML/DL-based interpretation of point clouds to classify models.	
VR/AR	[80]	Human-robot collaboration. Facilitates task planning and supervision through bidirectional communication and asset control.	
	[81]	Urban planning and design. Multiple viewpoints and usability testing from nonexpert stakeholders involved in the building project.	

Table 4. Decision support enablers

Technology	References	Construction applications	Tools/ Techniques
CV	[82]	Bridge maintenance system. Includes image recognition to enhance inspection processes.	Algorithm: Mask R-CNN, DeepSORT, Self-designed localisation, Fuzzy Logic, Edge detection, Neuro-fuzzy system, Optical Character Recognition, DeepLabv3 Software: Self-designed Revit, Blender
	[83] [84] [85]	FM. This includes movement recognition for maintenance operations, 3D structure reconstruction from CAD drawings, and street-view images.	
ML	[87]	Urban management. Contribute to efficiently building and maintaining base data for geospatial DT, including virtual 3D cities, building indoor models, and BIM.	Algorithm: Tree-based classification, Clustering, Association, Categorizing, YOLOV3, SVM models, GA Network structure: PointNet neural network (PNN), CNN, Deep bi-directional Recurrent Neural Networks (DBRNN), LSTM, Back-propagation neural network (BPNN), Deep Residual Networks (DRN), Iterative Closest Point (ICP), Random sample consensus (RANSAC), KPConv, Monte Carlo tree search (MCTS), Multivariate regression mode, Non-dominated sorting GA
	[88] [89] [90]	Improve energy efficiency. Include energy management through interacting with occupants, smart building design, and integrating semantic models.	
	[91] [93] [94]	Safety management. Develop a security system for a three-floor elevator in a commercial building setting and an indoor safety management system based on DT. Propose a threat assessment framework for construction sites. Identifying essential entities from the electrical and fire-safety domain.	
	[95]	Construction equipment monitoring. Evaluate asset performance in various conditions.	
	[96]	Building performance optimisation. Integrate with CPS in a building environment and provide theoretical information and practical reference for developing the indoor environmental control system.	
	[97] [98]	On-site construction optimisation. Improve construction workflow schedule and optimise the structure of building components.	
	[99] [100]	Structure design optimisation. Support structure evaluation dynamically and minimise wind-induced vibration.	

2.1.5 Computation Enabler

The computational aspect of the DT system is employed for data analytics and solution generation, covering the optimisation models and facilitating intelligent decision-making service [101]. As illustrated in Table 5, most research focuses on building facilities due to their similarity to manufacturing processes, which have been extensively studied. Conversely, research on structural and environmental applications is less prevalent, as they represent unique scenarios in the building industry. Extracting practical knowledge from heterogeneous data is challenging, making it essential to identify appropriate methodologies and tools for achieving and optimising the target function, which is crucial to DT construction.

Table 5. Computation enabler

	Structure	Facilities	Environment
Computation	<ul style="list-style-type: none"> • Robust and automated façade segmentation using point cloud data [102]; • Four-stage holistic assessment for maintenance planning [103]; • Extended DT creation in building life cycle management [26]; • Classic ML model comparison study on PdM performance on track structure [104]; • AI clustering and hazard index calculation-based decision support techniques for road maintenance [105]. 	<ul style="list-style-type: none"> • Operation errors removal through feature extraction and ML techniques [106]; • Ensemble ML model for malfunction expedite detection [107]; • Automated risk detection and mitigation, including obstacles and human activity [108]; • Maintenance schedule prioritisation through hidden risk analysis [109]; • Systemic empirical knowledge-enhanced neural network model for automatic failure management [110]; • Reinforcement learning approach improving maintenance schedule based on tack inspection [111]; • RNN-based surrogate model for dynamic maintenance optimisation [112]; • Explainable unsupervised deep learning model for failure detection [113]. 	<ul style="list-style-type: none"> • Predict and optimise occupant's comfort by combining statistic and ML models [114]; • AI-driven reasoning and service process for automatic building management system [115]; • Ensemble classifier learning algorithm tailored for imbalance scenarios during three-stage CBM [116]; • Ensemble failure prediction algorithm for indoor air quality [117]; • Combination of clustering ML and rule-based strategies for malfunction reduction [118].

Classical ML models, such as clustering and regression algorithms, are commonly applied to fundamental tasks like occupants' comfort improvement [114], building management automation [115][26], and maintenance scheduling [105][104]. Furthermore, advanced functions utilised deep learning methods, including explainable unsupervised models [113], RNN-based methods [112], and reinforcement learning [111]. Besides, M. Jasiulewicz-Kaczmarek and A. Gola [106] enhanced the feature extraction process to reduce the operation errors, while Lin et al. [116] developed an ensemble learning method tailored for imbalance. Failure prediction commonly demands higher algorithmic and analytical ability than management optimisation

because of its real-time reliability requirements. Existing studies on failure reduction are mainly implemented on facilities [107][108]and environment [117][118], with structural-related studies focusing more on the management optimisation level [102]. Integrating the data-driven and model-driven methods enhances the failure management performance. For example, Weerapura et al. [110] incorporated empirical knowledge into the neural network design, while S. Agostinelli and F. Cumo [118] combined ML solutions and rule-based strategies. Employing robust and automated computational algorithms in building maintenance and management ensures comprehensive evaluation, improves predictive accuracy, and supports proactive decision-making.

2.2 PdM in Building Industry

PdM has emerged as a critical strategy in the building industry to enhance the efficiency, reliability, and sustainability of building facilities and the environment. This part explores the existing research on PdM in the building industry, offering a comprehensive analysis of the PdM framework, application, and applications.

2.2.1 PdM Framework

The conceptualisation and establishment of PdM frameworks for building assets is the cornerstone of system implementation, and it has been studied by many researchers with different concerns. Recent technological developments in Construction 4.0, like BIM, IoT, and ML, have shown great potential to facilitate PdM implementation and improve FM efficiency in the building industry. Shalabi and Turkan [119] proposed a BIM-based framework that integrated the output data from different building management systems to detect energy overconsumption. Also, Villa et al. [120] and Cheng et al. [121] integrated IoT systems and BIM models into the PdM framework of building facilities and validated the approach feasibility through fault detection and maintenance planning for mechanical, electrical and plumbing (MEP) components in buildings. ML and big data mainly serve as forecast functions during PdM. The data-driven PdM framework considering distributed computing and ML has been designed and examined by Katona and Panfilov [122] and Bouabdallaoui et al. [123] to strengthen the decision support model during PdM implementation. Santiago et al. [124]

presented a big data service architecture for the failure inspection and prediction of the heating, ventilation, and air conditioning (HVAC) system, while Razali et al. [125] proposed a novel big data procedure for PdM, which considers both well-structured statistical data and non-visible data like web search queries and social media.

2.2.2 PdM Methodology

1) Model-based PdM. Building facilities are constantly deteriorating as service time increases, which makes deterioration analysis and RUL estimation become the traditional focus during PdM implementation. To establish a foundational comprehension of the operational paradigms for critical building facilities, Chandra et al. [126] conducted an in-depth examination of failure mechanisms and statistics for vital building assets, encompassing HVAC, power distribution, elevators, and lighting systems, while Chen et al. [127] incorporated reliability analysis technology and scheduling optimisation method to develop an expert system model for the maintenance and management of building facilities. To overcome the subjectivity and uncertainties associated with the current methods, Ahmed et al. [128] proposed a hybrid approach combining genetic algorithms (GA) and fuzzy Markovian models to develop an effective stochastic deterioration prediction model, which incorporates recently observed conditions to enhance the accuracy of the forecasting procedure. Most existing PdM implementations for building assets concentrate on the HVAC systems because of their substantial importance and representativeness. RUL estimation is convenient to achieve with some mathematical approaches like the Generalized Likelihood Ratio procedure [129] and Monte Carlo-based simulation techniques [130]. Utilising the data mining approach, Yang et al. [131] developed a failure mode and effects analysis method utilising word-order documents to achieve the failure prediction and root exploration, while Gálvez et al. [132] proposed a hybrid approach integrating the physical-based and data-driven models to solve the data lacking problem during the PdM for an HVAC system.

2) Data-driven PdM. PdM is a multifaceted, data-centric procedure requiring the amount of monitoring data extracted from both the asset side and maintenance process, which emphasises the indispensable role of employing ML techniques to enhance its effectiveness and utility. Table 6 provides an overview of ML-driven PdM methods

applied in the building industry, and the emphasis of existing publications is the MEP facilities like HVAC and elevator systems, which are crucial elements to ensure building performance. These publications are categorised by different PdM functions, including fault detection, fault prognostics, condition monitoring, and strategies optimisation.

Table 6. ML methods in PdM in building facilities

Target	Ref	Approach	Data source
<i>Fault detection</i>			
HVAC system	[132]	Adaptive boost	Real sensor system, soft sensor system
	[133]	Variational AE	Sensor system
	[122]	Gaussian model	Sensor & actuator system
	[134]	Logistic regression (LR), random forest (RF)	Sensor system
	[135]	Regression tree (RT) models	Sensor & actuator system, work-order database
	[124]	RF	Sensor system
	[136]	Decision trees, support vector machine (SVM)	Sensor system
	[137]	Grey model, autoregressive model	Quick access recorder
Traction elevators	[138]	AE model, RF	Sensor system
<i>Fault prognostics</i>			
Lift	[139]	RF	Sensor system, maintenance records
	[140]	Artificial neural network (ANN)	Sensor system
Multiple	[141]	Hierarchical extended k-nearest neighbours (KNN)	Multiple distributed slave meters
HVAC system	[123]	LSTM	Building automation system, IoT device
Complaints	[142]	Multi-layer perceptron (MLP) models	Building operator records, meteorological records
<i>Condition monitoring</i>			
Healthcare facilities	[128]	GA, fuzzy logic,	Historical datasets
MEP components	[121]	ANN, SVM	BIM, IoT device, FM system
Grids	[143]	KNN	Sensor system
Multiple	[144]	Backpropagation MLP	Maintenance records
<i>Strategies optimisation</i>			
Healthcare facilities	[145]	GA, K-means clustering algorithms	Maintenance records
	[146]	Decision trees, KNN, Naïve Bayes	Expert surveys
Concrete roof	[147]	Fuzzy logic	Literature surveys

2.2.3 PdM Application

1) Building facility. The building facilities garnered the most attention from the implementation of PdM in the construction industry because of its notable similarities

with the mechanical equipment in the manufacturing industry. Villa et al. [120] and Cheng et al. [121] demonstrated the integration of IoT systems and BIM models into a PdM framework for building facilities and achieving the fault detection and maintenance schedule. Also, Chandra et al. [126] explored failure mechanisms and statistics for critical building assets such as HVAC, power distribution, elevators, and lighting systems, while Chen et al. [127] utilised reliability analysis and scheduling optimisation to establish an expert system for building facilities management. HVAC stands as the most representative facility, with a significant research focus within the PdM-building field. Santiago et al. [124] introduced a big data service architecture aimed at inspecting and predicting failures in HVAC systems. Concurrently, Gálvez et al. [132] suggested a hybrid approach that combines physical-based and data-driven models to address data scarcity issues in the PdM of HVAC systems. Besides, Xia et al. [148] proposed a multi-layer self-attention method based on LSTM to improve the accuracy and robustness of the current RUL estimation approach, which contributes to the reduction of maintenance costs and unexpected failures.

2) Indoor climate. PdM for indoor environments is a proactive approach that leverages data analysis, sensor technology, and advanced algorithms to optimise the maintenance and management of indoor conditions, including climate, lighting, energy, etc. Kapoor et al. [16] systematically discuss IAQ parameters in existing studies relevant to Indian school classrooms over the past decade. Moreover, Wei et al. [149] developed the AE with LSTM cells for anomaly detection tasks in indoor climates and validated the proposed method through a practical dataset collected in New Zealand, while Li and Cai [150] introduced an innovative demand-controlled ventilation strategy aiming to contain COVID-19 contagion in the indoor environment while saving energy consumption. To facilitate the lighting and temperature control in shared and open offices, Jung et al. [151] collected temperature, humidity, and brightness information using an IoT device to train an LSTM model for anomaly analysis in shared and open offices. For energy consumption maintenance, Himeur et al. [152] reviewed existing AI-enabled anomaly detection frameworks for building energy consumption and introduced a comprehensive taxonomy based on adopted modules and parameters for the analysis of existing algorithms. Besides, deep anomaly detection is of great significance in the optimisation of the equilibrium between building energy production and demand, and Malki et al. [153] investigated the potential to integrate ML-based

anomaly detection approaches to catch trends and changes at an early stage and improve the maintenance of the energy systems as well as Copiaco et al. [154] proposed a novel supervised deep transfer learning approach using 2D image representation as input features, which provides more possibilities to encode data detect pertinent features than traditional 1D energy time series dataset. Vertical green living walls (VGW) are recognised as a nature-centric strategy to enhance air quality and enhance comfort within the construction sector. Liu et al. [17] explored the possibility of deploying ML-based anomaly detection methods to vertical plant wall systems to enhance automation and improve intelligence to realise the PdM of the indoor climate. Besides, Helman et al. [155] developed a comprehensive high-throughput monitoring system utilising remote sensing technology in a VGW workplace, which integrates multiple data acquirement systems such as sensor systems, cameras, gas-exchange measurements, and AI algorithms into an analytical and visualisation of an online platform.

2.3 Research Gaps

The growing interest in PdM implementations in the building industry indicates that researchers and stakeholders are realising the potential benefits associated with proactive FM strategies. Although current studies have made commendable strides in exploring PdM applications in this field, it is essential to acknowledge the presence of notable research gaps that warrant further investigation. These gaps encompass multifaceted aspects of PdM implementation, ranging from developing comprehensive PdM frameworks tailored to the building sector's unique requirements to refining methodologies, such as data analysis and predictive modelling, that can effectively address the distinctive challenges posed by building infrastructure. Key research gaps are summarised as follows:

1) Integration of Industry 4.0 Technologies with DT:

Pros: Numerous studies have successfully integrated Industry 4.0 technologies (e.g., IoT, AI, and robotics) with DT systems for construction-specific applications, demonstrating the potential for enhanced data collection, monitoring, and automation.

Cons: However, many of these studies focus on isolated use cases rather than a unified framework that combines technical aspects, project management, and the building

lifecycle. There is a gap in developing a holistic DT framework that seamlessly integrates Industry 4.0 technologies to optimise building operations and long-term asset management. The proposed research seeks to address this by offering a unified DT-enabled PdM system that enhances operational efficiency and decision-making.

2) Application and Methodology in PdM for the Building and Construction Sector:

Pros: Existing research highlights the effectiveness of PdM in various industries, particularly manufacturing, where data-driven methods have been widely adopted.

Cons: The B&C sector remains underexplored, with limited studies providing a comprehensive overview of PdM applications, methodologies, and trends tailored to the unique needs of buildings. The current literature often lacks a critical synthesis of how PdM can be effectively adapted for complex building infrastructures. The proposed research aims to fill this gap by offering a tailored PdM methodology specific to the building industry's requirements, including detailed modelling that addresses these complexities.

3) Failure Prediction and Data Labelling Issues:

Pros: Real-time monitoring and fault detection systems have been well-researched, contributing valuable insights into asset health and performance monitoring.

Cons: However, most studies neglect the predictive aspect of maintenance, which limits the ability of building operators to plan for future maintenance needs. Moreover, the challenge of dealing with unlabelled or partially labelled data remains a significant obstacle, as data labelling requires considerable domain expertise and is resource intensive. The gap lies in developing robust failure prediction models that can effectively handle both labelled and unlabelled data. The proposed research will focus on addressing this gap by exploring machine learning algorithms that enhance failure prediction accuracy, reducing the reliance on fully labelled data.

4) Development of DT-enhanced and integrated PdM Systems:

Pros: The use of IoT technologies for real-time monitoring has laid a strong foundation for PdM applications, providing detailed insights into asset health. **Cons:** Despite these advancements, the potential of DTs remains underutilised in PdM research. While IoT enables data collection, DTs offer dynamic, real-time simulations and predictions

across the lifecycle of building components. The proposed research will focus on fully leveraging DTs to enhance PdM systems, providing more comprehensive predictive insights, optimising resource allocation, and improving decision-making capabilities throughout the building lifecycle.

4) PdM for Indoor Climate:

Pros: Current PdM solutions have proven effective in facility-related applications (e.g., HVAC, lighting, elevators), ensuring the smooth functioning of essential building systems. **Cons:** Despite the importance of indoor environmental quality, relatively little attention has been given to PdM for indoor climate systems, which directly impact occupant comfort, health, and building sustainability. This gap presents an opportunity to extend PdM frameworks to address indoor environments. The proposed research will incorporate indoor climate considerations into the PdM system, focusing on predicting failures in heating, cooling, and ventilation systems to improve energy efficiency and occupant well-being.

2.4 Research Opportunities and Proposed Model

To address these research gaps, the proposed study will focus on the following modelling approaches:

- **A Unified DT Framework:** This approach will integrate Industry 4.0 technologies, such as IoT, AI, and big data analytics, within a DT architecture. The framework will not only enhance operational efficiency but also incorporate project management and lifecycle considerations, providing a holistic solution for long-term asset management and decision-making in the building industry.
- **Advanced Predictive Models:** The study will focus on developing sophisticated machine learning models capable of processing both labelled and unlabelled data. These models will improve the accuracy of failure prediction, allowing for more proactive maintenance strategies while minimising the labour-intensive task of manual data labelling.

- **PdM for Indoor Climate:** Recognizing the gap in current PdM applications, the study will extend PdM systems to address indoor climate management. This will include the development of models that ensure environmental sustainability, energy efficiency, and occupant comfort, aligning with the growing emphasis on green building practices and occupant well-being.
- **DT-enhanced PdM Solutions:** The study will explore the untapped potential of DT technology in PdM. By leveraging real-time monitoring and 3D visualisation, the research will optimise maintenance processes and long-term asset management, offering a dynamic and responsive approach to building system management that enhances resilience and operational performance.

In summary, while the existing literature has made notable contributions to PdM implementations in the building industry, several critical gaps remain. These include the need for a unified DT framework integrating Industry 4.0 technologies, a deeper understanding of PdM specific to the B&C sector, advanced failure prediction models that handle labelled and unlabelled data, the inclusion of indoor climate considerations, and the full exploration of DT's potential in PdM solutions. The proposed research aims to address these gaps by developing innovative models and frameworks that push the boundaries of current PdM practices, offering significant advancements in building management, sustainability, and operational efficiency.

Chapter 3 The Survey of DT in Building Industry

This chapter presents a comprehensive investigation of the DT paradigm within the building domain. Firstly, we systematically review the convergence of DT technology and Industry 4.0 enablers within the building industry. Subsequently, the conceptual framework underpinning DT integration in construction is proposed based on the previous systematic survey of existing studies. Moreover, the author employs the Six M methodology to analyse the DT-related studies considering the construction lifecycle. Finally, Theoretical insights and the future trajectory of DT's deployment, performance, and implications are unearthed through this meticulous analysis, offering

a pragmatic roadmap for harnessing the full potential of DT to revolutionise the building industry.

3.1 Integration of Technologies using a DT-adapted Framework

The range of core enabling tools and techniques presented earlier varies in technological advancement and implementation difficulty. To frame these co-construction entities within a Construction 4.0 context, a unified DT-adapted architecture is proposed. Through the five-layered technology stack, as shown in Figure 6, each entity fits into the overall structure towards enabling smart construction applications. This architecture leverages previously featured layered DT hierarchies [31], [25] to provide an overview of the technological landscape for enhancing the B&C environment.

Starting with the cyber-physical layer, assets and resources are digitalised through a systematic approach to a virtual space, emphasising multi-source data acquisition, real-time two-way connectivity for asset monitoring and control, and cyber-physical information exchange. Within the physical component, raw data is acquired from construction assets and resources (E.g., working site, materials, equipment, and workforce) via a range of spatially dispersed sensors and communication devices, as shown in the data collection component. Following industrial communication protocols such as the OSI standard, site activity monitoring, and resource tracking can provide additional value to stakeholders regarding safety, productivity, and quality assurance. Additionally, instructions can be passed down to the physical entities using remote control and actuators to implement solutions under user supervision. Meanwhile, updated contextual heterogeneous data is mapped onto respective cyber entities through techniques such as point cloud mapping and BIM modelling. In the cyber component, common tools to facilitate data representation include BIM, simulation, point cloud, and high-fidelity 3D models. Depending on stakeholder/ project requirements, the type of simulation technology selected may vary between dynamic, discrete event, and agent-based to reflect construction activities accurately [156], whereas the use of point cloud models often requires a LiDAR system setup for real-time mapping. With mobile

robots being increasingly utilised in detection systems, information from dangerous environments can be safely collected, while structural and health monitoring activities can be automated through sensor networks and IoT systems [23].

Next, the data processing and computational layer comprise data treatment, storage and retrieval, and analytical processing modules to convert raw data into useful information and, subsequently, knowledge based on a systematic approach. While high-frequency data collection enables accurate analysis due to lower estimation variance, data cleaning and filtering processes are crucial towards ensuring quality input due to the higher probability of data redundancy and inaccurate recordings. Data conversion ensures format compatibility and standardises information flow into the subsequent computational modules, whereas data fusion enables multiple data sources to be integrated to generate consistent, useful, and accurate information. Semantic modelling refers to the method in which information is organised and allows reasoning models to infer useful knowledge based on the networks and relationships established between data nodes [39]. Following that, the processed data are stored in remote servers which host the databases and data lake. The selection of database types depends on the intended application and is highlighted previously in Section 3.1.2. To derive meaningful knowledge, the computation layer draws on multimedia processing, data mining, cloud/ edge computing, algorithms, and ML approaches to generate insights and analysis for use in the functional layer.

Subsequently, the functional layer identifies core construction applications featured in existing studies and contains domain-specific knowledge such as the various ISO protocols for safety and sustainability, government regulations and guidelines, and stakeholder-centric preferences and requirements. Relevant knowledge obtained from the computational layer will be further refined to provide wisdom to the end-users, who can view the recommended solutions and current situation through the visualisation layer. Through the various interfaces and mixed reality displays, users can interact and control the physical assets, as well as implement system-generated solutions. The modules showcased within each layer can be integrated to fulfil specific use cases [87], and as such, potential technology combinations can be deployed to overcome advanced construction pain points.

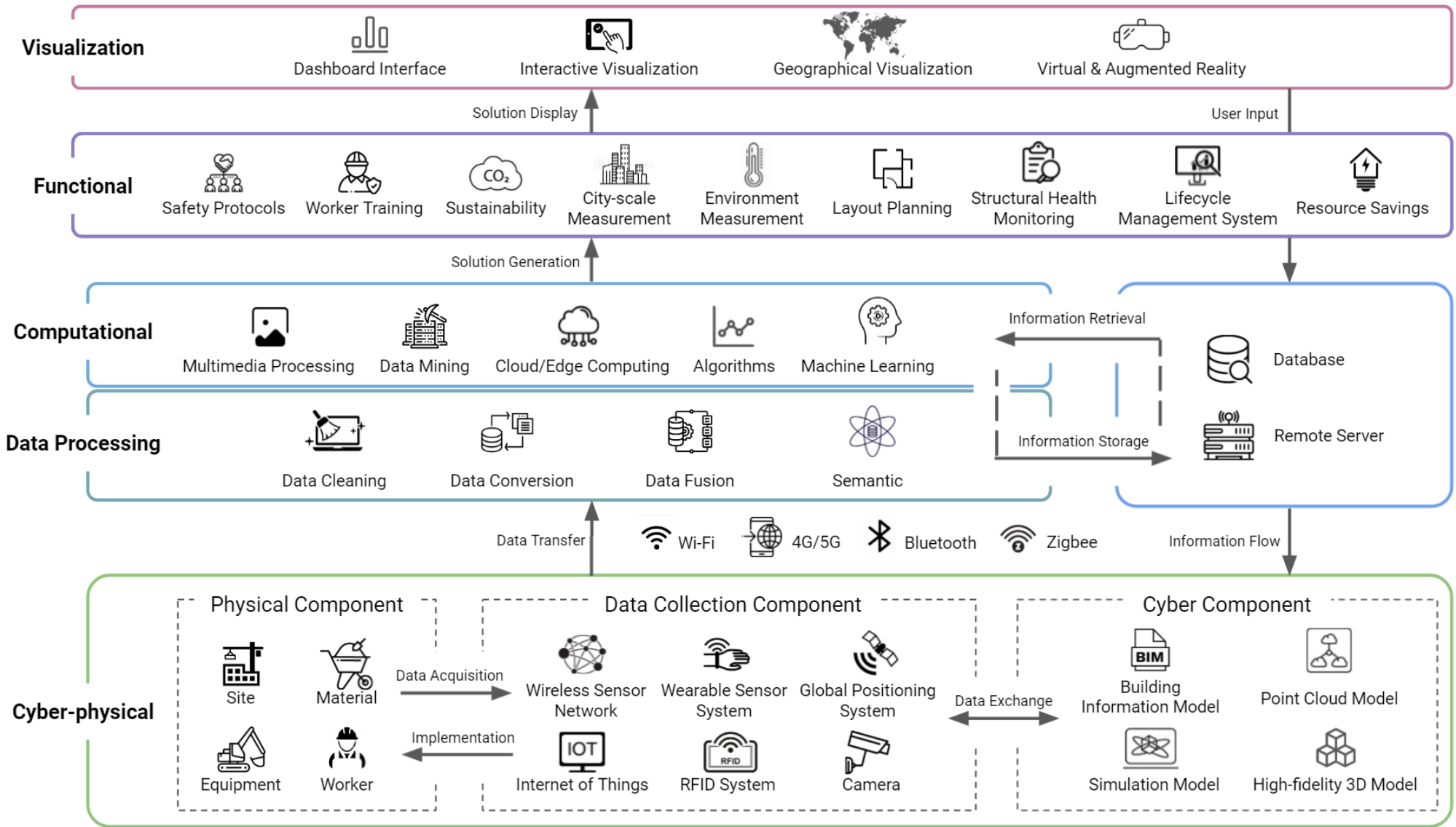


Figure 6. A unified technological architecture featuring co-construction elements for Construction 4.0

3.2 DT Perspectives on Construction Lifecycle Aspects Based on a Six M Methodology

This section presents a novel Six M methodology to represent the essential factors required for successful construction projects. Adapted from established reliability engineering paradigms with lifecycle consideration [157], this methodology consists of Machine, Manpower, Material, Measurement, Milieu, and Method (Six M) and is closely associated with the BLM approach [158]. Based on both the Six M and building lifecycle perspective, as illustrated in Figure 7, the benefits of DT are categorised accordingly. It is noted that DT technologies did not prominently support the ‘requirements identification’ and ‘project planning’ stages.

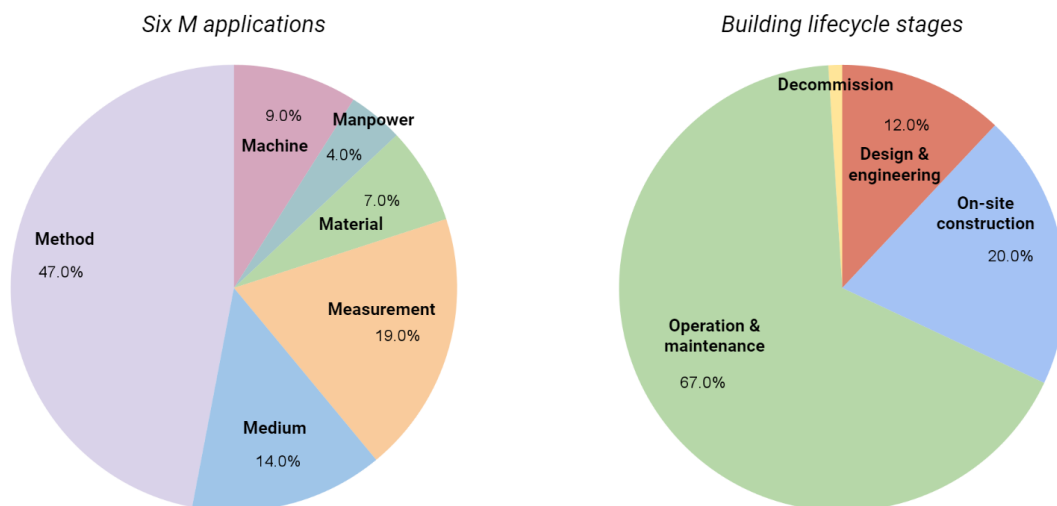


Figure 7. Distribution of existing applications on Six M and BLM aspects

3.2.1 Machine

The Machine aspect represents all physical assets related to equipment and machinery (E.g., truck, crane) used within the construction industry. DT technologies are often deployed on high-value assets to enhance efficiency and reduce breakdowns throughout the usage phase. Within the *on-site construction phase*, Zhang et al. [159] proposed a multi-system coupling mechanism to support a CPS-based hierarchical autonomous control of a tunnel boring machine (TBM), while Li et al. [160] developed a DT-driven virtual sensor structure for trailing suction hopper dredger to support pre-warning and

safe operations. Zhang et al. [161] proposed a unified robot-oriented framework for building automation and robotics. Liang et al. [162] established a robot control policy to handle repetitive tasks, while Zhang et al. [163] developed a dynamic data-driven modelling mechanism to allow robots to handle automatic pavilion constructions. Also, Lee et al. [164] proposed a DL approach to enhance the task allocation performance during robotic construction. In the *operations and maintenance phase*, Jiang et al. [165] simulated hoisting behaviours to avoid potential safety accidents for the tower crane, and Liu et al. [166] utilised the SVM approach to predict risk in advance for prefabricated elements hoisting. Besides, Kan et al. [167] proposed a layered CPS approach to plan and monitor mobile crane operations.

3.2.2 Manpower

The Manpower aspect involves the working force engaged throughout the construction stage and ranges from designers to equipment operators. Current DT studies in this aspect only focus on the *on-site construction phase*. For worker safety, Wu et al. [168] introduced a real-time visual warning system to proactively avoid dangerous entities for construction workers based on deep learning and mixed reality. For worker training, Akanmu et al. [169] showcased a CPS-driven postural training platform for workers to practice construction operations via a VR environment, while Sepasgozar [170] integrated VR/AR and DT technologies to educate stakeholders in the design, development, and implementation of a tunnel boring project as part of an education pedagogy.

3.2.3 Material

The Material aspect includes raw materials and intermediate products such as precast models, and recent studies mainly focus on material performance and tracking. In the *design and engineering phase*, He et al. [171] utilised a BIM-enabled fabrication methodology with material configurations to advance 3D printing capabilities in construction. Also, Orozco-Messana et al. [172] presented a novel solution for building envelope regeneration through leveraging Phase Change Materials elements. In the *on-site construction phase*, Marini et al. [173] presented a CPS-based method to improve the traceability and radiological detection of construction material. In the *operations*

and maintenance phase, Meža et al. [174] explored the use of secondary raw materials (SRM) in terms of feasibility, suitability, and sustainability in the long run. In the *decommissioning phase*, Züst et al. [175] enhanced construction sustainability by evaluating excavation and demolition material flows via economic and circular economy perspectives.

3.2.4 Measurement

The measurement aspect refers to the transformation of information drawn into descriptions and quantities, which can be extended to include the value, cost, and price of construction work. Moving towards a digitalised landscape, data acquisition and status monitoring of both physical objects and target environments are essential for the development of infrastructures in a speedy, dependable, and sustainable manner. In the *on-site construction phase*, Hao et al. [176] reduced carbon emissions through a BIM-based approach. In the *operations and maintenance phase*, Chiachío et al. [177] proposed a DT framework for structural monitoring tasks within the civil engineering domain; many researchers implemented practical health monitoring systems for different structures and scenarios [178], including beam string structures [179], large-scale structures [180], timber buildings [181], bridges [182] [183], underpass [184], and road [185]. For non-destructive asset evaluations, Angjeliu et al. [186] developed an expedited modelling method for dynamic reconstruction simulations, while Kong et al. [187] presented an ML approach based on percussive diagnostic techniques to determine conditions of bolted joints. To assist in failure avoidance, Mahmoodian et al. [188] developed CPS-enabled monitoring systems to provide early warning capabilities and corrective instructions.

3.2.5 Milieu

The Milieu aspect represents the physical environment in which work is carried out and includes ambient information, terrain type, and surrounding layout. In the *on-site construction phase*, Zhang et al. [189] integrated DT and the extension of the level of details in BIM into a framework to support construction site monitoring and management. Jiang et al. [190] [191] proposed a DT-enabled system to achieve the real-time management of on-site assembling during modular integrated construction. In the

operations and maintenance phase, both building and city environments are targeted as potential optimisation areas. For indoor environment management, CPS-enabled systems were deployed to boost thermal comfort for both homes and offices based on user preferences and energy efficiency [192], [193], [194], [195]. Meanwhile, Zhao et al. [18] visualised indoor conditions and energy consumption parameters, while Shahinmoghadam et al. [196] developed a VR platform to monitor thermal comfort conditions. Zhang et al. [197] developed an automatic humidity control system for heritage sites via computational fluid dynamics simulation, and Bonci et al. [198] developed a BIM-based CPS platform to evaluate and optimise building performance. For building occupancy monitoring, Gomes et al. [199] proposed a context-aware recommendation system for co-working environments, while Seghezzi et al. [200] presented an occupancy-oriented building management system to optimise cleaning operations. For smart city development, Schrotter, Hürzeler [201], White et al. [202], Cho and Kim [203], and Wu et al. [204] developed DT cities to enhance transparency, transport efficiency, and urban planning.

3.2.6 Method

The Method aspect covers the approaches used to improve B&C efficiency throughout the lifecycle. In the *design and engineering phase*, building shape optimisation can allow planners and architects to minimise environmental influence when designing buildings. Wei et al. [205] proposed an assessment approach for the wood panelised components during off-site construction to progress to the next construction step. Böke et al. [206] demonstrated the CPS-enabled automation of adaptive façade functions ranging from solar shading to sound insulation, whereas Ding and Kareem [207] utilised building shape morphing and evaluation techniques to reduce wind load significance to satisfy building drift and comfort requirements. For BEM, Porsani et al. [208] and Demianenko et al. [209] proposed a workflow and framework for automating energy analysis based on BIM. For indoor environment design, Jia et al. [210] developed a platform-based method for rapid prototyping and explored design spaces to improve design performance. In the on-site construction phase, for safety management, Jiang et al. [211] established CPS-based risk data synchronisations with warning and scene reconstruction mechanisms, while Liu et al. [212] proposed a hoisting safety risk management framework for prefabricated buildings with

considerations to relations between risk factors. To improve construction logistics, Greif et al. [213] implemented a decision support system for silo dispatch and replenishment via fill-level monitoring. As for quality assessment, Tran et al. [214] ensured the 3D geometric quality of as-built prefabricated façades through comparisons between the as-designed and as-built digital models. To enhance sustainability, Yang et al. [215] contributed to long-term city development by summarising publications considering DT integrated with intelligent green buildings. Çetin et al. [216] explored how to apply circular principles during the construction lifecycle.

In the *O&M phase*, asset management organises resources through structured and competent means [217]. To enhance the operation and management flow of buildings, Zhao et al. [218] and Quirk et al. [219] proposed conceptual frameworks for DT-based FM systems while detailed solutions for specific assets, including pumps [220], tunnel [221], bridge [222], and airport [223] [224] have also been explored. To lower energy consumption, Francisco et al. [225] benchmarked daily electricity usage according to the strategic period while innovative approaches, including activity monitoring [226] [227] and resilient buildings achievement [228] [229], are proposed. Bass et al. [230] and Huang et al. [231] achieved a regional energy-saving system through urban-scale energy modelling. To achieve NZEB [232], Hosseini Haghghi [233] enhanced the interoperability between urban building data and energy consumption evaluation. To enhance lifecycle management, Yitmen et al. [234] analysed the impacts of using cognitive DT systems in various lifecycle aspects based on applicability, interoperability, and integrability. Furthermore, several studies have illustrated DT-enabled decision support capabilities in infrastructure projects [235], [236], [237] with emphasis on sustainability and vulnerability. Lastly, Zu and Dai [238] highlighted a distributed path-planning strategy to reduce crowd-induced casualties during building evacuation.

3.3 Future Directions for DT in Construction 4.0

This section reviews potential trends highlighted from existing studies and outlines 11 directions to advance DT-enhanced systems in construction. These future directions are

categorised from technical, application, environmental, and management perspectives, as shown in Table 8.

Starting with system and technology enhancement, the development of diverse and multi-function sensor systems would facilitate data collection in complex and harsh environments through intelligent, miniaturisation, and integrative functionalities to support cellular networking, GPS, and robotics. AI-enhanced functionalities such as ML, CV, and optimisation algorithms can significantly improve process efficiency and provide better analysis and solutions. Multi-function and integrated DT systems aim to incorporate functionalities with higher operation performance to include additional project considerations within the same platform, such as environment monitoring, safety management, and building evacuation.

Next, a wider scope of implementation would provide more industrial relevance as DT systems can take on increased functional roles and alleviate pain points based on a holistic outlook. Multi-asset servitisation requires multi-source datasets to manage high-quantity assets and leverage resource data to enhance overall B&C operations. City-scale DT systems highlight the shift from building-oriented DT systems towards the mapping and management of virtual cities for mass administration and urban planning cases. Broad industry implementations will provide a wider sense of realism to enterprises and industry stakeholders by implementing DT-enabled solutions with an emphasis on industrial protocols and guidelines to resolve real-life situations feasibly. Encompassing the entire lifecycle refers to the expansion of DT solutions to include knowledge from design to demolition stages to better manage bottlenecks economically and efficiently.

The circular economy promotes sustainability and is emerging as a crucial factor in the modern construction environment. Sustainable construction mainly targets resource savings, emission reduction, and waste management in the on-site, operation, and decommissioning stages. Meanwhile, lean concept integration ensures the efficiency and environmental footprint of prefabricated production operations through reduced material and energy consumption in the design and manufacturing stages.

Lastly, DT has the potential to improve project management from both a time and cost perspective. The Time-based analysis utilises DT-enabled solutions to optimise project timelines, mitigate disruptions, and reduce the risk of delays. Economic considerations factor in the financial perspectives and ensure that the recommended solutions are feasible while operating within the preferred business model.

With these directions to enhance DT capabilities, other functional techniques derived from previous work, such as complex environment path planning [239], BIM-enabled detection methods [50], and other construction enablers, can leverage DT systems to enhance construction robustness and resilience.

3.4 Leveraging Future Directions for DT-enabled PdM in the Building Industry

This section examined how insights and directions discussed in Section 3.4 can be applied to enhance the DT-enabled PdM systems in the building industry.

1) System and Technology Enhancement:

In the PdM system, advanced sensor networks would be critical for monitoring the health and performance of building components. Intelligent, miniaturised sensors with integrative functionalities, such as GPS and robotics, could enable real-time, continuous data collection even in challenging building environments. Besides, AI-enabling technologies can be utilised to analyse the collected data, predict potential failures, and optimise maintenance schedules. Optimisation algorithms could further enhance resource allocation for maintenance tasks, improving system efficiency. Lastly, the PdM system can be integrated with multiple functionalities, such as monitoring environmental conditions and controlling safety, thus contributing to broader building management goals.

2) Wider Scope of Implementation:

In building maintenance, a DT-enabled PdM system can enhance efficiency through multi-asset servitisation by managing multiple building assets simultaneously, such as elevators, HVAC systems, and electrical systems. By utilising multi-source datasets,

the system can allocate maintenance resources effectively across various assets, optimising overall building performance. Besides, for large projects, such as smart cities or multi-building complexes, the PdM system can be integrated into city-scale DT systems, enabling centralised management of maintenance activities across numerous buildings. This centralisation helps standardise maintenance practices and improves operational efficiency on a larger scale. Additionally, the PdM system should encompass the entire lifecycle of building components, incorporating data from the design and construction phases. This lifecycle approach leads to more accurate predictive models and more effective maintenance strategies, improving both performance and cost-effectiveness over the long term.

3) Project Management Improvements:

DT-enabled PdM systems can significantly enhance sustainability, efficiency, and cost-effectiveness in building operations. Firstly, time-based analysis within a DT-enabled PdM system allows for failure predictions and the implementation of just-in-time maintenance scheduling. This proactive approach reduces emergency repairs, optimises resource utilisation, and mitigates the risks and costs associated with unexpected breakdowns. Besides, lean concept integration further enhances this by streamlining maintenance processes, reducing material and energy consumption, and minimising downtime. Lean principles ensure that maintenance activities are minimally disruptive, contributing to overall operational efficiency. Finally, economic considerations are equally important, as the PdM system not only lowers operational costs by reducing reactive maintenance but also extends asset lifespans. The financial feasibility of the system must be carefully analysed, ensuring it aligns with industry standards and business models and offering a cost-effective solution for long-term building management.

In conclusion, the 11 future directions provide a comprehensive roadmap for developing and refining the DT-enabled PdM system methodology, ensuring that it is technically robust, widely applicable, environmentally sustainable, and economically viable for the building industry.

3.5 Remark

The versatility and scalability of DT solutions are evident in many industries stretching from aerospace to healthcare. Emphasising the benefits of DT as a platform within the B&C industry, this chapter reviews related studies over the past six years to derive an architecture showcasing the integration of Industry 4.0 technologies as functional modules within DT systems. In addition, an in-depth analysis based on a novel Six M methodology and lifecycle perspectives provides an outlook on the advantages brought forth by DT-enabled systems for various construction functional roles. These findings are outlined below in two broad perspectives to provide insights on construction DT applications for stakeholders.

Horizontal technological perspective (Scalability of DT applications). As DT systems are primarily deployed for environment monitoring and resource tracking roles within specific use cases, there is a need to investigate the role of DT in handling multi-asset integration and complex scenarios. To overcome this challenge, the Six M methodology is highlighted to ensure the scalability of construction DT systems. Consisting of essential factors for successful project outcomes, the Machine, Manpower, Material, Measurement, Milieu, and Method aspect provides a reference model for academics and industries in the implementation of future DT models.

Vertical technological perspective (Advancement of DT systems). To further refine the role of DT systems in distinct operations, a five-layer DT-oriented architecture consisting of cyber-physical, data processing, computational, functional, and visualisation layers highlights how Industry 4.0 technologies can be incorporated. Furthermore, the breakdown of specific tools and techniques previously utilised provides a practical guide towards identifying research gaps and deploying feasible DT-enabled solutions.

As challenges faced typically involve complex resource relations, fluctuating environment conditions, and the lack of high-quality datasets, this resulted in fewer automatic decision support and recommendation systems within the Machine, Manpower, and Material aspects. Thus, 11 future directions are identified, including

diverse and multi-function sensor systems, AI-enhanced functionalities, multi-function and integrated DT systems, multi-asset servitisation, city-scale DT systems, broad industry implementations, full lifecycle encompassment, construction sustainability, lean concept integration, time-based analysis, and economic considerations. The authors hope that this research can be regarded as a guideline for more DT-related research and discussion towards innovative construction applications and Construction 4.0.

Table 7. Benefits of DT applications categorised by Six M aspects based on B&C lifecycle stages

Six M aspect	Lifecycle stage	Construction function	Reference	DT-enabled benefits
Machine	On-site construction	Intelligent equipment control	[159] [160]	Reduce steady-state errors and safety risks.
		Automatic construction robot	[161] [162] [163] [164]	Improve context observation to implement robot control policy, enhance the generative design and robotic construction through real-time perception-modelling, and achieve real-time bidirectional communication and supervision of remote collaboration between workers and robots.
	Operations and maintenance	Safety management	[165] [166]	Improve object detection confidence level in the digital triplet security system.
Asset management		[167]	Enhance bidirectional coordination between virtual and physical assets and establish context-aware capabilities for configuration and workflow efficiency.	
Manpower	On-site construction	Worker safety	[168]	Synchronise information in dynamic and complex environments to process hazards.
		Worker training	[169] [170]	Decrease training risk through a virtual practice platform and improve the learning effects of construction practitioners.
Material	Design engineering and	Structure design optimisation	[171]	Provide more accurate models to support the design validation of 3D-printed modules.
		Reuse and recycling	[172]	Reduce material consumption and waste generation through building component reuse.
	On-site construction	Material information tracking	[173]	Improve traceability and radiological detection of construction material.
	Operations and maintenance	Durability and response monitoring	[174]	Facilitate material circularity by exploring properties and responses of SRM.
	Decommissioning	Reuse and recycling	[175]	Guide material flows towards a sustainable material flow through quantitative assessment.
Measurement	On-site construction	Greenhouse gas emissions tracking	[176]	Improve the potential for establishing energy conservation and emission reduction strategy through real-time GHG emissions monitoring.
	Operations and maintenance	SHM	[177] [178] [179] [180] [181] [182] [183] [184] [185] [186] [187] [188]	Provide promising paradigms for real-time and continuous SHM application, including structural damage detection, safety assessment, failure avoidance and maintenance operations assistance.
Milieu	On-site construction	Construction site monitoring	[189] [190] [191]	Improve construction digitalisation through automatic detection and monitoring of construction site and assembly progress.
	Operations and maintenance	Indoor environment management	[192] [193] [194] [195] [18] [196] [197] [198]	Benefit visually dynamic common platforms for intelligent indoor management functions, including real-time monitoring, safety maintenance, thermal comfort, and reducing resource consumption.
		Building occupancy monitoring	[199] [200]	Improve space utilisation and sensor system efficiency and accuracy through real-time building occupancy monitoring and intelligent algorithms.
		Smart city development	[201] [202] [203] [204]	Easier demonstration and transparency of administration tasks, urban planning, and policy to the public through visualisation and analysis of digital prototypes.
Method	Design engineering and	Building shape/profile optimisation	[205] [206] [207]	Automate façade functions development, minimise the influence of wind load through dynamic façade, provide a cost-effective method to satisfy serviceability limits and optimise the shape of the concrete roof structure with complex geometry for energy saving.
		BEM	[208] [209]	Enhance the interoperability between BIM and BEM in the building design phase.

	Indoor environment design	[210]	Enable rapid prototyping of applications to improve design performance by reusing hardware and software on shared infrastructures.
On-site construction	Safety management	[211] [212]	Enhance safety management in construction sites through risk factors analysis, proactive risk control, and threat assessment.
	Construction logistic	[213]	Support decision-making during silo dispatch and replenishment activities.
	Quality assessment	[214]	Facilitate the visual quality assessment of as-built prefabricated façades during the construction process.
	Sustainability enhancement	[215] [216]	Support data synchronisation and blockchain integration for traceability and incorporate the smart product-service paradigm.
Operations and maintenance	Asset management	[218] [219] [220] [221] [222] [223] [224]	Better access to siloed data and support the development of asset management applications such as real-time monitoring and more intelligent decision-making for cognitive buildings.
	Energy reduction	[225] [226] [227] [228] [229] [230] [231] [233]	Promote energy-saving construction to achieve energy-reduction goals through accurate energy simulation analysis, encourage energy-efficient behaviours, and intelligent matching of residents and activities.
	Lifecycle management	[234] [235] [236] [237]	Enable cognitive features in assets to support sustainability and vulnerability assessments and maintain quality throughout the construction lifecycle.
	Building evacuation	[238]	Provide guidance information for efficient building evacuation in emergencies.

Table 8. Future directions for DT in Construction 4.0

Category	Reference	Future Direction	Description
Technology enhancement	[23] [26] [37] [156]	Diverse and multi-function sensor systems	Develop advanced miniature sensors with intelligent and integrative features to support GPS, 4/5G, and robotics for performance improvements.
	[25] [87] [90] [93] [96] [240]	AI-enhanced functionalities	Automate and accelerate learning, reasoning, and perceiving from extensive datasets to tackle higher-order tasks such as detection, prediction, optimisation, and planning.
Application scope	[35] [46] [219] [241] [242]	Multi-function and integrated DT systems	Enhance computation capabilities to include higher quality simulation and solution accuracy as well as faster processing time to support visualisations and evaluations.
	[84] [95] [243]	Multi-asset servitisation	Integrated solutions using assets and resources to enhance recognition, tracking, and management operations.
	[201] [202] [244] [245] [246]	City-scale DT systems	Validate current DT architecture on a broader scale and expand DT applications from building to community and city levels to provide the foundation for the optimisation of city services.
	[36] [49] [54] [59] [167] [176] [247]	Broad industry implementations Encompass full lifecycle	Incorporate complex multi-asset scenarios based on real-life practices to suit industrial needs with an information-rich DT model. Achieve an efficient DT system that can be used to plan, design, operate, maintain and demolish economically and environmentally throughout the whole lifecycle of the construction project.
Circular economy	[31] [32] [225]	Sustainable construction	Improve resource efficiency, tracking, and reducing emissions, extend asset lifespans, and enhance waste management through functional component monitoring and analysis in each lifecycle stage.
	[62] [63] [248]	Lean concept integration	Integrate lean concepts within digital solutions to enhance resource-sustainable infrastructure projects or implement lean manufacturing approaches for PPVC production.
Benefits analysis	[249]	Time-based analysis	Explore the influence of DT solutions on project timelines with comprehensive dataset analysis.
	[250] [251]	Economic considerations	Ascertain the financial viability of DT adoption and the use of DT solutions to achieve cost savings.
	[252]		

Chapter 4 The Survey of PdM in Building Industry

Numerous studies aiming at enhancing maintenance practices within the building industry have recently emerged under the context of the rapid advancement in the ICT field. In this chapter, we explore the current development landscape of PdM within the building domain. Our investigation employed the keyword network analysis approach to extract research clusters and identify the critical factors in existing PdM-related studies. Besides, we explore the role of PdM in the building industry through a structured Three-by-Three M analysis approach. Lastly, we discuss the research challenges and offer insights into future directions within this domain. Drawing from the analysis of 122 recent papers over the last five years, this chapter serves as a fundamental reference for PdM applications in the B&C industry and paves the way for future developments.

4.1 Research Requirement and Summary

Maintenance is a process of ensuring the reliability, safety, and longevity of complex systems and products, which is widely valued in many industries, particularly in fields such as manufacturing, transportation, and facilities management, where the maintenance of critical components can have a significant impact on safety, efficiency, and overall performance [253]. Currently, the mainstream approaches are reactive and preventive maintenance, which lack the capability of condition assessment and potential malfunctions prediction. The latest maintenance paradigm, known as Maintenance 4.0 [6], leverages the enabling technologies from Industry 4.0 to implement more sophisticated and proactive techniques, such as failure prediction, fault classification, and schedule optimisation. These emerging methodologies can be assimilated into PdM, which rectifies the limitations of existing maintenance approaches. PdM is a data-driven method aimed at monitoring the condition states and forecasting the failure events of the target asset. The implementation of PdM firstly requires real-time monitoring of equipment, and then its decision-making process is

heavily dependent on the integration of multi-source data and the computational ability of AI algorithms.

In our opinion, the implementation methodologies and utilisation scenario of PdM solutions in the B&C industry are two main aspects. Our study aims to solve the aforementioned problems with three objectives:

- 1) Furnishing a comprehensive examination of methodologies and technologies for feasible deployment of PdM in the B&C industry.
- 2) Examining the role of PdM in complicated B&C projections while considering significant factors.
- 3) Analysing current research challenges and providing instructive work towards potential directions for PdM in the B&C industry.

In this chapter, recent PdM publications in the B&C industry are analysed systematically through a keywords network to explore enabling technologies and prevalent application scenarios in a preliminary manner. Furthermore, methodologies and implementations are consolidated within the Six M methodology (Method, Measurement, Manpower, Machine, Material, and Milieu), which is critical for the successful execution of any B&C-related project.

4.2 Research Methodology

This review focuses on academic articles related to PdM in the B&C industry over the past five years, including frameworks, methodologies, implementations, etc. The five-year time range was chosen to reflect the most recent advancements and trends in PdM research within the B&C sector, ensuring that the review captures the latest technological and methodological developments in this rapidly evolving field. The study refers to the PRISMA pipeline [254] to ensure a structured and transparent review process, while the analytical methodology integrates techniques from prior previous studies [255][256][257][258][259].

The identification of studies is guided by the PRISMA framework, as depicted in Figure 8, which delineates the process into identification, screening, and included components. Scopus, the largest academic database, was selected to conduct the literature search due to its comprehensive coverage of high-quality peer-reviewed publications across a wide range of disciplines. Two sets of keywords associated with PdM and the B&C industry were employed to ensure comprehensive coverage. The search process can be replicated using the following pseudocode:

*Topic = (“Condition-based Maintenance” OR “Predictive Maintenance”
AND “Building” OR “Construction” OR “Civil Engineering” OR
“Architecture, Engineering, and Construction Industry”)*
Time Span: 2018–2022
Language: English
Type = “Article” OR “Review”
Source = “Journal”

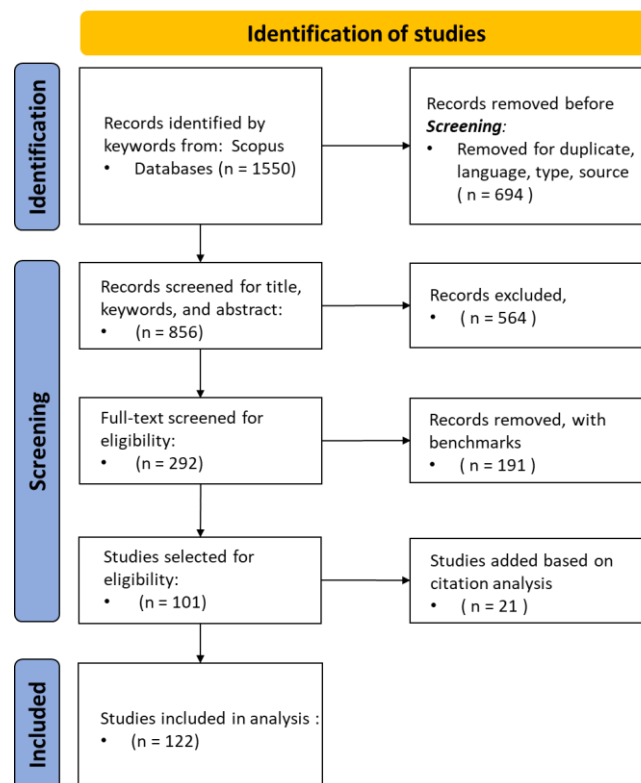


Figure 8. Identification of Studies

Subsequently, the records are scrutinised and screened through titles, keywords, abstracts, and full-text documents. The following selection benchmarks are utilised to

refine the inclusion of papers relevant to this review: (1) PdM applications and scenarios are selected, including construction machinery and buildings assets during the construction and maintenance stage, which contains conceptual and empirical discussion about PdM, allowing critical technologies to be highlighted. (2) PdM frameworks are examined to provide an overview of designs, methodologies, and functions involved, which supports the identification of challenges for maintenance evolution. (3) Studies directly and indirectly involving PdM concepts are examined, even those without mentioning PdM in the title, keywords, or abstracts, contributing to the discussion of PdM for future perspectives.

The outlined workflow and criteria facilitate a comprehensive review process, ensuring an extensive and appropriate scope of literature. Consequently, 122 publications were identified, and 16 supplementary works were cited to enhance the credibility and validity of the survey.

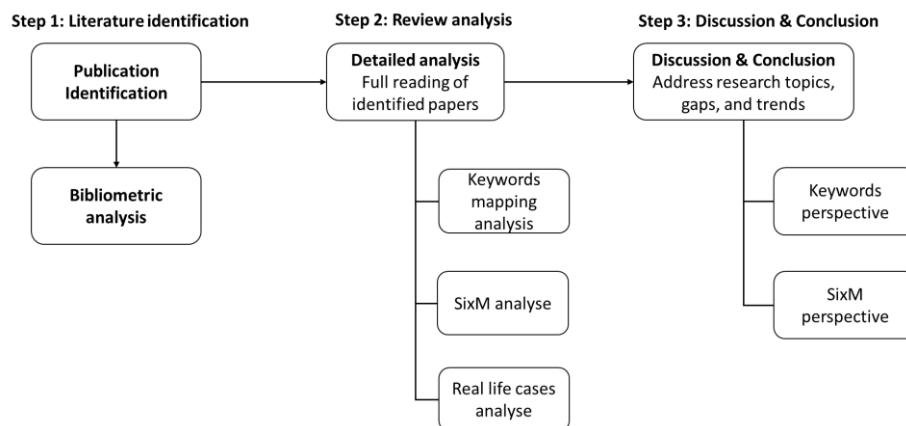


Figure 9. An overview of the research methodology

Figure 9 demonstrates an overview of the methodology employed in this study. Following the publication identification, a bibliometric analysis is performed to discern trends and patterns within the extant literature, such as research popularity and field distribution. Then, the detailed analysis contains the keyword mapping, the proposed SixM analysis, and real-life case studies. Lastly, a discussion and conclusion are conducted using both keywords and SixM perspectives. Research gaps, limitations, and future directions are analysed and explored.

Furthermore, Figure 10 illustrates the publication count by country, highlighting the global distribution of the field and identifying leading nations in research output. Most

research originates from developed countries or regions, such as Europe, the United States, and Australia, reflecting the significant demand for building maintenance in these areas. This trend is associated with the advanced stage of infrastructure development and the corresponding emphasis on maintenance in developed regions. In contrast, developing countries are predominantly focused on the construction stage, resulting in comparatively less research dedicated to the maintenance phase.

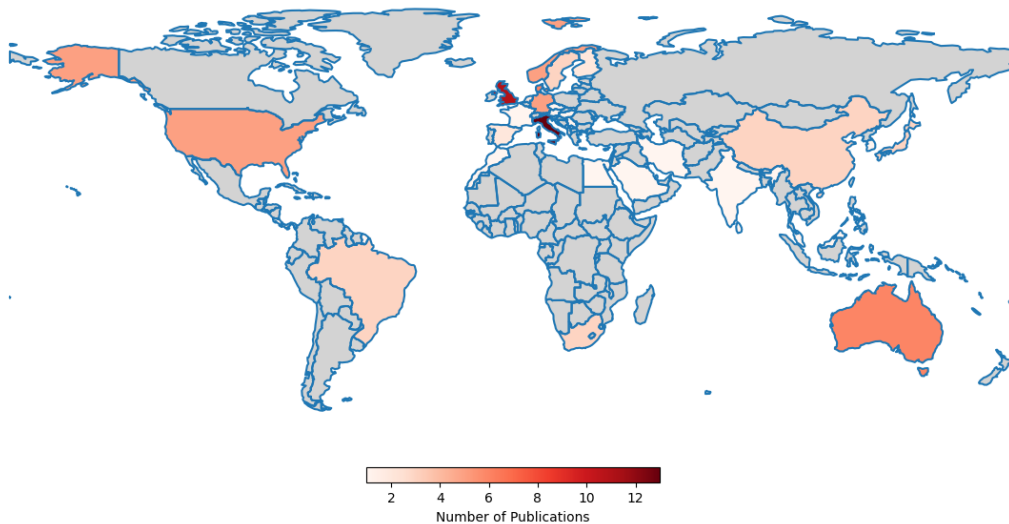


Figure 10. Publication counts by country

4.3 Keywords Network Analyses

Keywords play the most critical role in identifying the main concern of research [260]. The keywords network is established in VOSviewer by text processing and clustering algorithms [261]. Apart from author keywords, “index keywords” are also utilised in this scientometric analysis to include any keywords neglected by authors. The keyword occurrence count in the paper usually represents its main scope and research concern, while co-occurrences mean two keywords utilised in the title, abstract, or keywords list of the same article.

To conduct accurate and meaningful analyses, keyword filtering techniques are adopted to optimise the keyword selection process. A thesaurus is established to merge the keywords in different expressions with the same meaning, and then meaningless keywords like methodology and approach are manually removed. The keywords co-occurrence network shown in Figure 11 is a visualisation representation based on

bibliometric analysis, which contains keywords with a minimum of 4 co-occurrences. As a result, a total of 59 out of 1391 keywords meet the requirement and are tabulated in Table 9.

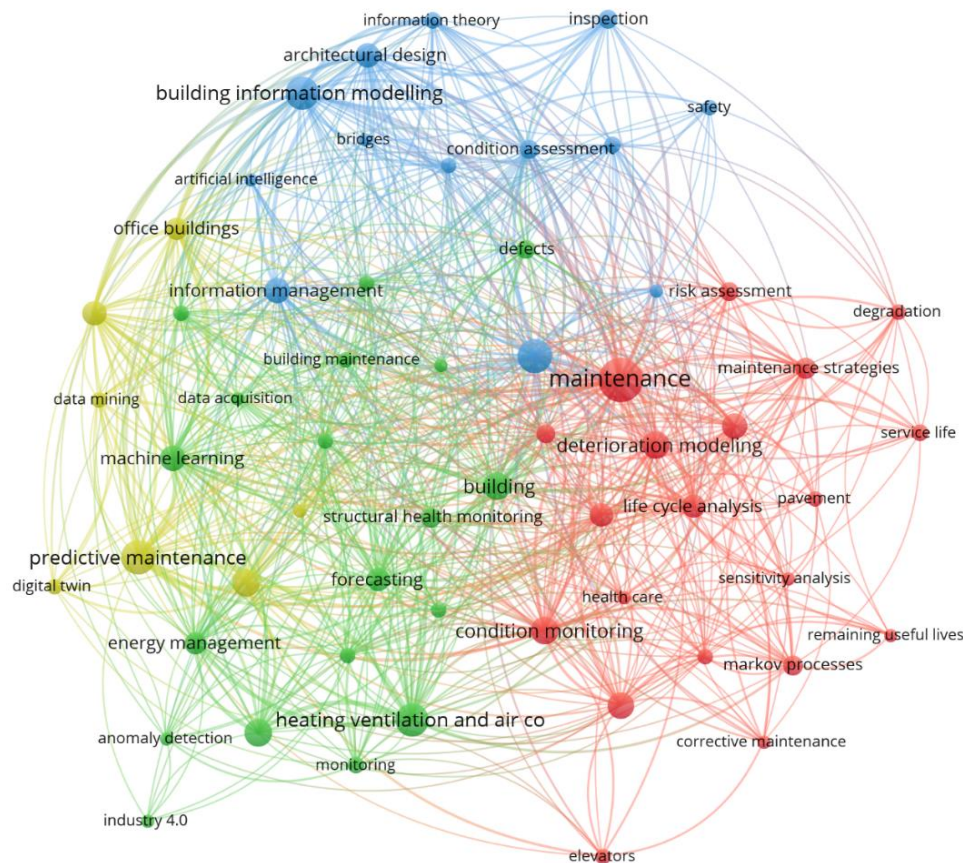


Figure 11. Network visualisation of keywords co-occurrence map about PdM in the B&C industry

Keywords are clustered through the VOS clustering approach [262] and colour-labelled in Figure 11; the colour of keywords signifies the cluster to which they belong. The circles and labels represent the keywords, while the popularity of keywords is represented by the circle size and the font size of the label. Keyword relevance is measured by the distance among them in Figure 11. A shorter distance among different keywords in the network means a stronger relationship, while a longer distance indicates a weaker relationship. Figure 11 only displays the strongest relationship among the keyword's networks. Four clusters with their keywords are present in Table 9, and their attributes are analysed in the following sections.

Table 9. Scientometric analysis of keywords about PdM in B&C industry

Keywords	No. of links	Total link strength	Occurrences	Avg. citations	Avg. pub. year
Cluster 1 (Red Coded, 20 items)					
condition-based maintenance	34	58	14	8.86	2020.93
condition monitoring	39	89	15	16.60	2021.13
corrective maintenance	13	21	4	20.50	2020.25
cost analysis	33	63	13	15.31	2020.54
degradation	14	19	5	9.60	2020.20
deterioration modelling	40	75	15	14.73	2020.60
elevators	11	18	5	3.40	2021.20
health care	16	21	4	8.00	2020.50
life cycle analysis	34	64	11	18.55	2020.64
maintenance	50	183	39	12.08	2020.59
maintenance strategies	31	50	9	29.33	2020.11
Markov processes	22	36	8	4.75	2020.88
monte carlo methods	22	27	5	10.40	2021.40
pavement	11	16	5	3.40	2021.00
preventive maintenance	34	55	11	22.36	2020.73
remaining useful lives	9	13	4	10.50	2020.50
repair	27	38	7	44.57	2020.57
risk assessment	27	34	7	4.71	2021.14
sensitivity analysis	13	16	4	4.50	2021.00
service life	14	21	6	14.00	2020.17
Cluster 2 (Green Coded, 19 items)					
anomaly detection	16	17	4	43.75	2020.00
building	49	102	16	24.69	2020.25
building maintenance	22	29	5	5.80	2021.40
data acquisition	28	37	4	57.00	2020.25
decision trees	19	20	4	19.50	2020.00
deep learning	26	33	5	26.40	2021.20
defects	29	42	7	27.29	2020.57
energy management	32	60	10	14.20	2020.70
fault detection	27	53	15	12.20	2020.80
forecasting	41	70	11	16.45	2021.09
Heating, ventilation and air conditioning	34	96	22	11.27	2020.73
industry 4.0	8	8	4	4.50	2021.25
learning systems	29	43	5	64.60	2019.20
machine learning	36	81	13	28.31	2020.54
monitoring	21	28	6	26.33	2020.50
operation and maintenance	30	39	5	24.80	2020.80
predictive analytics	25	34	5	10.80	2020.80
railroads	14	15	5	4.20	2021.40
structural health monitoring	18	26	8	25.25	2019.63
Cluster 3 (Blue Coded, 13 items)					
architectural design	38	100	12	31.50	2020.75
artificial intelligence	18	19	4	14.75	2020.25
bridges	15	17	4	7.25	2020.00
building information modelling	38	109	22	22.95	2020.36
condition assessment	29	45	7	41.29	2020.14
construction industry	29	42	6	16.17	2021.17
decision support systems	47	126	24	13.54	2020.71
information management	38	87	13	22.62	2020.54

information theory	26	51	6	44.50	2020.83
inspection	18	39	7	21.43	2020.71
safety	12	16	5	12.00	2020.60
structural analysis	18	24	4	4.25	2020.75
visualisation	19	32	5	23.40	2019.80
Cluster 4 (Yellow Coded, seven items)					
data mining	19	23	5	12.80	2020.60
digital twin	15	22	5	27.80	2021.00
facility management	36	81	12	30.92	2020.75
Internet of things	39	90	15	22.87	2020.73
office buildings	40	80	10	33.90	2021.00
performance assessment	12	13	4	6.50	2021.00
predictive maintenance	36	106	23	13.48	2020.83

1) Number of Links

This attribute represents the number of links a keyword has with others. The two keywords with the most links are maintenance (50) and building (49), as the investigation concerns PdM in the B&C area. The third keyword is the decision support system (47), revealing the nature of PdM. This is followed by forecasting (41), deterioration modelling (40), office buildings (40), condition monitoring (39), and the Internet of Things (39). The first four in the list indicate that PdM focuses on monitoring and forecasting functions for building-related assets, which are the main progress and improvement compared with traditional maintenance. IoT shows this is the most critical and foundational supporting technology for PdM implementation in the B&C industry. The lowest number of links belongs to Industry 4.0 (8), which shows that current PdM research in the B&C industry has not fully utilised and integrated enabling technology and knowledge from the Industry 4.0 perspective. Therefore, future research should explore the potential of cooperating with Industry 4.0 enablers with PdM in the B&C area. Other lower keywords, such as remaining useful lives (RUL) (9) and pavement (11), are functions provided and targets applied, which means relatively fewer research efforts and industry applications in these areas.

2) Total Link Strength

The strength of a link indicates the number of documents using different keywords in conjunction; more co-occurrence means a higher strength of links. Total link strength indicates the total strength of the keywords' links, which can be accepted as the influence index of keywords in this area. According to Table 9, the keywords with the highest total link strength are maintenance (183) again, followed by decision support system (126), building information modelling (109), predictive maintenance (106), and

building (102). The emerging term is BIM and PdM. PdM is expected because it is the research concern. Compared with IoT, BIM has fewer but stronger links, meaning BIM has lower popularity but more profound impacts than IoT in this area. Digitalisation in the B&C industry began with computer-aided engineering and 3D information models, which were first utilised by architectural engineers during the design stage. This led to the expansion of BIM-based studies in other life stages, such as the O&M phase. Indeed, the collection, storage, management, and utilisation of information are regarded as the largest challenges throughout the PdM implementation. The research about the combination of IoT and BIM has the promise to enable PdM to achieve efficiency improvement, service life extension, and other advantages. Therefore, BIM integration is the key to the PdM in the B&C industry and will occupy a prominent place in future research. The gap in Industry 4.0 (8), performance assessment (13), and RUL (13) issues in PdM research are the same as the analysis for the number of links and need to be concentrated on in the future.

3) Occurrence Frequency

Occurrence frequency indicates the number of keyword occurrences in articles, showing the main concern of the article. A higher occurrence frequency of keywords indicates an intense focus on the particular field. Its pattern and trend are similar to the total link strength. The keywords that scored the highest occurrence frequency are maintenance (39), decision support system (24), predictive maintenance (23), heating, ventilation and air conditioning (22), and BIM (22). The new revelation is that the predominant application of PdM solutions within the B&C industry lies in HVAC system maintenance. The low boundary of occurrence frequency comprises numerous keywords as the threshold for co-occurrences is set at 4, which contains a limited spectrum of application scenarios, like healthcare (4) and bridges (4), as well as underutilised methods including AI (4) and deep learning (4).

4) Average Citations

This attribute indicates the total citation counts received by all publications utilising the keyword divided by publication counts. Average citation scores are computed free from the research area, while the PdM in the B&C industry is involved in multiple disciplines (Figure 11). Repair (44.57), condition assessment (41.29), facility management (30.92), maintenance strategies (29.33), and machine learning (28.31) are the top five keywords,

indicating attractive topics which are more likely to be cited. Future studies should focus on the ML-driven task related to the first four categories in the list. The gap mainly exists in pavement (3.40), elevators (3.40), railroads (4.20), structural analysis (4.25), and Markov process (4.75), which means the article related to these topics are not easy to gain attention from researchers. To synthesise, in the application domain, researchers tend to favour the utilisation and citation of studies that demonstrate adaptability and universality. For the methodology side, it has been observed that ML techniques offer more informative insights and reference value compared to traditional numerical methods, such as Markov processes.

5) Average Published Year

The average publication year of the documents provides insight into the chronology of keywords' occurrence in the relevant papers. The closer year of publication, the newer research topic represented by keywords. It is noted that keywords such as Monte Carlo methods (2021.40), elevators (2021.20), deep learning (2021.20), condition monitoring (2021.13), and digital twin (2021.00) are relatively recent and frequently used in conjunction with PdM research in the B&C sector. The chronological sequence of these keywords suggests that in recent years, researchers have been directing their attention towards exploring intelligent, visualised, and real-time solutions for increasingly building-related assets through advanced computing methods.

4.4 Three-by-Three M-based Literature Analyses

Keyword analyses in Section 4.3 provide a preliminary understanding of the prevalence and technological status of PdM in the B&C industry. However, a comprehensive review of the research situation requires a systematic approach that incorporates a comprehensive perspective and meticulous summarisation. The refined Six M approach, derived from [259] and originally developed from Ishikawa's diagram [263], is utilised to analyse the crucial determinants of successful PdM projects in the B&C sector. The schematic representation of the Six M methodology is shown in Figure 12. Milieu, Machine, and Material (Hard Ms) are utilised to distinguish the targets and scenarios of PdM, while Manpower, Method, and Measurement (Soft Ms) categorise their methodology. M concepts are listed in Table 10. The maintenance action is

achieved through manual operation, intelligent optimisation, automatic bidirectional information transfer, etc., between the two M groups. The outermost layer refers to the division of the construction lifecycle [158]. PdM is predominantly applied in the on-site construction phase for construction machinery and the O&M phase for building assets.

This section presents an in-depth analysis of the literature about PdM in the B&C industry, and an extensive review of existing studies has been summarised in Table 11 using the Six M approach. To further clarify the details of the Soft M, Table 12 provides a comprehensive illustration of the various PdM tools and techniques that have been applied in the B&C industry. It should be noted that the table only listed the algorithms, software and hardware mentioned in the article. As authors typically only highlight noteworthy tools or techniques in their articles, the absence of information in the table (indicated by "NA") implies that the author did not emphasise such details, rather than no relevant requirement or support. Additionally, the distribution of the Soft MS and Hard MS is presented in Figure 13. These tables and figures provide us with a visual representation of the prevalence of different M aspects in the B&C industry and the extent to which they have been valued in practical situations.

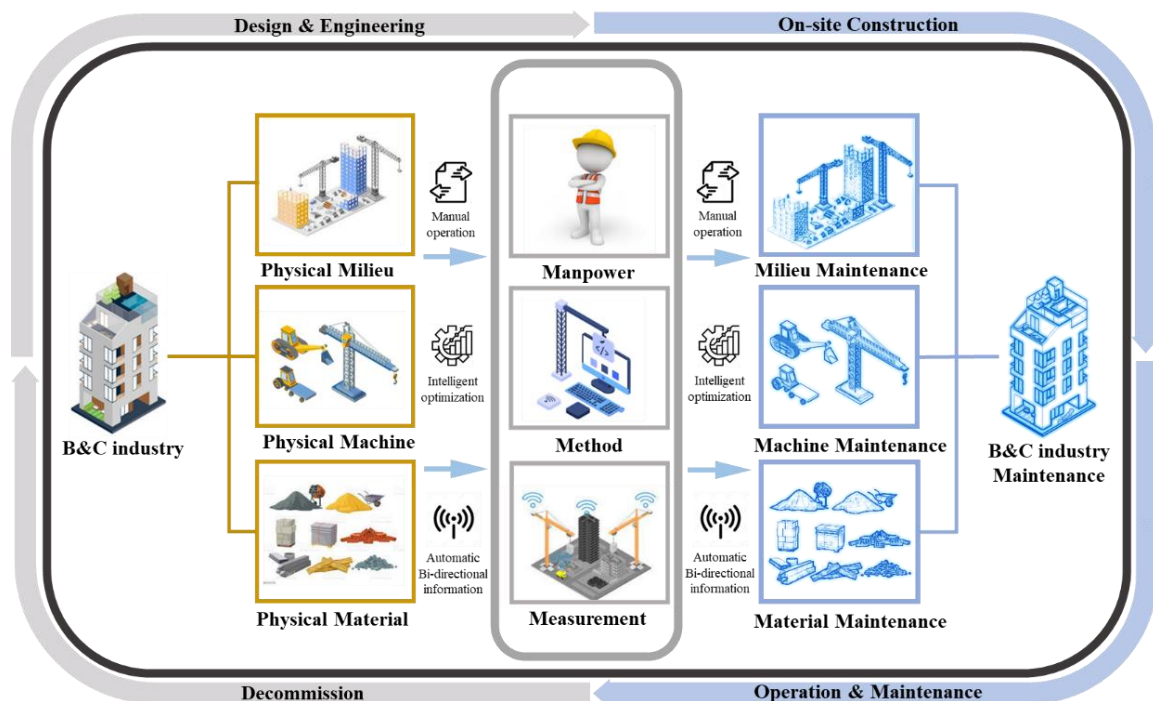


Figure 12. The Three-by-Three M approach for PdM in B&C industry

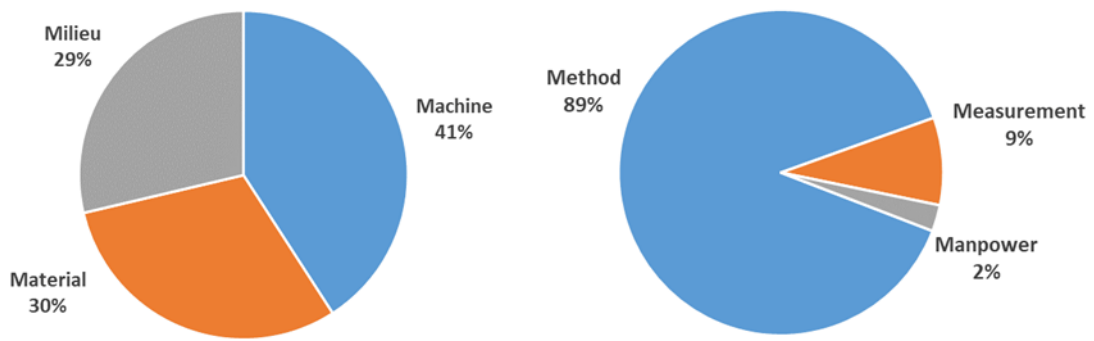


Figure 13. Distribution of Hard MS Applications & Soft MS methodology

Table 10. M definition in the Three-by-Three M method

M	Definition
Method:	Refers to the procedures and processes to improve the PdM performance in the B&C industry. This category addresses issues related to work procedures, work sequence, and workmanship.
Measurement:	Refers to the methodology for the conversion of graphical information into numerical information, and description can be expanded to evaluate the building assets and on-site construction processing. This category addresses issues related to accurate measurement and data transformation.
Manpower	Refers to the workforce-centric solutions for PdM implementation in the B&C sector, which range from designers to operators and supervisors. This category addresses issues related to training, experience, and workload.
Machine:	Refers to building facilities (HVAC, lift, etc.) and construction equipment (crane, TBM, etc.) utilised in the B&C industry. Issues in this category are mainly related to repair, malfunctions, and obsolescence.
Material:	Refers to all other B&C-related raw materials and components except machines such as pavement, walls, tiles, roofs, etc. Issues in this category are mainly related to quality, availability, and compatibility.
Milieu:	Refers to the practice environment where the assignment is conducted, such as buildings or bridges, etc. Issues in this category are mainly related to site management, ambient information, and surrounding layout.

Table 11. The Three-by-Three M matrix for literature analysis

	Method							Measurement		Manpower
	Fault identification	Fault prognosis	Degradation analyses	Strategies optimisation	Life cycle management	Asset management	Defects detection	SHM	Labour management	
Machine	HVAC Systems	[264], [106], [120], [123], [133], [265], [220], [266], [119], [141], [267], [268], [269], [270], [271]	[135], [131], [272], [273], [274]	[275]	[276], [277]	-	[278], [279]	-	-	-
	Elevator & lift system	[280], [138], [281]	[139]	-	[282], [283]	-	[284], [285]	-	-	-
	Lighting system	[286]	-	-	-	-	-	-	-	-
	Fan system	-	-	-	-	-	-	-	-	-
	Heat Meters	-	[287]	-	-	-	-	-	-	-
	Fire safety equipment	-	-	-	-	-	[288]	-	-	-
	TBM	-	-	-	[289]	-	[290]	-	-	-
	knuckle boom crane	-	-	-	-	-	[291]	-	-	-
	Dragline subsystem	[292]	-	-	-	-	-	-	-	-
	Slurry pump	-	-	-	-	[293]	-	-	-	-
	Switch machine	-	-	-	-	[294]	-	-	-	-
	Smart grid facility	[295]	-	-	-	-	[143]	-	-	-
	Wind Turbines	-	[296]	-	-	-	-	-	-	-
	Material	Multiple building structures	-	-	[297]	-	-	-	[298]	-
Lining Structure		-	-	-	-	[299], [300]	-	-	-	-
Roof structure		-	-	-	[301], [147], [302]	-	-	-	-	-
Concrete beams		-	-	-	-	[303], [304]	[305]	-	-	-
Cladding systems (Facades)		-	-	-	[306], [307]	-	[308]	[309], [310], [311]	-	-
External Thermal Insulation Composite		-	-	-	[312], [313]	-	-	-	-	-

	Systems (ETICS)								
	Building envelope elements	-	-	-	[314]	-	-	-	-
	Pavement structure	-	-	[315], [316]	[317], [318]	[319]	-	-	[320]
	Metallic frame structure	-	-	-	-	-	-	-	[37]
	Masonry structure	-	-	-	-	-	[321]	-	[322]
	Water diversion structure	-	-	-	-	-	[323]	-	-
	Pipe System	-	-	-	-	-	[324]	[325]	-
	Steel wire rope, Cable	-	-	-	-	[326], [327]	-	-	-
	Railway related components	-	[328], [329]	-	-	-	-	-	[330], [331]
	Public building	[332]	-	-	[333], [334], [335], [336]	-	[337], [338], [339], [340]	-	[341]
	Timber building	-	-	-	[343]	[344]	-	-	-
	Heritage buildings	-	-	-	-	-	-	-	[345]
	Healthcare building	-	-	[128]	[145], [346], [347]	-	[348]	-	-
	Bridge system	[188]	-	[349], [350]	[237]	-	[351], [352], [353]	-	[354], [355]
Milieu	Penetration dock	-	-	-	-	-	[356]	-	-
	Offshore wind farms	-	-	-	[357]	-	-	-	-
	Oil and Gas Field	-	-	-	[358]	-	-	-	-
	Smart city	[359]	-	-	-	-	-	-	-
	Business parks	-	-	-	-	-	-	-	[360]
	University laboratory	[361]	-	-	-	-	-	-	-
	Confined spaces	-	-	-	-	-	-	-	[362]

Table 12. Detail analysis of Soft Ms

M	Function	Tools/Techniques
	Fault identification	<p>Algorithm/Methodology: <i>ML cluster:</i> Bayesian network model, Calibration techniques in ML classification models, Clustering-evolutionary hybrid approach, GA, Semantic modelling, Bayesian optimisation algorithm, SVM, Principal component analysis (PCA), Bayesian change point detection methodology, Particle swarm optimisation (PSO), Clustering-PSO enhanced algorithm, Hierarchical extended KNN, RF classifier/regression, Moving window algorithm, Classification and RT, Feature extraction approaches, <i>ANN:</i> 2-D CNN, Variational AE techniques, AE and LSTM-encoder-decoder, Deep AE. <i>Numerical cluster:</i> Variance, Lag-1 autocorrelation function (ACF1), Power spectrum (PS), Monte Carlo simulation, Gaussian distributions, improved Gaussian mixture model with PCA approach</p> <p>Software: <i>Programming:</i> MATLAB, Microsoft Visual Studio Community 2019, JavaScript, WebGL; <i>Modelling:</i> Autodesk Revit, Autodesk Forge Platform, TRNSYS, HVACSIM+ simulation platform, TSI ProHood Air Capture, Autodesk Revit Scan-to-BIM plugin, EnergyPlus, DesignBuilder; <i>Data management:</i> Netica software, GraphDB, Oracle, MySQL, Microsoft Azure, Amazon Web Services, Google Cloud; <i>FM:</i> Elevint</p> <p>Hardware: Hood model PH731 TX5 laser scanner, Geographic Information System (GIS), Sensor system (Tilt sensor, strain gauge, vibrometer, Current, Voltage, Temperature, Ambient temperature, Humidity, Photoresistor, Distributed slave meter, measurement control unit (MCU), Sensor board: Arduino MCU, Raspberry Pi 3B</p>
		Method
	Degradation analyses	<p>Algorithm/Methodology: <i>ML cluster:</i> Nonlinear optimisation method, Metropolis-Hastings algorithm (MHA), Hybrid GA-based fuzzy Markova model, Markov chains, 3D scene reconstruction algorithms (SfMVSM), <i>Numerical cluster:</i> Markov-chain Monte Carlo (MCMC) method, Markov degradation models, NLO models, Quasi-Poisson regression, the least square interpolation, Piecewise deterministic Markov process (PDMP), Chapman–Kolmogorov equation, Monte Carlo simulation</p> <p>Software: <i>Programming:</i> MATLAB, Fuzzy Logic Toolbox in MATLAB, <i>Modelling:</i> Autodesk Civil3D, Autodesk Dynamo, Autodesk Revit, Autodesk Recap, Autodesk Inventor; <i>Data management:</i> Microsoft Excel, Flux, Konstru, Speckle; <i>Knowledge:</i> Bridge deterioration knowledge ontology</p>

	<p>Hardware: UVA (DJI Phantom 4 Pro), Traffic speed deflectometer (TSD), Pavement assessment vehicle (iPAVe), Geophysical survey systems, Model 4108 1.0 GHz horn</p>
<p>Strategies optimisation</p>	<p>Algorithm/Methodology: ML cluster: Reinforcement learning (proximal policy optimisation algorithm), Hybrid hierarchical k-means clustering, GA, Decision-tree based classification methodology, Non-dominated sorting GA (NSGAI), RF, Vector autoregression (VAR), Vector error correction model (VECM), Autoregressive moving average (ARMA) model, Sentiment analysis, PCA Numerical cluster: Weibull distribution, Log-normal for varying rates, Multiple linear regression equation, Binary linear programming, The least squares method, Bayesian structural time-series model, MCMC Decision support cluster: Multiple objective optimisation-based clustering approach, Fuzzy logic-based multiple criteria decision-making (MCDM), Fuzzy axiomatic design (FAD), Additive ratio assessment (ARAS), Stepwise weight assessment ratio analysis (SWARA), Weighted additive sum product assessment (WASPAS), Fuzzy analytic hierarchy process. Knowledge cluster: Qualitative studies (grounded theory), Knowledge management model, Semi-structured interviews, Questionary, Market survey, Granger causality, Visual survey, SQL-based association rule mining</p> <p>Software: <i>Programming:</i> MATLAB, Python, R; <i>Modelling:</i> Autodesk CAD, Autodesk Revit, Autodesk Dynamo, CSI-Bridge, Open Studio; <i>Data management:</i> NVivo; <i>Model optimisation:</i> CPLEX, General algebraic modelling system (GAMS); <i>Sustainment Management System:</i> USACE-CERL SMS, SMS BUILDER; <i>Construction Cost Estimating:</i> RSMean; <i>Crack Inspection:</i> SCC-MATS</p> <p>Hardware: NA</p>
<p>Life cycle management</p>	<p>Algorithm/Methodology: ML cluster: SVR, Hidden Markov model, Fuzzy logic, ANN, Generalised linear regression, Ensemble inference models, Kalman filtering, Exhaustive Chi-square automatic interaction detection algorithm, Static/dynamic tests, PCA Numerical cluster: Stochastic maintenance models, Petri nets, Weibull distribution, Fused root mean square index FRMSI, Deterministic and probabilistic approach, FEM, Monte Carlo simulation with statistical distributions Decision support cluster: Multi-criteria analysis</p> <p>Software: <i>Programming:</i> MATLAB; <i>Data management:</i> IBM SPSS, Hardware: MTS 64.106 tensile machine</p>
<p>Asset management</p>	<p>Algorithm/Methodology: ML cluster: Bayesian network model, Text mining, MLP models, Autoregressive integrated moving average (ARIMA) models, Fuzzy set theory, Fuzzy logic, X-means clustering algorithm, SVM, Regression tree (classification/prediction), KNN algorithm, ANN Numerical cluster: Markov chain model with distribution fitting techniques, Markov chains, Simon model, Statistical method, Independent and identically distributed (IID) test, k-s test, Failure mode effects and criticality analysis (FMECA); Empirical calculation method (Creep numerical model, Shrinkage numerical model), Finite element modelling, Experimental and numerical investigations, Dempster-Shafer (D-S) theory, Analytic network process (ANP), Hausdorff distance, Maximum deviation squares, Nonlinear finite element (FE) program</p>

	<p>Decision support cluster: MCDM methods, Analytical hierarchy process (AHP), weighting method</p> <p>Knowledge cluster: Questionnaire survey, Evidential reasoning, Expert interviews, IDEF0 methodology, GIS</p> <p>Software: <i>Programming:</i> Microsoft Visual Studio, MATLAB, Software R, C++, C#, Python, Arduino; <i>Modelling:</i> Bridge information model (BrIM), HoloBridge, Autodesk Unity, Autodesk Revit, Autodesk Dynamo, OBJ Exporter, iDM Revit Plugin, COBie Toolkit, Autodesk COBie Extension, Autodesk Inventor, AgenaRisk, Ansys Workbench, AutoCAD, Apple SceneKit, Apple ARKit; <i>Data management:</i> Cloud database, InfluxD, Microsoft excel, Easyfit</p> <p>Hardware: Microsoft Hololens, Drones, UBIDOTSplatform, Sensor system (Arduino microcontroller, DHT11 sensors, PT550 light sensors, Power meter, Current Transformer (CT), Ambient sensors (temperature, humidity, etc.))</p>
Defects detection	<p>Algorithm/Methodology: ML cluster: Hierarchical feature selection (HFS) Algorithm, Least-squares SVM (LS-TWSVM) algorithm, SVM, KNN algorithm, one-vs-all LR algorithm, Repeated random subsampling validation, Monte Carlo cross-validation approach, Class activation mapping (CAM), Fuzzy logic, Neural fuzzy system. <i>ANN:</i> Mask R-CNN, CNN, pre-trained CNN classifier, ResNet-50</p> <p>Numerical cluster: Finite elements analysis (FEA), Least-squares twin, Hybrid feature selection, 2D Continuous Wavelet Transform (CWT), Kullback–Leibler divergence, Failure mode and effects analysis (FMEA)</p> <p>Knowledge cluster: Prototyping, Questionnaires, Expert interviews.</p> <p>Software: <i>Modelling:</i> Autodesk Revit, Autodesk Dynamo; <i>FM:</i> Computer Aided FM software, SmartLab, Microsoft Excel</p> <p>Hardware: Unmanned aerial vehicles (UAVs) (DJI Phantom 4 RTK et al.), Robot, Scanner, DSLR cameras, Camera (Nikon D810 et al.), Sensor systems (Ultrasonic sensor, PixHawk autopilot, PX4 flow sensor, IoT sensors)</p>
Measurement	<p>Algorithm/Methodology: ML cluster: ML tools, Advanced CV-based method, VAR, Vector error correction model (VECM), ARMA model, PCA</p> <p>Measurement cluster: Advanced LiDAR-based method, Strain-based displacement estimation; 3D eddy current pulsed thermography, 3D fusion and visualisation method, X-ray CT, Deconvolution interferometry using ambient vibration, Ambient noise inter geometry, Weibull or log-normal for varying rates</p> <p>Numerical cluster: FEM, Polyreference least squares complex frequency domain method (p-LSCF), Covariance driven SSI method (SSI-COV), Pruned exact linear time (PELT) method, Differential time-square-root</p> <p>Knowledge cluster: Granger causality</p>
SHM	<p>Software: <i>Programming:</i> MATLAB, C++; <i>Modelling:</i> David 3D system, Point cloud library; <i>Data Management:</i> SPSS, Distributed database, InfluxDB, JSON-LD; <i>SHM:</i> MOVA, MOSS</p> <p>Hardware: Dual-camera system, LiDAR system, 3D eddy current pulsed thermography system (Induction Heater, Signal Generator, Infrared Camera, 3D Scanner, Induction Coil, Rail Track Sample), 6 Degree-of-Freedom (DoF) system-in-package LSM6DSL device, Sensor system (ST Microelectronics, low-power MCU; GPS, Pressure sensor, IMU, Accelerometer, Strain gauge, Linear variable differential transformer, MEMS-based IMU, Sensor Area Network (SAN) bus)</p>

Manpower	Labour management	<p>Algorithm/Methodology: <i>ML cluster:</i> Data analytics algorithms, Textual classification algorithm, Data mining algorithm, Term frequency, Term frequency-inverse document frequency), Frequent itemset analysis, Frequent itemset analysis + sliding window, RF, Hierarchical RF <i>Numerical cluster:</i> Monte-Carlo simulation <i>Decision support cluster:</i> Quantitative decision-making models,</p> <p>Software: <i>Modelling:</i> Autodesk Revit, Autodesk Architecture; <i>Programming:</i> Visual Studio, JAVA IDE, HTML, JavaScript, C#; <i>Data management:</i> SQL Server; <i>Geographic information:</i> GIS, ArcMap, ESRI Shapefile, GeoJSON; <i>Space safety monitoring:</i> CoSMoS</p> <p>Hardware: TolesB mote platform, Digital sensor-based networks</p>
-----------------	--------------------------	--

4.4.1 Hard Ms

1) Machine:

Machinery maintenance accounts for over 40% of all studies, and it can be categorised into building facilities and construction machines. The PdM research on building facilities covers a wide range of systems, while the study of construction machinery is constrained to particular types. Additionally, the number of documents under each type of construction machine is fewer, while building facilities-related studies are more intensive. PdM for the HVAC system is a highly sought-after field, accounting for more than half of the publications in machine aspects. One explanation for this discrepancy is the complexity of construction machinery and the challenging operating environment, both of which significantly increase the difficulties associated with the implementation of PdM. Future research should potentially draw upon the PdM knowledge gained from studies on building facilities like HVAC systems and extend them to additional assets, thereby further strengthening the comprehensive of PdM in the B&C industry at machine aspects.

2) Material:

Material-related research occupies nearly one-third of the focus for PdM within the B&C sector, which contains different building structures made up of various materials. The scope of the considered structure covers various components from roof to facades as well as elements like slabs, beams, columns, etc. For material types, PdM for masonry, metallic frame, concrete, and ceramic elements have been implemented. It is noteworthy that some research has proposed unified methods that can be applied to multiple targets. For instance, [297] proposed a novel deficiency and deterioration

prediction method that employs snapshot data for 12 different building elements, covering both material and machine aspects. Apart from building structures, some MEP-related materials like steel wire ropes, cables, and sewer pipe systems have also been explored, which play a significant role in the maintenance of B&C projects. In general, the application of PdM is extensive and average without obvious partiality at the material level.

3) Milieu:

Approximately 29% of the literature in the field is dedicated to research on Milieu-related topics, and it can be classified into building and community levels according to the scope of maintenance implementation. The building level refers to a single building, such as a hospital or bridge, while the community level represents multiple buildings. The scope of the building level includes buildings with various structures and functions, such as common residential or office buildings, timber constructions, heritage buildings, and healthcare facilities. The community level is relatively complex and can range from penetration docks to offshore wind farms, oil and gas fields, and even smart cities, which typically encompass a diversity of building combinations with varying functions and types. Though maintenance efforts at the milieu level primarily concentrate on key indicators for the main functional aspects of the target environment, they still result in a substantial improvement of the milieu performance and cast light on facility-level maintenance in the B&C industry.

4.4.2 Soft Ms

1) Method:

The method accounts for the majority of methodology, contributing most of the functions and services for PdM and supporting a wide range of maintenance aspects, from machine, material, and milieu within the B&C industry. This category encompasses not only micro-level techniques such as fault identification, fault prognosis, and degradation analysis but also macro-level functions, including strategies optimisation, lifecycle management, and asset management.

Fault identification involves malfunction detection and fault type determination, which are the most popular subjects among Method-related publications. The application is

primarily in the machine field, with only limited studies focusing on Milieu's concerns regarding individual building and bridge systems. In terms of technical realisation, ML techniques such as decision trees, RF, and ANN are extensively used, while some numerical algorithms are employed for feature extraction. The implementation of these algorithms is dependent on the extensive data acquired from the scanner system, sensor network, and other support information providers like GIS. A variety of programming, modelling, and data management software is adopted to support this process. Fault prognosis is the prediction of potential failures in a system or component through an analysis of its current and historical performance, which represents a more advanced and intricate capability. The application is mainly concentrated in the Machine field, and the number of studies is limited. The technical support requirements of fault prognosis and fault identification are quite similar, with the main difference being that the prognosis algorithm pays more attention to regression instead of classification models. Degradation analysis refers to the process of analysing the gradual decline or reduction in the performance of a system over time. The application of this method is evenly distributed in each M aspect. The predominant tools employed for degradation analysis are numerical models, such as MCMC-related models, while the application of ML algorithms is relatively limited. The use of fuzzy logic is frequently mentioned and applied in this area, with MATLAB and its fuzzy logic toolbox being the primary development software over other popular coding languages. The highlighted hardware of robot-related devices plays a crucial role in enhancing the efficiency and diversity of data collection beyond the conventional sensor system.

Strategic optimisation refers to the decision-making process for the achievement of the desired outcome with the optimal and efficient utilisation of resources, which typically contains cost-effectiveness improvement and schedule optimisation. This is the second most prevalent direction in the Method aspect, with a majority application in the Milieu area. The solution achievement utilises a diverse array of algorithms and techniques, encompassing ML clustering, numerical clustering, decision support clustering, and knowledge clustering. This is because strategic optimisation contains a wide range of problems, and each requires a unique approach. As for the developing process, such research projects can be efficiently handled through the common programming languages, complemented by data management and 3D modelling software, while investigators seldom highlight the hardware requirements necessary for this stage.

Lifecycle management pertains to the systematic and structured method for managing and enhancing the full life cycle of a product, service, or system. The focus is primarily directed towards Material aspects to prolong the lifespan of building structures and materials. The algorithm primarily originates from the fields of ML and numerical analysis. Decision support techniques such as multi-criteria analysis have also been applied by some researchers. Nevertheless, the software and hardware requirements associated with lifecycle management are not emphasised in detail in the available literature. Asset management is the process of acquiring, maintaining, and disposing of assets systematically and cost-effectively. The prerequisite for realising asset management is usually to realise asset monitoring and assessment. Research in this area primarily focuses on the Machine and Milieu aspects, with limited attention to the Material area. Tools and techniques share similarities with strategy optimisation, with a special emphasis on the sensor systems utilised in the hardware.

2) Measurement:

Although the second highest number of research studies within the Soft M pertains to Measurement, only 9% of studies adopted related methods. The functions and services within this realm encompass defect detection and SHM. Defect detection in the B&C industry mainly refers to the process of identifying common defects, including cracks, holes, discolouration, stains, and warping of materials. In addition to conventional ML methods, numerical methods such as FEA have been adopted by many researchers in this area, although they serve more as a supplementary and supportive role to ML. Knowledge-based methods like questionnaires and expert interviews are also used to establish and label training sets. The utilisation of UAVs, robots, and scanners is prominently emphasised in the hardware requirements for data collection, constituting a significant difference from the method approach, which mainly relies on sensor systems. SHM refers to the process of continuously monitoring the structural integrity and performance of a building, bridge, or other civil structure. SHM and defect detection share many similarities in tool requirements and algorithm selection, while SHM has a broader application scope that covers both Material and Milieu domains. It is noteworthy that the application of Measurement methodologies in Machine aspects remains an under-explored area. The reason is that PdM for Machine aspects is mainly accomplished through the monitoring of mechanical and electrical parameters rather than the measurement of indirect parameters such as external surfaces or working

environment. The exploration of leveraging Measurement-related knowledge to enhance or streamline the PdM implementation on the Machine side merits the attention of future researchers.

3) Manpower:

Research related to Manpower solutions is the least prevalent, with only three articles accounting for a mere 2% of all research. These studies are classified within the category of labour management and are solely focused on Milieu aspects. The primary objective of labour management is to minimise labour costs and streamline manual decision-making during maintenance operations, which is achieved through the implementation of a visualised and intelligent task assignment system. Additionally, worker safety management during maintenance is also considered, and it is achieved through the integration of sensor systems and BIM technology. The development of these systems utilises algorithms from ML, numerical analysis, and decision support fields and is assisted by common modelling, programming, and data management software, as well as sensor systems. Although numerous studies examine PdM from a strategic perspective, including maintenance schedule optimisation, they have a narrow focus on the potential economic and efficiency influence of advanced technologies or equipment. Current research overlooks the critical role of human factors in PdM, which is a highly under-explored area. Future research endeavours must integrate manpower-driven approaches into the implementation of PdM.

4.5 Discussion

According to the review and analysis of recent PdM research in the B&C industry, the existing challenges and future directions were further discussed in this section.

4.5.1 Existing Challenges and Gaps

PdM-related studies in the B&C industry have attracted increasing attention since 2018. The PdM exploration in the B&C industry is less than a decade due to the lag in the adoption of advanced technology within the B&C area. In general, many PdM studies originate from the perspectives of manufacturing, materials science, and even sociology, such as lift systems, wall surfaces, and economic analysis. Therefore, these

approaches often overlook the unique characteristics of the B&C industry and present challenges in integrating comprehensive PdM solutions for this industry. For technical support, advanced predictive analytics remains a challenge due to the stringent data and algorithm requirements. Most articles are limited to asset monitoring and management. Specifically, research on fault prognosis is much less than fault detection, and failure prediction should consider not only failure time but also the type and severity of the failure, which is not widely addressed in the current stage. For methodology adoption, current solutions are always designed from Method or Measurement perspectives, ignoring Manpower-related solutions, such as task assignment, worker training, and safety management during maintenance. Additionally, the B&C industry encompasses different life stages with varying characteristics and requirements, yet current studies primarily focus on the O&M phase, neglecting the potential benefits of considering maintenance from the early stages, such as the design stage.

4.5.2 Implications for Future Research and Practice

DT enables PdM in the B&C industry.

The scient metric analysis of keywords in Section 4.3 reveals a trend of integration with DT technology in existing studies. This is because PdM and DT share numerous commonalities in the technical foundation and user requirements, including real-time bi-directional information transmission and virtual model creation. Combining DT-related technologies and methodologies will significantly advance the development of PdM in the B&C industry. To clarify the trend and sort out the evolution of PdM, the author named the clusters in Table 9 and Figure 11 according to the field of keywords.

1) Physical-numerical PdM (red):

The keywords in this cluster mainly concentrate on the fields of deterioration modelling, condition-based maintenance, and strategies analysis, whose methods are mainly derived from numerical and decision support, such as Monte Carlo and MCDM. Though these methods do not have an overreliance on massive data and intelligent algorithms, they necessitate high levels of background information and professional knowledge about the maintenance target. This renders these solutions deficient in terms of scalability and generalizability.

2) Physical-intelligent PdM (green):

Detection and forecasting-related services have become the focus of this cluster, with a substantial number of keywords related to ML for technical support. Enabling techniques under the ML and Industry 4.0 umbrella have been valued and utilised to improve the performance of PdM solutions. This category reduces the reliance on professional skills and knowledge, though it imposes certain demands for data acquisition. Critically, the prediction or detection algorithms from different maintenance domain processes are interoperable and universal, which can effectively promote the research of PdM.

3) Virtual-intelligent PdM (blue):

BIM and AI have emerged as the core technologies within this cluster, and their application has elevated to decision support, condition assessment, and information management. As the level of intelligence is further advanced, visualisation and modelling techniques are also integrated, leading to the introduction of 'virtual' in the cluster name. Research in this cluster further enhances the accessibility and operability of maintenance tasks. The clear depiction of the principle and structure of the maintenance target in the model reduces the demands for specialised knowledge and skills. However, it significantly increases the requirements for data collection due to the model establishment of the model.

4) Virtual-physical PdM (yellow):

Advanced FM and performance assessment can be achieved through continuous information exchange utilising reliable channels like DT and IoT. Data mining has also been applied to deal with massive data requirements. This type of maintenance requires a large amount of data support, including both target-related operating parameters and environment-related data like temperature. Then, with the aid of automation and robotics, the maintenance process may become as streamlined as a mouse-click operation. The implementation of these techniques will strongly improve the effectiveness of maintenance and significantly increase the cost-efficiency of assets.

This analysis is not only dependent on a systematic literature review but also inspired by the model-based and data-driven distinguishment [272], which is currently the

mainstream classification method for PdM research. The above analysis highlights that DT-enabled PdM will be an important direction in the future of the B&C field.

Future studies will promote PdM in the B&C industry.

According to the Three-by-Three analyses, a comprehensive overview of the implementation of the PdM in the B&C industry is presented. The future study refers to aspects within Six M that have not been fully considered or addressed in the process, which results in inefficiencies or other problems. Discussions on such M, including Milieu, Measure, and Manpower, as well as recommendations, are provided to offer guidance for future research.

In terms of application, Milieu is the most neglected direction at present despite being the best representation of the characteristics of the B&C industry during maintenance. The derivation of the Machine and Material aspects from research in manufacturing and materials science has resulted in a higher level of technological accumulation and inheritance of knowledge from the parent industry. In contrast, research on the Milieu aspect is often specifically designed for B&C projects, with a focus on the civil environment, such as buildings and bridges. The Milieu-related research considers the characteristics and lifecycle perspective in construction projects, which is more conducive to enriching the Six M overview and promoting the development of PdM in the B&C field. It deserves the attention of future researchers.

With regards to methodology, Measurement and Manpower-related approaches have huge potential for improvement. The current implementation of PdM relies heavily on professional knowledge and operating parameters like voltage and current. However, it is challenging to implement PdM through the measurement of indirect factors such as photographs, scans, and ambient temperature. In practice, the unexpected maintenance task is the result of both internal and external factors. Thus, it is worth investigating the development of PdM methods that utilise indirect information from a Measurement perspective. In addition, Manpower's role is even more neglected in current research. Despite the capability of current solutions to generate maintenance plans partially or fully automatically, most actual maintenance operations are still performed manually. As a result, the role played by Manpower is crucial and requires further consideration,

particularly in terms of workload optimisation, workforce training, and safety administration during maintenance operations.

4.5.3 Remark

The functionality and applicability of PdM deployments are obvious in different industries ranging from manufacturing to construction, along with the development of the data-driven method. With growing awareness of PdM implementation in the B&C area, there is a need to gain a deeper understanding of its overview of the application, methodology, research trends and gaps. Based on the review of 122 articles over the past five years related to PdM in the B&C industry and another 11 supplementary documents to enhance interpretability, keywords mapping and analyses have been conducted through VOSviewer and clusters are identified as Physical-numerical PdM, Physical-intelligent PdM, Virtual-intelligent PdM, and Virtual-physical PdM according to their embodiment in function and method. Additionally, to further investigate the role of PdM in the B&C industry as well as provide more valuable insights into the current state, Ishikawa's diagram invented for quality engineering is adapted into the Six M approach with Three-by-Three-Matrix form for in-depth analysis. Comprising critical aspects for achieving perfect project implementations, Machine, Material, Milieu from an application perspective and Method, Measurement, and Manpower from a methodology perspective. To our knowledge, this is the first review of PdM in the B&C industry. By critically evaluating the existing literature, the author hopes it will serve as a foundation reference and provide valuable insights for future practice and research in this field.

The results of keyword analyses indicate that PdM implementation in the B&C industry should be integrated with the DT perspective and leverage related enabling technologies and methodologies to improve solution performance and scalability. Adaptable and reused PdM solutions contribute significantly to the improvement of its promotion and cost efficiency. Besides, underestimated aspects of implementation and methodology identified by the Six M method should be enhanced in future practice and research to improve the comprehensive reference model. The limitation of this study is the extraction of bibliometric data and literature publications only from academic publications. Integrating practical and commercial works may generate more novel

insights and results. Future studies may gather information from the mentioned entities to enhance the analysis and attain more comprehensive and knowledgeable findings.

Chapter 5 A Semi-GAN-based Failure Prediction Approach for Building Facilities

The rapid growth of ICT and AI promotes the proactive paradigm shift of maintenance practices in the building industry, with the DT serving as the foundation framework to enhance data transferability and interoperability. This chapter introduces an innovative unified PdM framework rooted in the DT perspective for the building industry. This framework overcomes the constraints associated with previous facility-specific approaches, emphasising generality and scalability as core principles. Within this framework, a cutting-edge failure prediction methodology utilising Semi-supervised GAN emerges as a breakthrough, which leverages labelled and unlabelled data sources, addressing the challenges posed by the scarcity and costliness of labelled datasets. Additionally, a dedicated online platform has been developed to facilitate monitoring and prediction of information presentation. Experimental results showcased within this chapter underscore the effectiveness and superiority of the proposed method, as validated through extensive testing on public building facility datasets. This chapter comprehensively explores these transformative developments in the maintenance practices of building facilities, shedding light on the future potential of data-driven approaches to enhance operational efficiency and reliability.

5.1 Machine-Perspective PdM Implementation in Building Industry

The Three-by-Three M method introduced in Chapter 4 provides a systematic framework for developing industrial solutions and academic pursuits. The 'Machine' dimension is associated with critical building facilities like lifting systems, air conditioning units, and other related electromechanical equipment for the PdM-enhanced buildings domain. The PdM methodologies applied to such facilities are currently anchored in model-based paradigms. This requires extensive domain-specific expertise on the various facilities under consideration. However, there is an observable trajectory towards more innovative methods, where data-driven techniques are beginning to take precedence. This progression signifies a departure from conventional

practices and holds the potential to refine PdM strategies substantially. The integration of big data analytics, ML algorithms, and DT into PdM systems can lead to enhanced predictive accuracy, optimised maintenance schedules, and, ultimately, the facilitation of more durable and efficient FM. Such advancement is pivotal for fostering sustainability and resilience in building operations, aligning with the evolving exigencies of modern infrastructure maintenance.

5.2 Research Requirement and Summary

PdM is a complex data-driven strategy aiming to monitor the condition states and forecast the failure events of target assets, which heavily depends on vast real-time data and necessitates meticulous and exhaustive data labelling procedures.

The advancement of ICT and FM systems has significantly strengthened the data supplement, while the labelling task demands domain expertise and professional knowledge of target assets. Therefore, diminishing the reliance on labelled datasets of PdM solutions can greatly reduce maintenance costs and improve algorithm efficiency. Motivated by this, the author proposes a Semi-supervised failure prediction method for pivotal building facilities such as air conditioning systems and elevators based on the Semi-supervised GAN approach, which can also be embedded in the DT structure as the decision support component. The developed approach can extract the distribution and structural knowledge inherent in unlabelled datasets and then integrate them with the limited labelled dataset to complete the task.

5.3 DT-enabled PdM Framework in Building Industry

Despite the efforts made by some researchers, most existing PdM frameworks concentrate on individual building facilities like HVAC systems or elevators. There is a lack of a unified framework available that comprehensively elucidates the establishment of PdM within the building industry. Through leveraging previously featured layered DT hierarchies in the building industry [259], as shown in Figure 14, a 3-layer PdM architecture from a DT perspective has been proposed to promote the implementation of DT-enabled PdM service in the building industry.

Starting with the cyber-physical layer, building facilities like HVAC or lighting systems and the building environment from the physical side are digitalised through a systematic approach to a virtual space, emphasising real-time and bi-directional data exchange. Adhering to established public communication protocols, such as diverse communication technologies (e.g. Wi-Fi, 4G/5G, Bluetooth, Zigbee) and the OSI standard, the monitoring of status and operation can provide amplified benefits to stakeholders regarding PdM service. Also, advice and instructions can be provided to the physical side under user supervision. The selection of the simulation model should be dependent on the maintenance target and task specification, with the BIM model typically assuming the role of an integration platform. The information associated with facilities and surroundings should be visualised and seamlessly integrated utilising BIM, which is an important feature of PdM in the building industry [106].

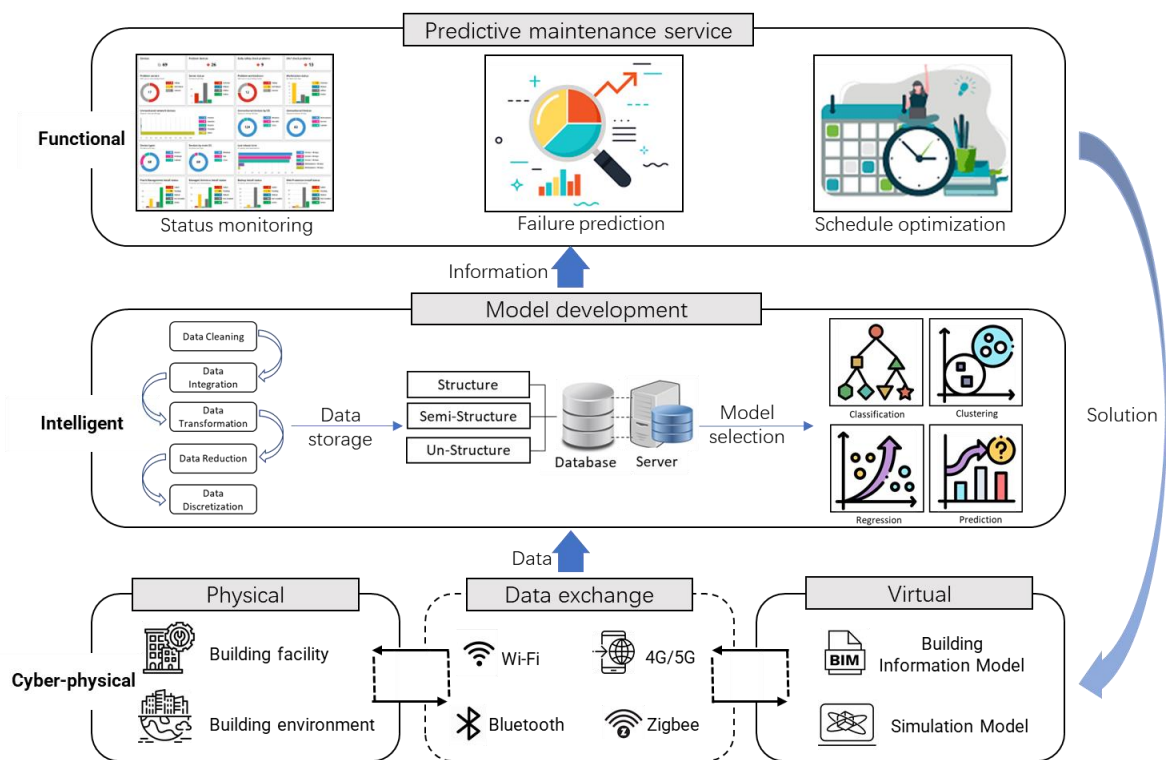


Figure 14. A unified framework of DT-enabled PdM in the building industry

Subsequently, the intelligent layer comprises the processes of data processing and model development. Typically, data processing is responsible for ensuring data quality and suitability for input into the model, as raw data collected from practical field operations often suffer from redundancies and inaccuracies due to irregular operations

and harsh environments. After data processing, suitable data are stored for further model development. Structured and semi-structured data can be stored and managed in a distributed database system, while structured query language and Hadoop distributed file systems can serve non-structured data [363]. Following that, model selection is conducted to facilitate data analytics and decision support in accordance with stakeholder or project requirements. Common model categories encompass classification, clustering, regression, and prediction models, which are tailored to specific data-driven PdM services.

The knowledge derived from the intelligent layer is refined to provide valuable insights to end-users, while the functional layer identifies the core categories of PdM services as highlighted in existing studies. Condition monitoring, failure prediction, and schedule optimisation can be perceived as progressive stages in the evolution of PdM services, wherein the demand for data and algorithms intensifies as these services are upgraded. Condition monitoring and failure prediction mainly rely on operational data from the facilities, whereas schedule optimisation requires careful consideration of maintenance records. In comparison to state monitoring, failure prediction imposes heightened algorithmic requirements, emphasising the need for further enhancements.

5.4 GAN

The initial proposal of the GAN concept can be attributed to Goodfellow et al. [364] in 2014, representing a powerful unsupervised ML technique that enables the generation of synthetic data samples exhibiting a structure and distribution that closely resemble those of the authentic training dataset. As shown in Figure 15, the standard GAN contains two neural networks: the Generator and Discriminator. The Generator aims to create indiscernible synthetic samples from actual samples, while the Discriminator aims to distinguish between real and generated samples accurately. The Generator takes random noise as input, often in the vector form containing random values drawn from a specific probability distribution (e.g., Gaussian distribution), which proceeds to produce synthetic data as the output. Typically, the Generator encompasses one or more hidden layers composed of neurons to process the random noise input and transform it into data meant to resemble real data. The Discriminator takes real data from practical

datasets and synthetic data as input, which are usually vectors containing values of different properties, and then utilises hidden layers to produce a binary output to evaluate the authenticity of the input. The objective function of the Generator and Discriminator can be expressed as $\min_G V(G, D)$ and $\max_D V(G, D)$, respectively, where G refers to the Generator while D denotes the Discriminator, V represents the value function that measures the difference between the generated and real samples.

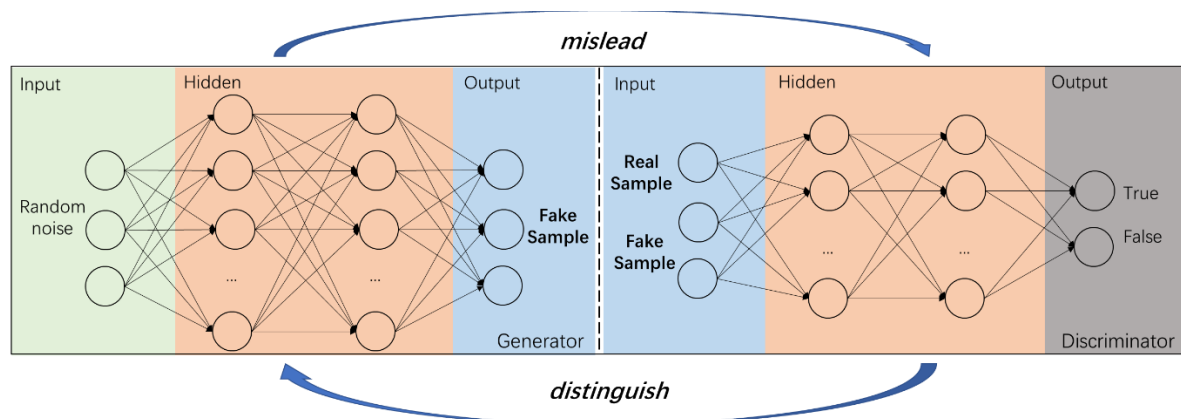


Figure 15. Schematic of the GAN approach

5.5 Semi-supervised GAN

According to Salimans et al. [365], the conventional GAN approach can be adapted into a semi-supervised learning method, allowing for the utilisation of both labelled and unlabelled datasets during the Discriminator's learning process. The Semi-supervised GAN is also composed of two networks: The Generator and the transformed Discriminator, which now acts as a Classifier. The input and output of the generator are the same as those of the standard GAN, while the main difference is between the discriminator and classifier. The input of the Classifier contains the synthetic data generated from the Generator, the unlabelled and labelled dataset acquired from the real field experiments. Specifically, the unlabelled dataset consists of vectors containing monitoring feature values without any failure or normal label, while the labelled dataset includes all of them. Unlike the binary discrimination task in standard GANs, the Classifier in Semi-supervised GAN performs multiclass classification, as it needs to not only determine the real or fake class of the input data sample but also assign it to the appropriate class in the output layer. Figure 16 shows the structure of the Semi-

supervised GAN, while Table 13 illustrates the loss function and parameter update rules of the approach.

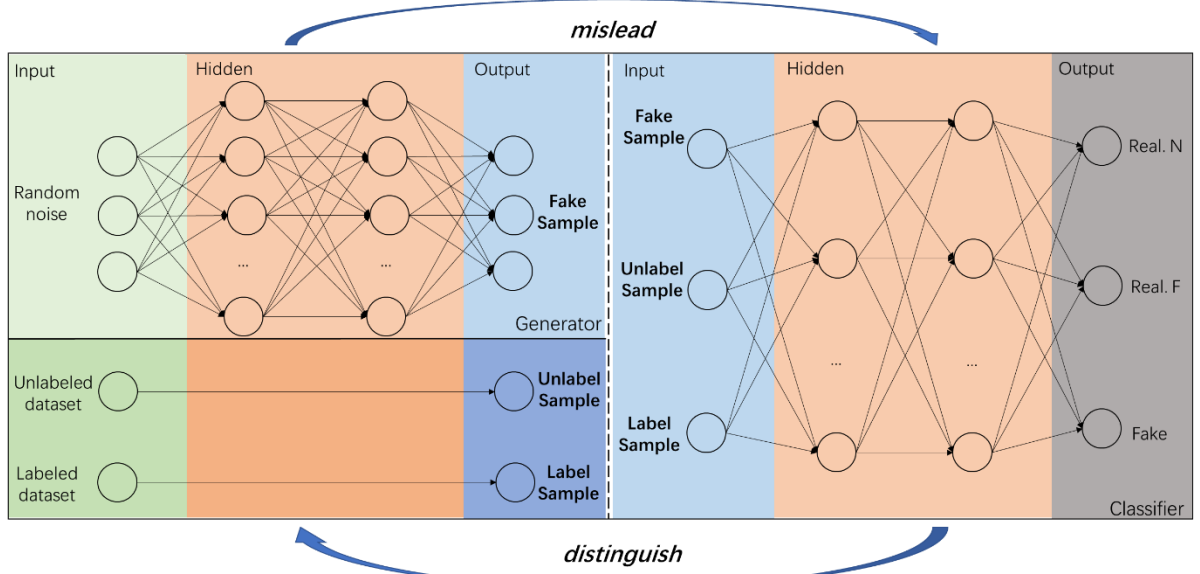


Figure 16. Schematic of Semi-supervised GAN approach

Table 13. Loss function and parameter update

	Classifier	Generator
Loss function	$L_{labelled} = -\frac{1}{m_l} \sum_{i=1}^{m_l} \log(P_c(y = t^{(i)} x_i^{(i)}))$ (1)	
	$L_{unlabelled} = -\frac{1}{m_u} \sum_{i=1}^{m_u} \log(1 - P_c(y = N x_u^{(i)}))$	$L_r = -\frac{1}{m_f} \sum_{i=1}^{m_f} \log(1 - P_c(y = N G(v^{(i)})))$ (5)
	$L_{fake} = -\frac{1}{m_f} \sum_{i=1}^{m_f} \log(P_c(y = N G(v^{(i)})))$ (3)	
Parameter update	$\theta_d = \theta_d - \eta \nabla_{\theta_d} L_i (i = labeled, unlabeled, fake)$ (4)	$\theta_g = \theta_g - \eta \nabla_{\theta_g} L_r$ (6)

For the above equations, $p_c(y = c_i | x)$ stands for the probability of x belongs to class i according to the Classifier. G refers to the Generator, and D denotes the Discriminator. m indicates the number of training samples in the dataset. L represents the loss function for different datasets, θ denotes the weight parameters of the neural network, η is the learning rate, ∇ is the gradient of the loss function. y refers to the predicted result. $x_i^{(l)}$ refers to the i^{th} labelled data, $x_u^{(i)}$ is the i^{th} unlabelled data, $v^{(i)}$ is the i^{th} random noise vector, and the $t^{(i)}$ represents the label of i^{th} labelled data, $t^{(i)} \in [0,1,2]$.

Below are three kinds of data samples and their target output of the Semi-supervised GAN approach (Table 14).

1) Labelled data:

The classifier should assign the labelled data sample to the corresponding class within the “real class”.

2) Unlabelled data:

The classifier should judge the unlabelled data sample as the “real class” and do not need to determine which class this data sample should belong to.

3) Generated data:

The Classifier should judge the generated data sample as a “fake class”.

Table 14. Example of data sample classification

	Real class		Fake class
	Normal	Failure	Fake
Normal samples	1	0	0
Failure samples	0	1	0
Unlabelled samples	0/1	0/1	0
Fake samples	0	0	1

The Semi-supervised GAN can utilise labelled and unlabeled data samples. The labelled dataset can be utilised by the Classifier directly, while the unlabelled dataset is utilised during the adversarial training process [366]. The Generator aims to generate fake samples similar to the unlabelled dataset, which means it can capture the implicit characteristics of the unlabeled dataset. The interplay between the Generator and the Classifier within the Semi-supervised GAN framework allows for the effective utilisation of both labelled and unlabeled data, leading to improved accuracy and performance of the Classifier. The input samples for the Classifier contain labelled, unlabeled, and fake data samples, while the input for the Generator is random noise.

5.6 Proposed Failure Prediction Method

This section presents the proposed methodology, including data processing, details of training, and the implementation framework.

5.6.1 Data Processing

Data processing is a fundamental and indispensable step in preparing the input data for successful deep-learning training, necessitated by anomalous data within the initial dataset, including instances of missing data and outliers. Additional techniques, such as standardisation and format conversion, are also applied to ensure the input data's

compatibility and reliability. The proper data processing significantly impacts the training outcomes, as the deep learning model heavily relies on appropriate and well-prepared input. The following steps have been implemented to effectively process the data according to the specific requirements of this research.

1) Data cleaning: Data samples containing missing values and irrelevant features are eliminated from the dataset. Removing samples with missing values ensures that the subsequent analysis is performed on complete and reliable data. Likewise, excluding irrelevant features helps streamline the dataset by focusing only on the most pertinent and informative variables for the given analysis.

2) Standardisation: Standardisation applies conversion routines to transform data into a favoured and uniform format using standard and custom business regulations. For the numerical features, the standardisation process begins by identifying the maximum and minimum values within the column. Then, calculate the standardised value for each data sample through the following equation, where x_i denotes the data sample from the original dataset while y_i represents the corresponding standardised data sample.

$$y_i = \frac{2(x_i - \min_{1 \leq i \leq n} \{x_i\})}{\max_{1 \leq i \leq n} \{x_i\} - \min_{1 \leq i \leq n} \{x_i\}} - 1 \quad (7)$$

3) RUL calculation: The calculation of RUL is critical to failure prediction, representing the time remaining before failure occurs for each data sample. The process begins by identifying the faulty data samples within the dataset. Subsequently, the time to failure is computed for each normal data sample by subtracting the timestamp of the most recent faulty sample from the corresponding normal sample's timestamp. The equation for calculating the time-to-failure value for each normal data sample is as follows, where the T refers to the value of the timestamp column.

$$T_{rul} = T_{normal} - T_{failure} \quad (8)$$

4) Label generation: The dataset aggregates the sensor data on a specific time interval, extracting a list of potentially indicative features to form one data sample. Based on the machine status records, RUL, and warning period, labels are assigned to each data

sample, which serves as the target value we aim to predict. Label 1 indicates that an alarm should be triggered due to the potential failure within the warning period, while label 0 signifies that the machine is expected to continue operating normally.

5.6.2 Details of Training

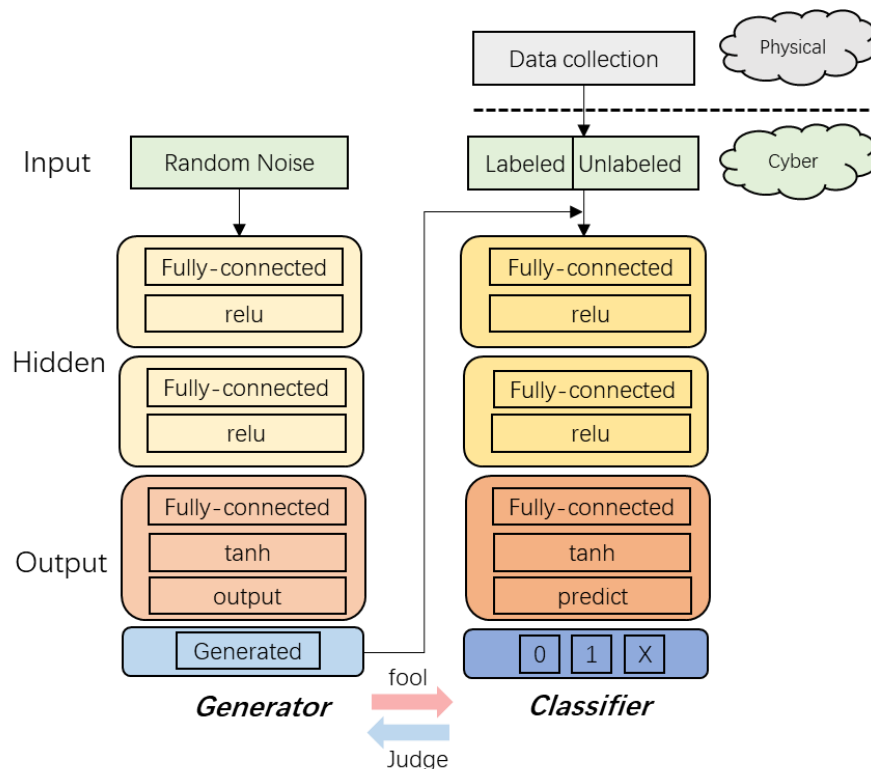


Figure 17. Sample of Semi-supervised GAN training

This section provides an in-depth overview of the model training process. Following the data processing steps, the problem of failure prediction is transformed into a classification problem that can be tackled using various existing ML algorithms. Figure 17 illustrates the structure and workflow of the proposed Semi-supervised GAN method. The trained Classifier serves as the failure prediction model during online deployment. The model training in this study is carried out using Python 3.8 within the Visual Studio Code environment. Detailed information on the proposed Semi-supervised GAN for failure prediction is presented below:

1) Hyperparameters settings.

The process of hyperparameter tuning is carried out in a systematic manner to ensure optimal performance of the proposed model, which consists of Generator and Classifier

networks. For the Classifier, the number of neurons in the input layer should be equal to the number of feature columns present in the dataset, while that of the output layer should match the number of classes. The Generator aims to generate the fake samples using random noise as input. Hence, the number of neurons for the output layer and the size of the random noise vector should align with the number of feature columns. Now, there are 3 classes of output (fake, normal, fault) and 5 different features in our study.

A neural network with more hidden neurons can represent more intricate patterns. The structure and distribution of the open-source dataset are relatively simple, making it more favourable to utilise a straightforward design for both the Generator and Classifier networks. Besides, the authors also discovered that employing a more complex Generator led to improved performance of the Classifier compared to utilising a simpler structure. This observation may be attributed to the fact that the fake data samples synthesised by a sophisticated Generator manifest a greater similarity to authentic data samples. This compels the Classifier to augment its capacity to learn the implied and structural knowledge of real samples (unlabelled dataset). Consequently, the Classifier becomes more adept at accurately distinguishing between real and fake data samples. Therefore, the experiment employed a Classifier network with a smaller number of neurons in the hidden layer, while the Generator network was relatively more complex in harnessing the above benefits.

The hyperparameter scopes of systematic experiments are shown in Table 15. The number of hidden layers is explored from 1 to 3 for both Generator and Classifier, while the number of neurons is also tested from 10 to 40. Different combinations of the number of neurons and layers are also explored in systematic experiments. According to the experiment results, the size of the generator network is chosen as (5, 40, 40, 5), and the size of the classifier is (5, 20, 10, 3). For the training process, the iteration number is set to 600, indicating the number of times the algorithm updates the model parameters using mini-batches of data. Additionally, the size of the minibatch is set to 120, specifying the number of data samples utilised in each iteration for updating the model's parameters. These hyperparameter settings are chosen to strike a balance between computational efficiency and model performance in the training of the Semi-supervised GAN model.

Table 15. Hyperparameters tuning

Hyperparameters	Range
Number of hidden layers	1, 2 , 3
Number of neurons in the hidden layer	10 , 20 , 40
Epoch number	100, 300, 600 , 1000
Learning rate	0.001, 0.01 , 0.1
Batch size	12, 60, 120 , 240

2) Optimisation algorithm.

Gradient descent is one of the most popular algorithms for optimising and, by far, the most common way to optimise neural networks. The stochastic gradient descent (SGD) technique is utilised in the model training and chosen for its favourable properties in achieving better local minima and convergence behaviour. It is a method to minimise an objective function $J(\theta)$ parameterised by a model's parameters θ by updating the parameters in the opposite direction of the gradient of the objective function $\nabla_{\theta}J(\theta)$ to the parameters. The learning rate η determines the size of the steps we take to reach a minimum.

SGD performs a parameter update for each training example $x^{(i)}$ and $y^{(i)}$. SGD eliminated redundancy by performing one update at a time so this method usually faster and can be utilised online.

$$\theta = \theta - \eta \cdot \nabla_{\theta}J(\theta) \quad (10)$$

The SGD technique is adopted to train the Semi-GAN because of better local minima and converge situations. In this study, the iteration number is set as 600, and the minibatch size is set as 120. Coding is as follows:

```
optimizerD = torch.optim.SGD(netD.parameters(), lr=0.01, momentum=0.0,
                             nesterov=False)
optimizerG = torch.optim.SGD(netG.parameters(), lr=0.01, momentum=0.0, nesterov=False)
```

3) Dropout approach.

To mitigate the overfitting issue during network training, the dropout approach [367] is employed. This technique involves randomly omitting a specified proportion of neurons during the training process. Consequently, these selected neurons are excluded from contributing to the parameter updates, thereby enhancing the model's generalisation capability. In this study, a dropout parameter of 0.2 is set for each layer,

indicating that 20% of the neurons are randomly dropped out during training to prevent overfitting and improve the overall performance and robustness of the model.

4) Synthetic data generation for imbalanced dataset

The effectiveness of ML algorithms heavily depends on the availability of a well-balanced training dataset. However, system malfunctions are far less frequent compared to normal operating conditions in real-world applications. This disparity leads to a significant imbalance in the number of data samples collected, with a much smaller proportion of failure event data compared to normal operation data. Such an imbalance presents a major challenge for training classifiers, as it can lead to a bias in the model's predictions, favouring the majority class—often the normal operating condition. This skewed representation reduces the model's ability to accurately identify and predict the minority class, in this case, system failures, which are often critical events. As a result, the classifier may exhibit poor generalisation, failing to detect rare but important failure instances, thereby diminishing its overall predictive performance and reliability in practical applications. Addressing this imbalance is crucial to improving the robustness and accuracy of machine learning models in detecting rare events.

To address this issue, synthetic data is generated specifically for the minority class, which in this study pertains to faulty data samples. Synthetic data refers to artificially contrived information created to safeguard privacy, evaluate the robustness of systems, or establish training sets for the calibration of ML algorithms. In our study, we adopted a widely used method to generate synthetic data specifically for the fault sample category. The feature values of the generated samples are obtained by averaging the corresponding feature values from two adjacent samples. The equation for generating the feature value of the synthetic fault sample is presented below, where V stands for the feature values of synthetic and failure data samples.

$$V_{sd} = \frac{V_{f1} + V_{f2}}{2} \quad (9)$$

By creating synthetic data to balance the dataset, the classifier becomes more adept at accurately predicting both normal and fault conditions, leading to improved overall performance.

5.6.3 The Implementation of the Failure Prediction Framework

A failure prediction framework based on the Semi-supervised GAN has been introduced in this section, which offers the advantage of not requiring prior knowledge about specific building facilities, making it a versatile and data-driven approach for predicting failure times and triggering alarms. By transforming the failure prediction task into a classification task with target labels, the proposed framework can be readily applied to various building assets, including HVAC systems and elevators. Figure 18 illustrates the practical application of this proposed paradigm, showcasing its adaptability and potential for real-world scenarios. The following shows the implementation steps:

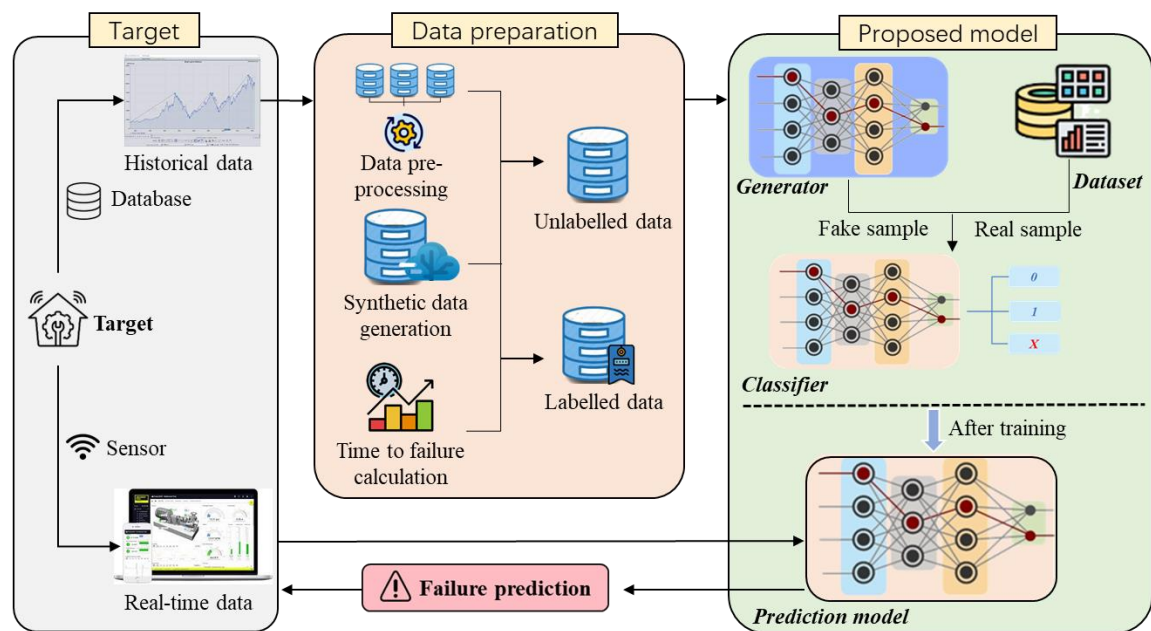


Figure 18. Proposed failure prediction framework based on Semi-supervised GAN

1) Data collection.

The data collection process involves gathering information from multiple sources, utilising advanced technologies such as IoT systems and Building Management Systems (BMS). These systems enable the acquisition of operational data and maintenance records from various building assets. However, a notable limitation is that only a small portion of the collected data can be manually labelled by domain experts

due to the labour-intensive and costly nature of the process. In this study, the dataset is provided by Microsoft, which is collected from 100 building-related facilities such as elevators and air-conditioning systems. The operational data, including metrics such as voltage, rotation, pressure, and vibration, are recorded every hour through the sensor system, while machine attributes are sourced from equipment manuals, and failure events are documented from sourced from maintenance records. Figure 19 illustrates sample data and the corresponding sources.

Time	Machine ID	Model	Age	Volt	Rotate	Pressure	Vibration	Failure
1/1/2015 6:00	1	3	18	176.2179	418.5041	113.0779	45.08769	Type 4

Data Source Legend: Sensor system Equipment manuals Maintenance record

Figure 19. Data sample & Corresponding sources

2) Proposed prediction model.

This step involves data processing and model training. The data processing phase contains three components: preprocessing, synthetic data generation, and time-to-failure calculation. During the preprocessing, data anomalies and errors are systematically identified and removed, while missing values are addressed through imputation techniques. Then, standardisation is utilised to ensure that all variables are on a consistent scale, which is a crucial prerequisite for effective algorithm implementation. Subsequently, the synthetic data generation technique is employed to rectify class imbalance issues, thereby improving the prediction accuracy. Finally, the time-to-failure calculation is a distinctive and pivotal operation in this model, which plays a decisive role in the label assignment of the data points with the help of a warning time setting. Specifically, instances with a value less than the predefined warning time are classified as failure events, whereas those exceeding the threshold are categorised as normal instances.

Leveraging the suitable and balanced dataset derived from the preceding data processing phase, the Semi-supervised GAN methodology is utilised to train the Generator and Classifier networks using both labelled and unlabelled datasets. Within this framework, the labelled dataset can enhance the Classifier's performance directly while the unlabelled data is contributed during the adversarial training process. The Generator aims to generate synthetic samples that closely mimic the distribution and characteristics of the unlabelled dataset, while the Classifier serves as a discerning

adversary, striving to distinguish between the real and fake input instances. Consequently, this collaborative training process enables both the Generator and Classifier networks to understand the underlying distribution and structure of the unlabelled dataset. Therefore, the Classifier can deliver more precise predictions of normal and failure events during the warning period.

Once the training is complete, the Generator network will be disregarded, and the well-trained Discriminator (Classifier) will be extracted and deployed as the prediction model for potential failures. The input data previously served as the training dataset will be replaced by real-time monitoring data, facilitating a transition into real-world implementation, where the model can make valuable contributions.

3) Online implementation.

An online failure prediction system will be developed on the well-trained classifier network from step (2). This system operates by taking real-time operating data from the target system as input and making predictions to determine if an alarm should be triggered. If the predicted target label indicates a failure (value 1), the system triggers an alarm, signalling a potential failure event within the warning period. The online failure prediction system utilises the trained model to monitor the real-time operation of the target system and deliver timely alerts upon predicting a failure, which facilitates proactive maintenance and intervention to mitigate potential risks and optimise system performance.

5.7 Experiment and Validation

This section provides an overview of the critical components of this research, including the introduction of the employed dataset, the methodology of experiment design, and the deployment of the online system, which substantiates the efficacy and practicality of the proposed framework.

5.7.1 Dataset Description

The proposed failure prediction framework based on Semi-supervised GAN is validated using the open-source dataset: *Predictive Maintenance: Exploratory Data Analysis*,

which was collected from 100 facilities between 1st Jan 2015 and 1st Jan 2016. The experiment relies on the utilisation of the following datasets, from which features and fault records have been extracted:

1) PdM_telemetry.csv: This dataset embodies the hourly averages of metrics such as voltage, rotation, pressure, and vibration accumulated from 100 machines over the course of the year 2015.

2) PdM_failures.csv: Each entry within this dataset signifies the substitution of a component due to a failure event. The timestamps in this data are approximated to the nearest hour, paralleling the hourly data collection in the telemetry data.

3) PdM_Machines.csv: Specific attributes of the machines, specifically their model type and respective ages.

To replicate a real-world scenario characterised by the limited availability of labelled data and a vast amount of unlabelled operational data, a deliberate imbalance is introduced in the training sets. The labels of unlabelled datasets are intentionally deleted and treated as unknown. The semi-supervised GAN is trained using both the labelled and the unlabelled training sets, and the failure prediction model is established on the obtained classifier network after training, which takes real-time operational data and generates an alarm.

5.7.2 Experiment Design

To evaluate the efficacy of the proposed method, a comparative analysis involving several widely utilised algorithms is conducted. Supervised learning algorithms include Decision Tree (DT), LR, KNN, SVM, RF, and Neural network (NN), and Semi-supervised learning approaches involving Self-SVM and Self-NN. Self-training is regarded as the most straightforward Semi-supervised learning technique, and the specific steps can be found in Table 16. The proposed approach and the two semi-supervised learning approaches can employ both labelled and unlabelled data sets, while the six supervised learning methods are exclusively trained using the labelled set

to develop the failure predictor. The testing set is constant with the balanced class distribution for each approach to allow for a fair comparison.

Table 16. Schematic of self-training

Algorithm 1 Self-training

```

1: Initialize:
2: Given  $(X_{train}, y_{train}) = (X_l, y_l)$ 
3: While stopping criteria are not met do
4:   Train classifier  $C_{int}$  from  $(X_{train}, y_{train})$ 
5:   Use  $C_{int}$  to predict class label  $y_u$  of  $X_u$ 
6:   Select confidence sample  $(X_{conf}, y_{conf}); (X_{conf}, y_{conf}) \in (X_u, y_u)$ 
7:   Remove selected unlabelled data  $X_u \leftarrow X_u - X_{conf}$ 
8:   Combine newly labelled data  $(X_{train}, y_{train}) \leftarrow (X_l, y_l) \cup (X_{conf}, y_{conf})$ 
9: end while

```

A predicted alarm accuracy is introduced to assess the performance of different approaches, a commonly adopted evaluation metric in this domain. This study aims to provide an alarm before the warning period, assuming that an alert is activated if the label of more than half of the data sample within the specified timeframe is predicted to be 1(fault). The prediction accuracy, defined as the ratio of accurately predicted samples in the testing set to the total number of samples in the testing set, is utilised to gauge the effectiveness of the methods under consideration.

The experiments in this study are carried out by the following guidelines to increase the reliability of the findings:

1) Ten different random seeds are used to train the neural network (or other classifiers) 10 times for each scenario. Ten prediction accuracy can be derived after testing set evaluation. The average of these 10 numbers is the outcome. This objective is intended to lessen the impact of various initialisations.

2) After 10 iterations of creating various datasets (labelled, unlabelled, validation, and testing sets), the various methodologies were assessed by using average outcomes. This objective is intended to lessen the impact of various dataset splits.

3) The grid search methodology was applied in the parameter tuning process of other comparative approaches to optimise the performance. Notably, the neural network

method utilises the same structure as the Discriminator of the proposed model. These settings aim to ensure a fair and meaningful comparison study.

5.7.3 Online Implementation

This section presents the visualisation platform developed using the Streamlit library in Python, which facilitates the establishment of data-centric web applications, as shown in Figure 20. It provides a streamlined Application Programming Interface (API) that empowers developers to efficiently build interactive applications with minimal frontend expertise. The key advantage of Streamlit lies in its hot reloading capability, which automatically refreshes the application upon code changes, eliminating the need for manual intervention. Furthermore, Streamlit enables seamless deployment on various platforms, including cloud services such as Heroku, AWS, or Azure, as well as Streamlit Sharing. The current input data of the system is retrieved from CSV files, while in the future, it can leverage APIs to obtain real-time data and be deployed to the server. Presently, the system provides real-time sensor information, encompassing voltage, rotation, pressure, vibration, and warning signals. An alarm value of 0 indicates normal conditions, while an alarm value of 1 signifies a potential failure within the next warning period (12 hours in this experiment). Additionally, the system incorporates a maintenance record page, allowing operators to document fault occurrence time and fault categories.

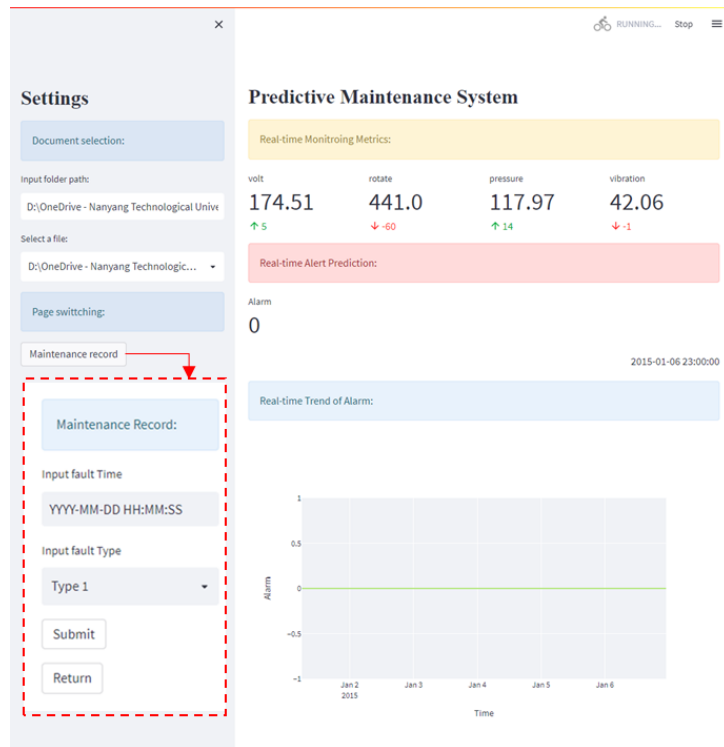


Figure 20. Online platform for PdM

5.8 Result and Discussion

This section presents the experimental results employing various random seeds and explores the effect of label size and the effect of data imbalance ratio.

5.8.1 Evaluation Metrics

In this study, accuracy is utilised as the primary evaluation metric to assess the performance of the proposed method and other benchmark models. The dataset used for model training and testing was collected over a one-year period, with data points recorded at one-hour intervals. The goal of the study is to develop a predictive model that can issue an alarm 24 hours before a system failure occurs, providing valuable early warning to pre-empt critical issues. To achieve this, we define a system alert based on the following logic: if more than 50% of the data samples from any 24 consecutive hours are predicted as 1 (indicating a fault), an alert will be activated. This approach helps reduce false alarms while ensuring that the model provides adequate notice of impending failures.

Table 17. True value & Predictive value

True value \ Predictive value	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (10)$$

As shown in Table 17 and formula (10), the prediction accuracy metric is calculated as the ratio of correctly predicted samples in the testing set to the total number of samples in the testing set. Specifically, if a sample is classified correctly—either as a fault or non-fault—it is considered accurate. This measure allows us to gauge the overall effectiveness of the model in distinguishing between normal and fault conditions. In the context of this study, high accuracy indicates that the model can reliably differentiate between normal operations and potential failures, thus providing actionable insights for maintenance teams to address issues proactively.

5.8.2 Comparison Under Different Random Seeds

According to the average testing accuracy of 10 experiments, this section shows a comparative analysis of the accuracy performances of various algorithms. These experiments were conducted with the same label data size (42 days) and data imbalance ratio (normal: fault = 1:1), utilising various random seeds. As shown in Figure 21, incorporating unlabelled data in the training process significantly enhances prediction accuracy. The proposed method demonstrates a notable improvement in prediction performance, ranging from 10% to 20% under the specified hyperparameter setting. Moreover, when comparing SVM with self-SVM and NN with self-NN, it is evident that the traditional self-training method did not yield enhanced accuracy for this dataset. The observed performance improvement can be attributed to the capacity of Semi-supervised GAN to extract knowledge of data distribution from a substantial number of unlabelled data samples. Consequently, the classifier becomes better equipped to predict occurrences of failure. In comparison to other conventional approaches, the proposed Semi-supervised GAN method exhibits higher accuracy in failure prediction.

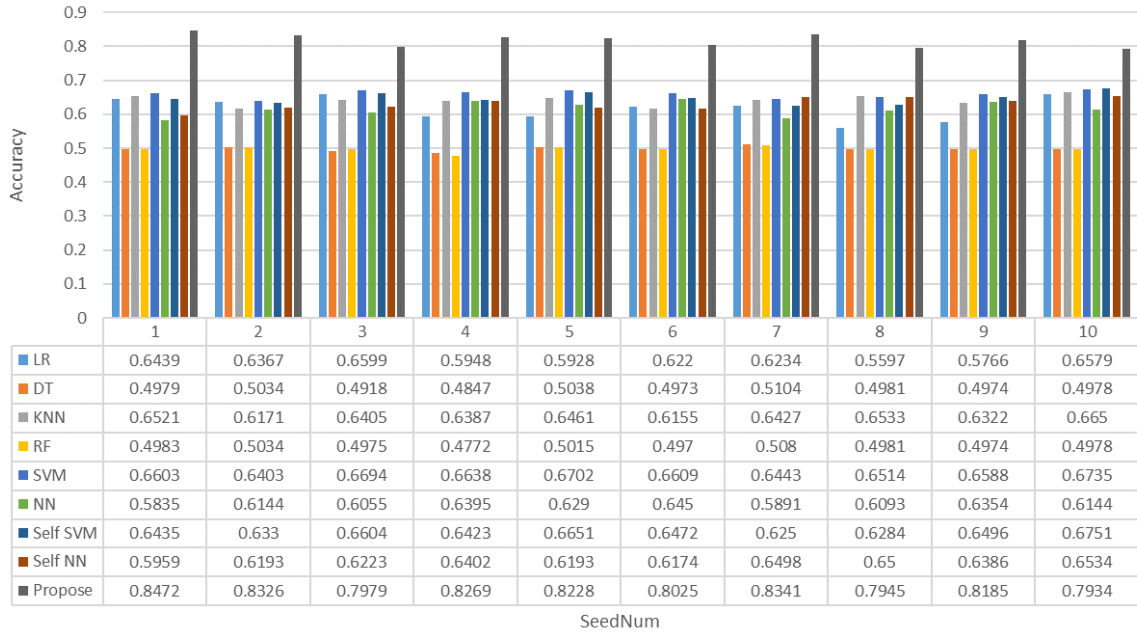


Figure 21. Comparison under different random seeds

5.8.3 Effect of the Labelled Data Size

To further illustrate the effectiveness of the suggested approach, the performance of the suggested approach is also compared with comparison algorithms across various labelled data sizes. In this section, the size of the unlabelled training set remains constant at 250 days, while the size of the labelled dataset is expanded from 7 days to 56 days. From Figure 22, we can observe that the increasing labelled data size can contribute to the improvement of prediction accuracy for most comparison algorithms. Figure 22 depicts the rapid increase in accuracy for the proposed method initially, followed by a tendency to plateau as the size of the labelled dataset increases. The proposed method demonstrates superior accuracy compared with existing supervised and semi-supervised algorithms. With 30 days of labelled samples, the suggested approach achieves an accuracy of nearly 80%, surpassing the approximately 60% accuracy achieved by existing algorithms. Additionally, this experiment highlights that though the proposed method reduces dependence on labelled data, the labelled dataset cannot be excessively small. For this specific dataset, the optimal performance, considering both accuracy and label cost, of the proposed approach is observed when the labelled dataset is 21 days.

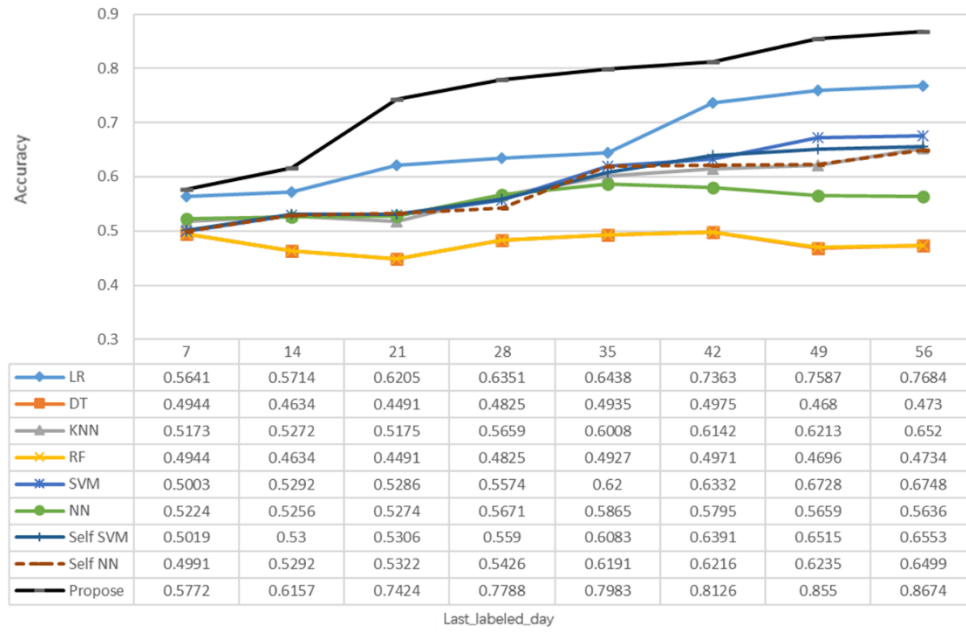


Figure 22. Effect of the labelled data size

5.8.4 Effect of the Data Imbalance Ratio

In addition to varying sizes of the labelled dataset, this study also investigates the performance of the proposed method under different data imbalance ratios. A data imbalance ratio refers to an uneven distribution of classes, which can lead to biased predictions by the classifier, particularly for the majority class. Hence, it is crucial to examine the performance of each technique across different data imbalance ratios.

This section provides a summary of the testing performance for both comparison approaches and the proposed approach under various data imbalance ratios ranging from 1/1 to 1/10. The accuracy values are calculated based on the average results obtained from 10 experiments, employing the same random seed number and labelled dataset size. Figure 23 illustrates that the performance of all approaches is influenced by the data imbalance ratio. As the data imbalance ratio increases, all algorithms eventually exhibit a decline in performance, with predictions tending to random guesses. It is noteworthy that the proposed algorithm consistently outperforms the comparison algorithms in most data imbalance ratio scenarios, except for the 1/10 imbalance ratio case for KNN. In that scenario, the accuracy performance of KNN is found to be 2% higher than that of the proposed method.

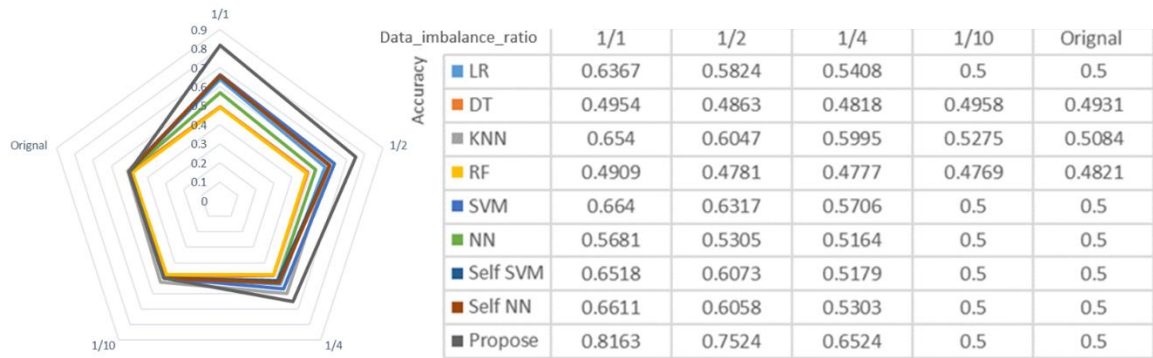


Figure 23. Effect of the data imbalance ratio

5.8.5 Discussion on Adaptability and Scalability

The proposed DT-enabled PdM framework is built upon a data-driven methodology, making it versatile and easy to apply across various building-related facilities, such as air conditioning, elevators, lighting systems, etc. As shown in Figure 24, the data-driven model leverages the vast number of metadata, operation data, and maintenance records collected from the target system to generate intelligent solutions such as pattern detection, failure prediction, and schedule optimisation [135].

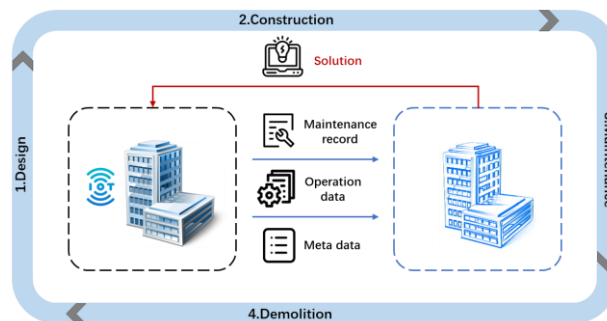


Figure 24. Data-driven method

Compared with traditional model-based approaches, adaptability and scalability are innate advantages brought by the data-driven method. The proposed framework relies on data support rather than domain rules, which can be applied to a wide range of systems, regardless of their complexity or operational differences. This flexibility allows it to continuously learn from new data, improving its predictive accuracy over time. Besides, the applicability and extensibility of the proposed method are also explored and supported by the experiment section.

Section 5.8.1 demonstrates the robustness of the method across different random seed settings, suggesting that the method can be adapted to diverse operational scenarios without being overly sensitive to variations in initial conditions. Section 5.8.2 explores the influence of labelled data size and highlights that the framework can operate effectively with a limited amount of labelled data, which enhances its scalability by reducing the dependence on costly and time-consuming manual labelling processes. Section 5.8.3 further underscores its adaptability, showing that the framework remains the superior performance of other benchmark models even under imbalanced data conditions. These three experiments comprehensively explore the common challenges encountered by solution scalability among various building facilities, particularly the scarcity of labelled data and the issue of data imbalance. In real-world applications, labelling monitoring data typically demands specialised engineering expertise, which is both time-consuming and labour-intensive. Moreover, most datasets collected in practice consist primarily of normal operational data, with fault samples representing only a small proportion.

In summary, the data-driven nature of the proposed framework allows it to be highly adaptable, scalable, and efficient across various facility management systems, offering a significant advantage over traditional rule-based methods.

5.9 Discussion on Practical Implementation

Although open-source data were utilised in this study, the author also conducted real-world experiments to explore and assess the potential feasibility of the implementation.

Table 18. System configuration

Unit	Description
Central Processing Unit	13th Gen Intel(R) Core (TM) i9-13900HX
Graphics Processing Unit	NVIDIA GeForce RTX 4070
Random Access Memory	32 GB
Operation System	Windows 11
Deep Learning Framework	<u>Pytorch 2.0.1</u>

Table 19. Training time comparison (label data size = 42 days, data imbalance ratio = 1:1)

Model	Wall Time	CPU Times Total
Proposed Semi-GAN	10.5 s	1min 4s

To evaluate the feasibility of the solution deployment, the author examined both the model's training time and the cost of the sensor system. Table 18 outlines the details of the implementation environment, showing the hardware configuration information during the experiment. Pytorch 2.0.1 was also selected as the DL framework during the model development. Based on the above hardware setup, Table 19 presents the training time for the proposed Semi-GAN method. The “Wall Time” refers to the total elapsed duration encompassing all operational phases and pauses, while the “CPU Times Total” represents the active computational periods utilised by the CPU, highlighting the efficiency and processing intensity of the training process.

In addition, the sensor system was deployed in a real-world scenario to validate the proposed framework. As depicted in Figure 25, pressure data were collected and transmitted to the laptop platform in real time. For the prototype system, existing laptops can meet the computational requirements, while the primary economic cost is attributed to the sensor system. To achieve more comprehensive monitoring capabilities, vibration, voltage, and current sensors would be necessary, as indicated by the data in the open-source dataset. During actual deployment, data were collected from two air conditioning units over a two-month period. However, there was a significant lack of fault-related information. Furthermore, these sensor devices were highly susceptible to environmental influences during data collection, leading to considerable noise in the data. As a result, the final experiment relied on publicly available datasets for effective model development. It is worth noting that, with advancements in ICT technologies, many air conditioning and elevator systems are now equipped with embedded status detection and data collection systems. Access to such confidential data from relevant organisations or companies would greatly enhance the development and optimisation of the algorithms.

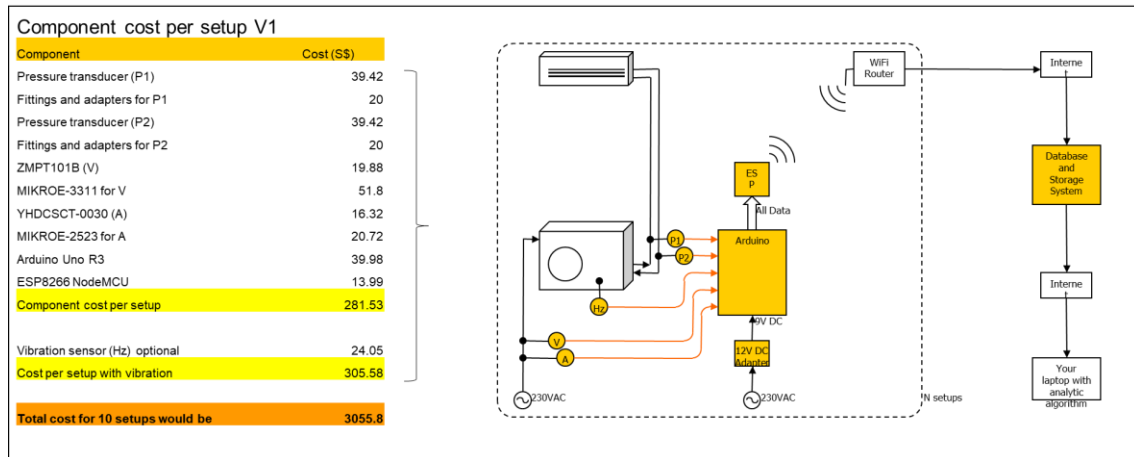


Figure 25. Sensor Deployment Blueprint & Final Plan

5.10 Remark

The adaptability and scalability of DT solutions have been widely recognised across various industries, such as aerospace and healthcare. Emphasising the advantages of employing DT as a platform for PdM in the building industry, this study first developed a unified three-layer architecture showcasing the PdM implementation from the DT perspective. A novel failure prediction approach based on semi-supervised GAN has been proposed for PdM solutions in buildings, which can be embedded in the DT-PdM framework as the failure prediction service provided in functional layers. The proposed method leverages both labelled and unlabeled data to enhance prediction accuracy, outperforming traditional methods by a margin of 10% to 20%, according to the comparative study. Notably, the proposed approach demonstrates the capability to achieve acceptable prediction accuracy with a reduced number of labelled data samples. This reduction in the required labelled dataset size is made possible by the utilisation of Semi-supervised GAN, which effectively harnesses the structural and distribution knowledge embedded within the unlabeled dataset. Furthermore, the experiment explores the accuracy performance of the proposed method across different labelled data sizes and data imbalance ratios. Lastly, a web application has been implemented to present monitoring data and prediction results. The findings highlight the robustness and adaptability of the proposed technique. In practical PdM scenarios, where typically there is a surplus of unlabeled data compared to labelled data, the proposed approach offers a straightforward implementation process with notable benefits. The previous

study mainly focuses on improving the prediction accuracy with the definite dataset, while this research provides a novel perspective to enhance practical implementations by diminishing the dependency on labelled datasets and maximising the utilisation of unlabelled datasets. Besides, the DT paradigm is introduced in this study, offering increased scalability and promising prospects for development. More intelligent algorithms and advanced functions can be developed and integrated by the DT platform, thereby improving the PdM solution comprehensively. These shifts promote the study from accuracy-centric to practicality-driven, and the integration of DT paradigms underscores the potential for future advancements.

Future studies need to enhance the current failure prediction framework to include the prediction of multiclass faults of the target asset. The current study only focuses on the prediction of fault occurrence due to limitations in the available dataset, while the prediction of failure time and failure types hold significant importance for PdM solutions. Besides, it is essential to expand the scope of analysis to encompass concurrent faults that may occur within the system. Multiple and concurrent faults introduce a greater level of complexity, requiring more sophisticated approaches to predict and classify them accurately. Addressing these aspects in future studies would contribute to a more comprehensive and robust failure prediction framework, enabling more accurate and reliable PdM practices.

Chapter 6 A Parallel LSTM-AE Failure Prediction Approach for Indoor Climate

This chapter introduces the emerging field of PdM within the building industry, specifically focusing on its extension into indoor climate management. In this chapter, RUL and TS techniques are employed to partition the prediction problem into supervised and unsupervised tasks, and a parallel prediction model employing the LSTM and AE methodologies is developed to enhance the predictive accuracy and reliability. A practical dataset collected from various buildings in Singapore is employed for the comparative analysis, which illustrates the effectiveness and reliability of the proposed model. Furthermore, this chapter presents a DT-enabled PdM framework customised for indoor climates, which is rigorously validated by creating an online platform designed to reconstruct 3D building models and offer real-time monitoring and alert systems for indoor climates.

6.1 Milieu-Perspective PdM Implementation in Building Industry

Following the Three-by-Three M analysis framework, the ‘Milieu’ pertains to the residential environment in the context of PdM for the building industry, including IAQ, lighting management, temperature control, etc. Unlike the ‘Machine’ field, data-driven approaches have become the normative method for managing milieu-related PdM because milieu-related targets are governed by straightforward principles and mechanisms like internationally recognised standards and protocols. Current research predominantly concentrates on real-time monitoring and anomaly detection within these environmental parameters. Nevertheless, there is a notable opportunity for enhancement in predictive functionalities and alert systems. The potential for improvement in this domain lies in advancing predictive analytics that can proactively signal potential environmental deviations before they become critical issues. Furthermore, integrating DT technology is a transformative prospect for the milieu-focused PdM. A DT system, the virtual replica of the building environment, could enable a more nuanced and dynamic environmental data interaction, leading to more

precise maintenance interventions. This convergence of DT technology with predictive analytics could usher in a new era of PdM, characterised by heightened efficiency, advanced foresight, and an overarching improvement in building environment maintenance.

6.2 Research Requirement and Summary

The assessment and enhancement of IAQ have gained prominence and are currently receiving increased attention within both academic and industry fields because the indoor environment significantly influences the quality of life and work productivity. The common supportive technology for IAQ monitoring and maintenance is IoT systems [368], which also serve as the foundation and supporting technology of the DT framework. Through the establishment of the visualisation model with bidirectional information exchange, the DT implementation enhances PdM strategies with improved interactive capability and optimised planning scheduling. The development of such ICT contributes a consequent amount of real-time data from different measurements (temperature, humidity, PM2.5, etc.) that can be utilised to evaluate, predict, and optimise indoor climate. Also, it promotes a significant transformation in the maintenance industry from the reactive to predictive paradigm, which represents a more proactive approach using real-time data acquisition techniques and ML algorithms to detect and predict failure events for the equipment and the environment [6].

Fault prediction in the PdM domain aims to estimate the probability of future failures within a defined time frame, which aids in planning maintenance activities and optimising resource allocation. Predictive tasks usually demand more rigorous data, sophisticated algorithms, and virtualisation compared to detection, which has led to fewer studies on failure prediction. Hence, harnessing environmental data with advanced algorithms and enhanced interaction platforms holds significant potential and value within the realm of indoor air pollution. In this study, we proposed a paralleled deep learning model for the failure prediction of IAQ, aiming to enhance the automation and reliability of PdM in the indoor climate. Besides, an online platform employing the DT paradigm has been established to facilitate visualisation and interaction.

6.3 DT-enabled PdM Framework for Indoor Climate

Existing PdM frameworks in the building industry mainly emphasise facilities and assets, while the consideration of indoor climates holds equal significance during the holistic PdM strategies execution; thus, it is important to understand the disparities. The framework for facilities encompasses various physical building-related components and systems aiming to address equipment maintenance and operational efficiency, whereas indoor climate pertains to the environmental conditions experienced by occupants centring on occupant comfort, air quality, and energy consumption. Building upon established layered DT structures in the building industry [259], a 3-layer PdM framework from a DT perspective has been proposed to facilitate the DT-enabled PdM implementation for the indoor climate, as illustrated in Figure 26 (up left).

Beginning with the physical-virtual layer, physical building environments should be digitalised through a systematic approach to a virtual space, while indoor environment-related data are collected through IoT devices. This pioneering step towards bidirectional data exchange enables smart communications between devices through the Internet. Unlike the PdM of building facilities, the digitalisation of indoor climate on the virtual side centres on the building environment rather than the target facilities. The virtual building serves as the DT platform, and the maintenance and forecast information is typically displayed through affiliated websites or interfaces. The IoT concept was proposed by Kevin Ashton [369] in 1999 through the achievement of automatic radio receivers utilising RFID tags, while the schematic diagram of the IoT system in this study is shown in Figure 26 (upright). A commercial IoT device named "kaiterra" is configured and deployed using the WiFi connection to systematically gather indoor climate-related information and upload it to the cloud dataset, which is accessible for display on both a dedicated website and a mobile application. For building virtualisation, our previous research has developed a comprehensive BIM reconstruction approach for existing buildings without as-design BIM called "IScan2BIM", and the pipeline is presented in Figure 26 (down). The target building is scanned to generate the point cloud dataset, and then the BIM model is reconstructed by the proposed AI engine, while the details of the structure and MEP reconstruction are introduced in [370] and [325].

Subsequently, the decision layer comprises two crucial processes: data-related and model-related. Required parameters and index values are collected by IoT devices, and then data processing techniques such as data cleaning and data conversion are utilised to enhance the data quality and generate a suitable dataset for the training process. After that, model development utilises various ML and statistical methods, such as regression analysis, time-series forecasting, and neural networks, to create predictive models tailored to specific tasks. The model is trained on historical data, enabling them to identify patterns and correlations within the indoor environment. By integrating real-time data and failure prediction models in the service layer, the PdM system can provide advanced services to stakeholders, including but not limited to failure alerts, proactive maintenance actions, energy consumption optimisation, and enhanced occupant comfort, among other benefits.

DT platform offers significant advantages for PdM in terms of indoor climate. By creating a virtual replica of the physical building environment, DT technology facilitates dynamic, real-time monitoring and in-depth analysis of multiple critical air quality parameters, and this continuous monitoring data streaming allows for proactive identification of potential issues, leading to more prompt anomaly detection and precise failure predictions. DT framework also offers the capacity to integrate diverse data sources, including sensor-generated information, operational metrics, and environmental variables. This real-time data integration augments the accuracy of failure prediction models and contributes to a comprehensive understanding of the intricate interplay of factors affecting indoor climate.

Moreover, with the ability to simulate and model the building's behaviour under different conditions, DT facilitates the execution of virtual simulations and meticulous "what-if" scenarios, enabling maintenance teams to assess the impact of different maintenance strategies and interventions before implementing them in the physical environment. This pre-emptive analytical prowess plays an instrumental role in the systematic evaluation of alternative maintenance plans, thereby affording an elevated level of precision in optimising decision-making processes. These meticulous assessments significantly improve maintenance procedures, leading to a more efficient and judicious approach to management and maintenance.

Furthermore, DT technology facilitates inherent scalability for failure prediction function by providing a comprehensive view of the asset's situations. The results of failure prediction are easily integrated with other PdM applications, such as FM, safety analysis, and emergency management. By providing real-time insights and simulations, DTs empower organisations to manage complex systems with greater precision and efficiency. This scalability is especially valuable in large-scale industrial environments, where many interconnected systems and components need to be continuously monitored and analysed for potential failures.

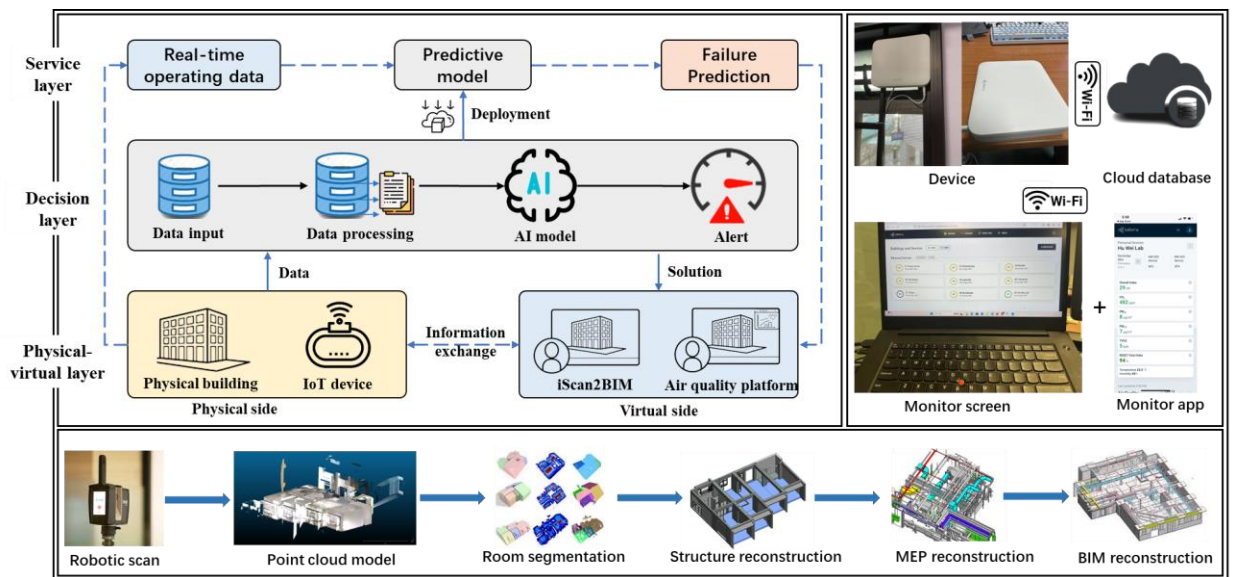


Figure 26. The framework of DT-enabled PdM for IAQ (up left), Prototype of hardware system (up right), iScan2BIM pipeline (down)

Overall, DT adoption in the PdM domain for indoor climate improved occupant comfort and well-being, mitigate potential risks and streamlined maintenance operations, thus contributing to more sustainable and intelligent management of indoor climate practices.

6.4 LSTM and AE

The principles of LSTM and AE are introduced in this section to provide a well-structured theoretical foundation for the proposed model.

6.4.1 LSTM

LSTM is a specialised RNN architecture designed to address the challenges of modelling long-term dependencies in sequential data. LSTM provided the capability of 'memory', enabling them to retain and utilise important information over previous time intervals, which allowed LSTM to preserve essential context and carry relevant knowledge through time, making it highly effective in capturing and utilising temporal patterns in sequences.

As shown in Figure 27, a standard LSTM unit comprises cells and a sophisticated gating mechanism consisting of forget input and output gates. Unlike traditional RNN, LSTM has two transfer stats: the cell state and the hidden state. The previous cell state (c^{t-1}) represents the current long-term memory of the network, which preserves previous information in sequence while the previous hidden state (h^{t-1}) can be interpreted as short-term memory, which is the output at a previous time point. The cell state changes slowly by adding new slices, while the hidden state can be very different at different positions. Details are as follows, and related equations are listed in Table 20:

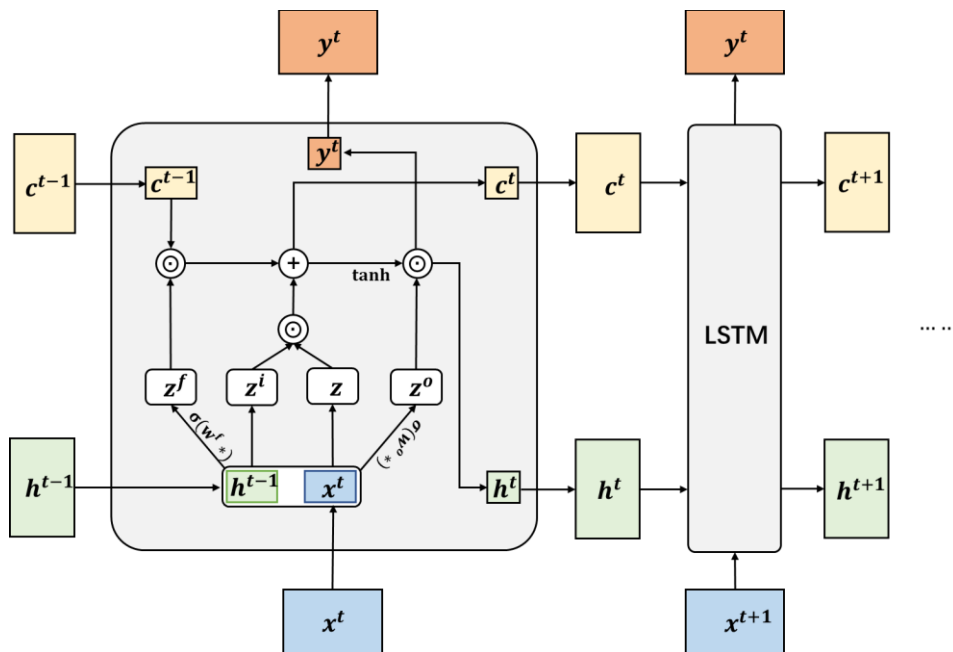


Figure 27. LSTM structure

1) Forget stage: Forget gate (z^f), as Eq. (10), is utilised to modify the information from the previous cell state (c^{t-1}), storing important information and dropping the useless.

2) Selective memory stage: Input gate (z^i), as Eq. (11), is used to conduct selective memory for the new information containing the previous hidden state (h^{t-1}) and new input data (x^t), valuable parts are selected and integrated into the cell state. The extracted information is presented as \vec{c}^t and the process is depicted in Eq. (12). The outcomes derived from the above two stages are combined to create the cell state (c^t) as Eq. (14), which is transmitted to the next time slot.

3) Output stage: Following Eq. (14), the LSTM determines the current output from the cell state using the output gate (z^o), and a new hidden layer state h^t is created with the new cell state c^t as shown in Eq. (15). The output y^t in LSTM is typically obtained by applying a transformation to the h^t like regular RNNs, illustrate in Eq. (16).

Subsequently, the cell state (c^t) becomes transitions to the previous cell state (c^{t-1}) and the hidden state (h^t) becomes a previous hidden state (h^{t-1}) in the next units. This iterative process persists until all input data are handled across all LSTM units.

Table 20. Equations explanation in LSTM

Stage	Function	Equation
Forget stage	Forget gate	$z^f = \sigma(W^f [x^t, h^{t-1}])$ (11)
Selective memory stage	Input gate	$z^i = \sigma(W^i [x^t, h^{t-1}])$ (12)
	Valuable extraction	$\vec{c}^t = \tanh(W^i [x^t, h^{t-1}])$ (13)
	New cell state	$c^t = z^f \odot c^{t-1} + z^i \odot \vec{c}^t$ (14)
Output stage	Output gate	$z^o = \sigma(W^o [x^t, h^{t-1}])$ (15)
	New hidden layer state	$h^t = z^o \odot \tanh(c^t)$ (16)
	Output	$y^t = \sigma(W^o h^t)$ (17)

For the above equations, σ is the activation function, x^t , h^t , and c^t refer to the current input, hidden state and cell state, respectively. W^f , W^i , and W^o represent the weights of forget, input, and output gates. \odot is the Hadamard product.

6.4.2 AE

The AE is a fundamental neural network architecture used in unsupervised learning tasks which aims to capture the most salient features and learn the meaningful representation of the input data. A typical AE structure contains one input layer, one output layer, and several hidden layers, and the training process is as follows:

1) Encoder: The encoding process aims to map the high-dimension input data to a low-dimension representation known as latent space or bottleneck, which is achieved by

progressively reducing the number of neurons in each network layer. This compressed representation is a condensed version of the original data, capturing its essential features more concisely and informatively.

2) Decoder: The objective of the decoder is to reconstruct the input data based on the compressed representation obtained from the encoder. The structure of the decoder is composed of hidden layers, and the output layer has an incrementally increased number of neurons.

3) Reconstruction: After the encoder-decoder procedure, reconstruction loss shown in Eq. (18) is calculated to minimise the difference between the input and output, which can be referenced with threshold setting and utilised for anomaly detection tasks [371].

$$L = \frac{1}{n} \sum_{n=1}^n |\hat{x}_t - x_t| \quad (18)$$

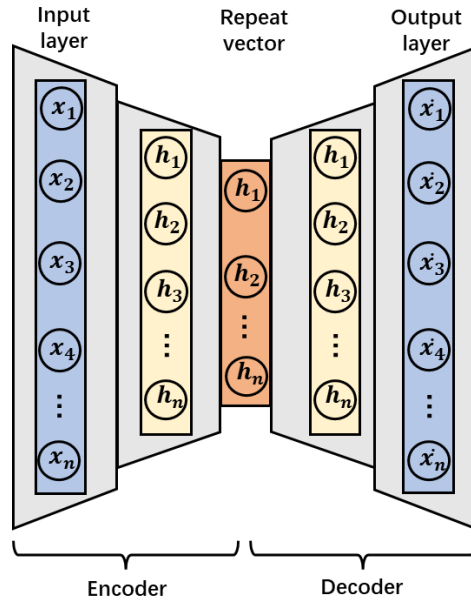


Figure 28. AE

4) Objective function: The objective, as expressed in Eq. (19), is to optimise the encoder (e) and decoder (d) by minimising the reconstruction loss, thereby facilitating the feature representation specific to the input data.

$$e \text{ or } d = \operatorname{argmin}_{e \text{ or } d} L(x, f(e(x))) \quad (19)$$

6.5 Parallel LSTM-AE Failure Prediction Method

This section presents a comprehensive exposition of the developed methodology, encompassing problem formulation, data preprocessing, and the proposed model.

6.5.1 Problem Formulation

In the indoor environment domain, the overall index of indoor climate should be maintained within a normal range to ensure the well-being and optimal performance of occupants and equipment. Otherwise, deviations from this normal scope can be seen as failure events due to poor air quality conditions. Therefore, the consistent monitoring and assessment of IAQ are essential aspects of PdM to promptly identify and mitigate any adverse effects on health, comfort, and operational efficiency. Our foundation of PdM for indoor climates centres on the notion that the data gathered by the IoT device will show certain indicative patterns before any failure occurrence. In the study, IoT data have been collected on a minute basis, from which we extracted features and formed the data sample. Based on the overall index collected and the United States Air Quality Index (USAQI) categorisation [372], as shown in Figure 29, we can label each data point as normal or failure.

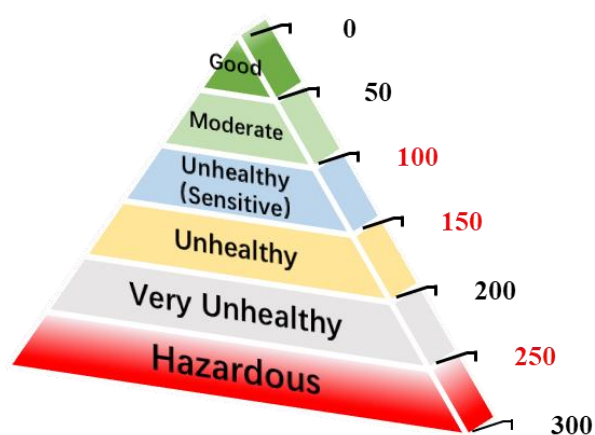


Figure 29. USAQI values and categories

The Overall Index (US) provides a holistic view of the air quality. Utilising the same methodology and breakpoints as the Environmental Protection Agency's air quality

inspection (AQI) standard [373], an index is calculated for each individual pollutant, and the highest index is displayed as the Overall Index. The index conversion algorithm is a piecewise linear function of the pollutant concentration. There is a discontinuous jump of one AQI unit at the boundary between Good and Moderate AQI categories, such as the boundary between Good and Moderate AQI values. To convert from concentration to AQI, the below equation is used, where I is the resulting index value, C is the pollutant concentration, C_{low} and C_{high} refer to the concentration breakpoint below and above C , respectively, I_{low} and I_{high} stand for the index breakpoint corresponding to C_{low} and C_{high} .

$$I = \frac{I_{high} - I_{low}}{C_{high} - C_{low}} (C - C_{low}) + I_{low} \quad (20)$$

Moreover, TS and RUL approaches have been utilised to process the acquired data, which can map the failure prediction problem to anomaly detection and classification tasks that can be solved by existing DL methods. The TS method is a prominent approach employed in prediction problems, particularly in time series analysis, where the goal is to forecast future values based on historical data. As depicted in Figure 30, by leveraging the information of predictor variable (X) at the immediate past ($t-1$) and the target variable (y) at the present (t), it capitalises on temporal dependencies within the data, enabling the model to capture patterns and trends that may influence the future outcomes. The RUL is another significant approach in the prognostic domain, which represents the time duration before failure happens for each data point. The calculation starts with the identification of the faulty point during the monitoring period. Subsequently, the time to failure is generated for every normal data sample through the timestamp subtraction of the next failure point from the corresponding normal sample. The principle is shown in Figure 31, and the equation is expressed as Eq. (20).

$$T_{rul} = T_{normal} - T_{failure} \quad (21)$$

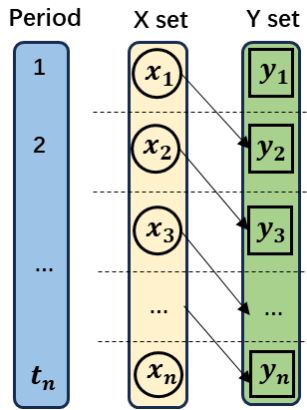


Figure 31. TS

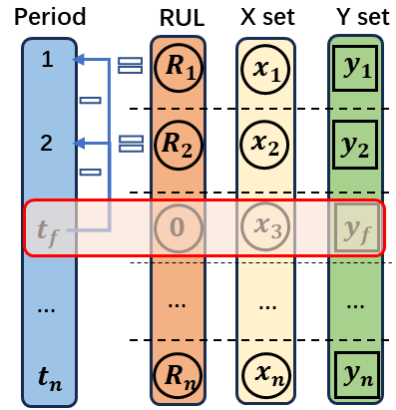


Figure 30. RUL

6.5.2 Data Preprocessing

Data preprocessing is a crucial phase in the data analysis pipeline, enhancing the quality and usability of the dataset. This section outlines several critical steps involved in our study:

1) Data cleaning: In this step, errors and irrelevant data arising during the collection process are removed to enhance dataset integrity. Notably, feature correlation analysis, as illustrated in Figure 32, demonstrates a substantial correlation between PM1.0 and PM2.5. Consequently, PM1.0 was removed because PM 2.5 is more informative and meaningful for future analysis.

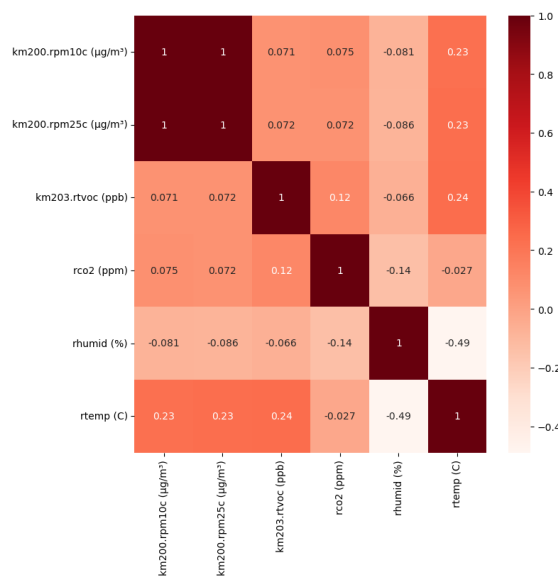


Figure 32. Feature correlation analysis

2) Data imputation: Data imputation techniques are applied to estimate or fill in the missing data, which is vital for maintaining completeness and accuracy. To address this issue, the average value of adjacent data points is utilised to generate the missing values through the following equation, where V_m refers to the missing value while $V_{m\pm 1}$ is the neighbour samples.

$$V_m = \frac{V_{m-1} + V_{m+1}}{2} \quad (22)$$

3) Standardisation: Standardizing the data involves rescaling or transforming variables to have a standard scale, often with a mean of zero and a standard deviation of one. It ensures that variables with different units or ranges do not unduly influence the results. For the numerical features, the standardisation values are calculated using the following equation (23).

$$y_i = \frac{2(x_i - \min_{1 \leq j \leq n} \{x_j\})}{\max_{1 \leq i \leq n} \{x_j\} - \min_{1 \leq i \leq n} \{x_j\}} - 1 \quad (23)$$

4) Data Splitting: The dataset is split into training and testing sets. The training set is used to develop models, while the testing set is reserved for evaluating model performance. This step is integral to assess the model's generalisation ability to the new dataset.

6.5.3 Proposed Model

1) Warning period: Warning time plays a crucial role in addressing accurate prediction, which refers to the duration between the identification of a potential failure event and the actual occurrence of the failure. Having an appropriate warning time allows operators and maintenance personnel to take necessary preventive actions, such as conducting maintenance activities or implementing operational changes, to avert or mitigate the consequences of the impending failure. To tackle the warning time, this study adopts a two-fold approach. Firstly, the RUL is utilised in the LSTM component of the parallel model. The data samples are labelled based on RUL values, where samples with RUL less than the defined warning period are assigned 1, indicating an impending failure within the next warning period. Secondly, the AE component

employs the TS method, wherein the shift value is aligned with the warning time, enabling the model to effectively analyse the anomalies in advance and provide insights into deviations from the normal operating behaviour, thereby contributing to the failure prediction.

2) Parallel model: The overview of our model is illustrated in Figure 33, which is a parallel integration of LSTM and AE methods designed to enhance the accuracy of failure prediction for indoor climate. The integration of these two components allows the model to leverage predictor variables (e.g., humidity, temperature) through the LSTM and past target variables (Overall index) by AE. This concurrent utilisation of predictors and targets provides a comprehensive understanding of the data dynamics and contributes to the model's superior performance. Furthermore, the combination of LSTM and AE enables the model to capture complex temporal patterns and underlying dependencies in the data, leading to more robust predictions and precise identification of anomalies. Overall, the parallel model offers a promising solution for addressing anomaly prediction tasks, leveraging the strengths of LSTM and AE techniques to achieve enhanced performance and utility in real-world applications.

3) AE-LSTM: The AE-LSTM has been utilised as the AE component because it is more suitable to handle sequential and time-dependent data than traditional AE. As shown in Figure 34, the 'overall index' from the original dataset is processed with the TS method first, and then the LSTM encoder establishes numerous LSTM units to examine the features and iteratively transform the input vector from high-dimension into the low-dimension format, eventually converging within the latent space. After that, the LSTM decoder reconstructs the original input from the compressed form in the latent space. Lastly, the evaluation of the reconstruction loss is performed to ascertain an optimal threshold value, which is subsequently employed in the test set to identify potential failures during the subsequent warning period.

4) Integrated prediction: Predictions from LSTM and AE components are integrated into the ultimate prediction employing a tandem approach of gradient boosting (GB) and logical disjunction (LD). GB amalgamates the predictions progressively from two individual models through an iterative optimisation process and nuanced weight adjustments, while LD aligns the final prediction with positive if at least one prediction

from the LSTM or AE component is positive. Empirical findings show that GB exhibits higher accuracy in the prediction of lower-severity anomalies, whereas LD demonstrates superior performance for anomalies of higher severity. The concurrent deployment of RUL and TS within the parallel model effectively leverages both predictor and target information and addresses the challenge of pre-emptive warning times in failure prediction, which can help operators and maintenance personnel make timely and informed decisions, facilitating proactive maintenance strategies and ensure a safer and more reliable operation of industrial systems.

6.5.4 Model Analysis

The proposed parallel AE and LSTM deep learning method for failure prediction in indoor air quality leverages the strengths of both models to enhance prediction performance. Each model presents distinct advantages, and integration methods like gradient boosting or logical disjunction optimise the final prediction output through strength combination.

AE Contribution:

AEs are unsupervised neural networks proficient in capturing patterns in high-dimensional data and detecting anomalies, particularly when the input deviates significantly from normal operating conditions [133]. In this context, the AE is highly effective in detecting sudden and serious faults, as such faults typically manifest as anomalous patterns in the data that deviate from normal indoor air quality conditions. The ability of AEs to reconstruct data and minimise reconstruction error allows them to raise alarms when inputs cannot be accurately reconstructed, indicating abnormal conditions. This characteristic accounts for the higher performance of AE for severe and sudden faults, as such anomalies are markedly distinct from the typical data patterns the model has learned.

LSTM Contribution:

LSTM networks are particularly powerful for time series prediction due to their ability to capture long-term dependencies in sequential data [374]. LSTM networks are designed to model temporal patterns and can track subtle trends or gradual deviations over time. In this situation, this is beneficial for capturing faults or system degradation

that develop slowly and progressively. However, LSTMs might underperform in detecting sudden faults because they rely heavily on sequential relationships, which can cause them to "smooth out" anomalies that appear abruptly and deviate sharply from the trend.

Integration Optimization:

The integration of AE and LSTM predictions optimises the overall model performance by leveraging the unique strengths of both models. In cases of sudden, severe faults, where AE's high sensitivity to anomalous data patterns is crucial, the integration ensures that such critical issues are captured by incorporating AE's alarms. Besides, for gradual faults or system drift, LSTM's ability to model temporal dependencies ensures that subtle, long-term degradation trends are identified, preventing the system from missing slowly developing failures.

By combining AE and LSTM outputs utilising gradient boosting or logical disjunction, the framework creates a more robust and balanced predictive system. Gradient boosting can adjust the contribution of each model based on their predictive accuracy in different scenarios, while logical disjunction ensures that if either model flags an issue, it is incorporated into the final prediction, improving the system's overall fault detection capability across a range of fault types. This hybrid approach enhances both the responsiveness to sudden faults and the ability to capture long-term trends, making the method highly effective for time series prediction in indoor air quality systems.

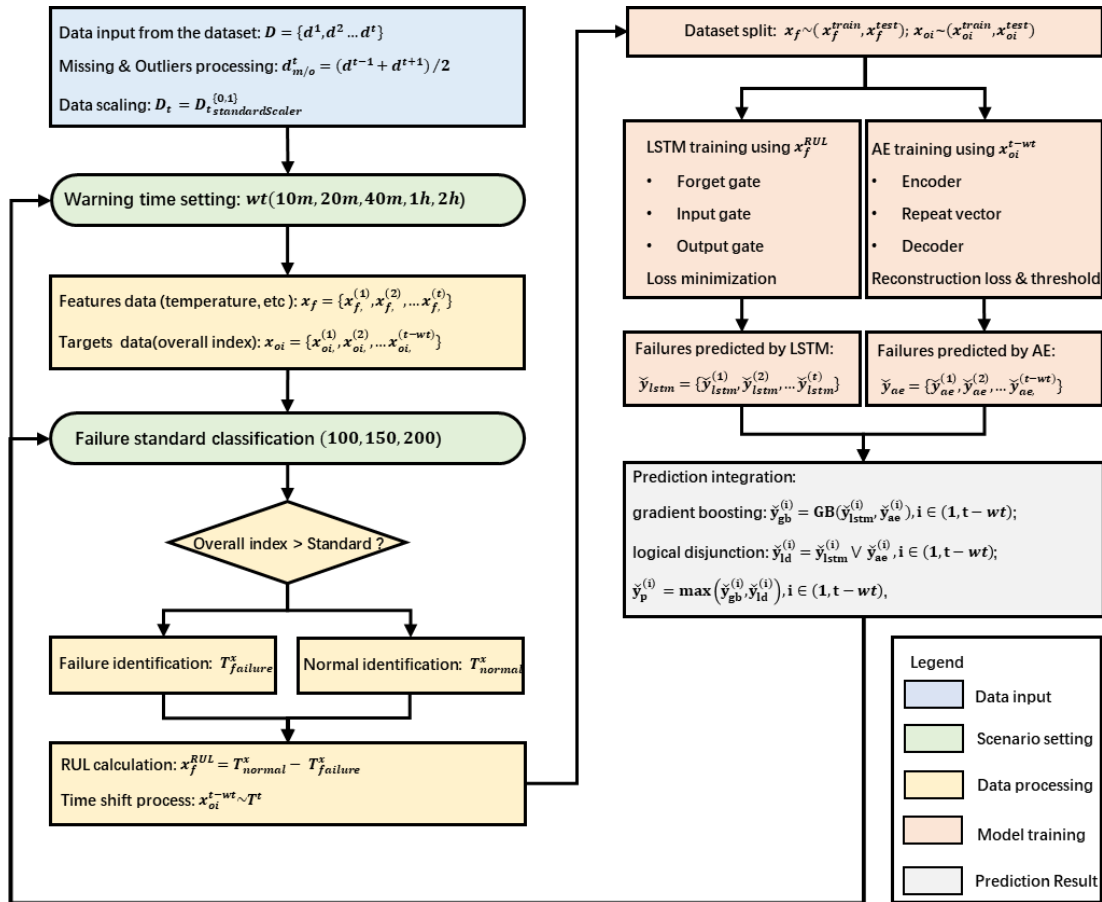


Figure 33. The schematic diagram of the proposed model

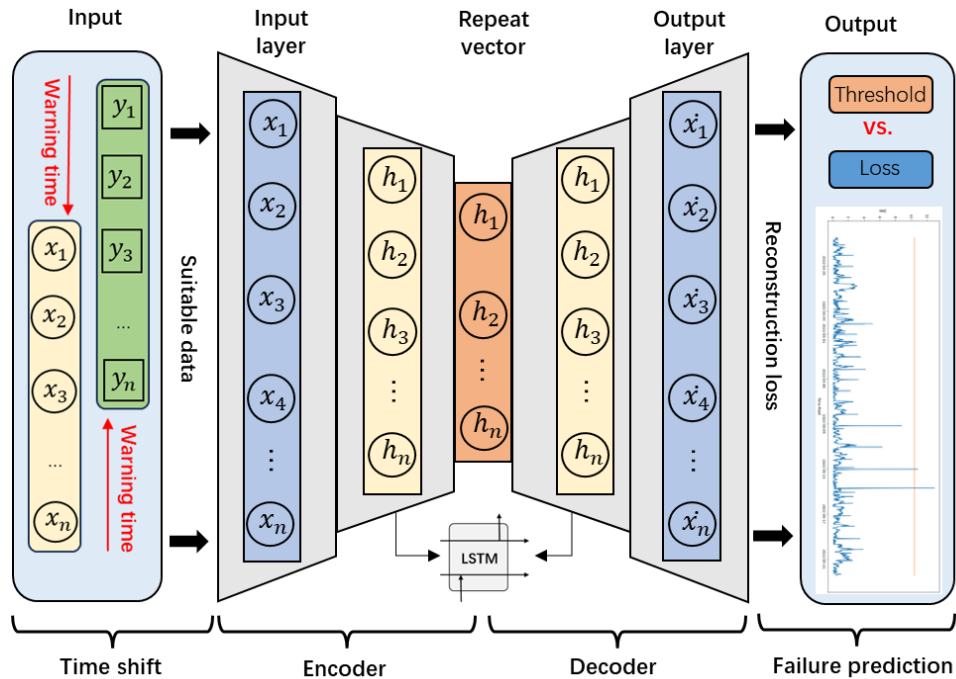


Figure 34. LSTM-AE model

6.6 Experiment and Validation

This section presents the practical dataset utilised for model development and the experimental design methodology.

6.6.1 Dataset Description

To ensure its broad applicability across various indoor environments, the dataset utilised for solution development in this study encompasses a diverse and comprehensive IoT dataset collected from different types of buildings, including commercial complexes, residential apartments, and educational institutions at various locations across Singapore, as shown in Figure 35. The dataset spans one year (2022.01.22 – 2023.01.22), capturing continuous measurement at a high frequency of one-minute intervals. The rich temporal and spatial information present in the dataset enables the exploration of intricate patterns and variations in indoor climate parameters, including temperature, humidity and CO_2 , etc.

Particulate Matter (PM) exists in various sizes and originates from sources like dust and pollutants. PM 2.5 and PM 1.0 refer to particles of different diameters and are of particular concern due to potential health impacts. Total Volatile Organic Compounds (TVOC) reflect the cumulative presence of indoor organic compounds originating from materials, cleaning products, and more. CO_2 concentration serves as an indicator of ventilation adequacy and its associated effects on occupants' well-being. Relative humidity denotes moisture content relative to the air's capacity at a given temperature and influences comfort and mould prevention, while temperature determines the warmth or coldness of the indoor environment. Elevated PM, TVOC, and CO_2 levels often signal poor ventilation, emphasising the importance of temperature and humidity control for occupant well-being. Monitoring and managing these aspects are integral to fostering a healthier, more comfortable indoor climate. This extensive and meticulously collected dataset is a valuable resource for developing and evaluating advanced PdM algorithms and solutions that can significantly enhance the PdM implementation for indoor climate and occupant comfort in various building settings.

In this study, a 4-month dataset of Lakeside station (black box in Figure 35) consisting of 172,800 data points is utilised with the consideration of the equilibrium performance of computational efficiency and experiment effectiveness. The training set spans three consecutive months, while the testing set comprises one month. Figure 36 provides an overview of the predictors and targets in the test set, and Table 21 presents their descriptive statistics. An examination of these statistics reveals that PM, TVOC, and CO₂ concentrations persistently maintain a baseline of low values for most of the recorded periods, interspersed with abrupt and conspicuous elevations. This characteristic pattern also extends to the distribution of the overall index, underscoring the pivotal influence of the above variables. Furthermore, humidity exhibits a consistent and even fluctuation centred around its mean value, whereas temperature demonstrates a gradual increase over time.

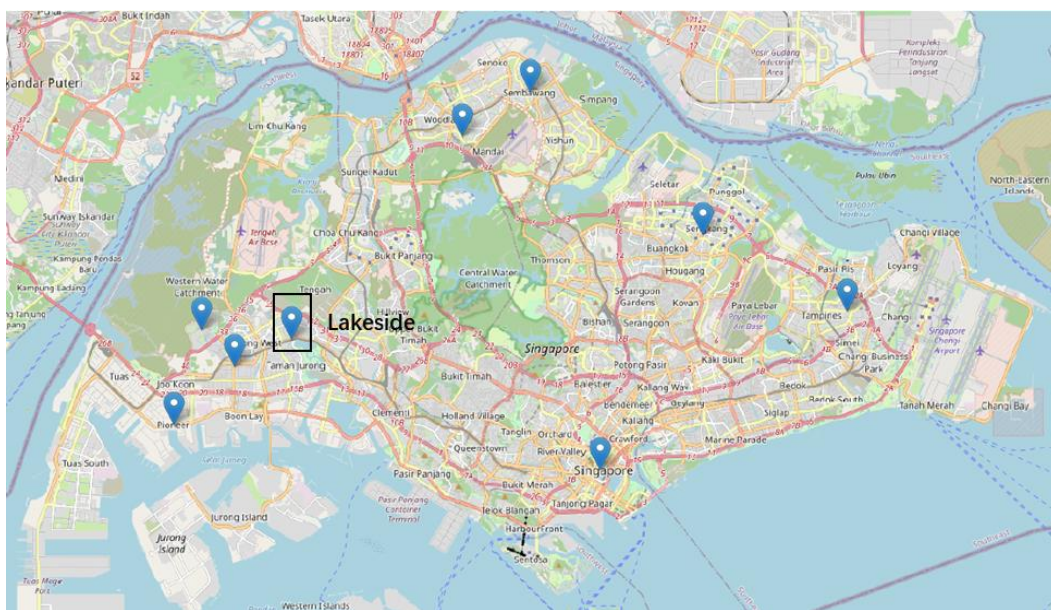


Figure 35. The location distribution of IoT devices

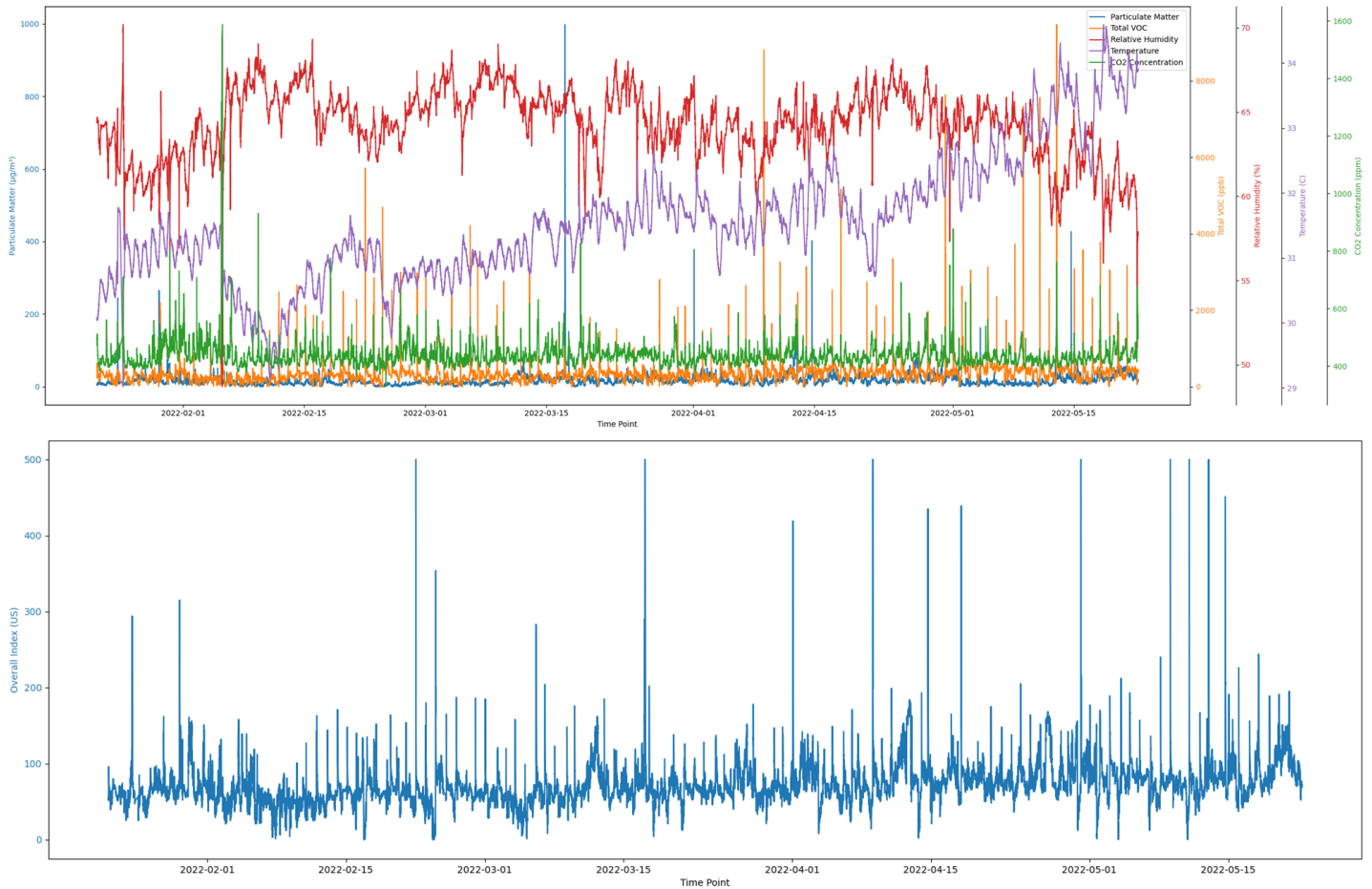


Figure 36. Predictors and targets of the test set

Table 21. The distribution metrics of the test set

Measurement	(Max, Min)	Mean	Median	Std Dev
(1) km200.rpm25c ($\mu\text{g}/\text{m}^3$)	(998,0)	17.06	14.00	18.76
(2) km203.rtvoc (ppb)	(9476,0)	335.71	309.00	225
(3) rhumid (%)	(70.19,48.7)	64.26	64.48	2.03
(4) rtemp (C)	(34.59,29.02)	31.56	31.46	0.94
(5) rco2 (ppm)	(1587,329)	437.04	429.00	47.16
(6) Overall Index (US)	(500,0)	72.25	68.00	26.7

6.6.2 Experiment Design

This section provides the experiment details, including experiment setup, evaluation metrics, and scenario setting.

Experiment Setup:

The specifics of the implementation environment are elucidated in Table 22, delineating the system configuration information during the model development. The hyperparameter tuning process was conducted systematically to ensure the optimal

performance of the model. This process involved evaluating various candidate values for each hyperparameter, as detailed in Table 23. To determine the optimal configuration, the model's performance was assessed using a grid search and cross-validation approach. In this method, different combinations of hyperparameters were evaluated by training the model on subsets of the training data, and performance metrics, including accuracy, precision, recall, and F1-score, were used to measure the model's effectiveness.

Table 22. Implementation environment

Unit	Description
Processor	13th Gen Intel(R) Core (TM) i7-13700K
RAM	32 GB
OS	Linux Office-server 5.4.0-148-generic
Packages used	TensorFlow 2.1.0

Table 23. Hyperparameters tuning

Hyperparameters	Description	Range
Learning rate	Step size for optimisation convergence	0.001, 0.01 , 0.1
Epoch number	Iteration count for training cycles	20, 50 , 100
Batch Size	Sample count processed per optimisation step	16 , 32, 64
Dropout	Neurons deactivated to prevent overfitting	0, 0.1 , 0.2
LSTM/AE layer & unit	Structure of neurons network	8, 16, 32 , 64
Optimisation function	Adjustment methods to minimise loss	Adam , SGD

Additionally, the model's behaviour was monitored across multiple iterations to ensure robustness against overfitting or underfitting. The final selection of hyperparameters was based on the configuration that consistently yielded the best performance on the validation data while balancing complexity and computational efficiency. Key considerations included learning rate, epoch number, batch size, and the network structure, ensuring that the model could generalise well to unseen data while maintaining computational feasibility. This approach allowed for the identification of hyperparameters that provided optimal predictive accuracy and stability in the experimental results.

Evaluation Metrics:

Accuracy, precision, recall, and F1 score are chosen for validating the proposed method, which is widely utilised in classification problems. Each metric addresses a unique perspective of model behaviour, ensuring that the evaluation captures the general aspects and other critical prediction scenarios, such as fake alarms and missed alarms, which are highly relevant in failure prediction contexts. As illustrated in Table 24, True

Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) are fundamental to the calculation of the performance metrics, providing the quantitative basis for the evaluation of model performance.

Table 24. Definition of TP, TN, FP, and FN

Term	Definition	Practical meaning
TP	Instances correctly classified as positive.	True failure.
TN	Instances correctly classified as negative.	True normal.
FP	Instances wrongly classified as positive.	Fake failure.
FN	Instances wrongly classified as negative.	Missed failure.

Accuracy measures the overall correctness of classifications by calculating the proportion of true predictions, including both true positives and true negatives, relative to the total number of predictions. While it provides a general indication of the model's performance, accuracy is insufficient in the PdM context, where the class imbalance between normal operations and failure events is prevalent. In such cases, high accuracy can be achieved by predominantly predicting the majority class, which is misleading. Therefore, additional metrics are required to evaluate the model's performance more precisely, particularly in detecting minority class instances such as different severity of failures.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (24)$$

Precision is chosen to assess the model's capacity to accurately identify positive instances, specifically actual failures, by calculating the proportion of true positives among all predicted positives. This metric is crucial in scenarios where minimising the cost of false positives is a priority, such as unnecessary maintenance actions or system downtime. Precision ensures a high probability of accuracy when the model predicts a failure, thereby reducing the occurrence of unnecessary maintenance interventions.

$$Precision = \frac{TP}{TP+FP} \quad (25)$$

Recall evaluates the model's ability to detect all actual positive instances (failures) by calculating the proportion of true positives among all actual positives. In the PdM area,

missing failure (false negatives) can result in significant operational risks and unexpected breakdowns. Therefore, recall is crucial for evaluating how effectively the model captures all failure events, even at the risk of increasing false positives.

$$Recall = \frac{TP}{TP+FN} \quad (26)$$

The F1 score was chosen to balance precision and recall, as it represents the harmonic mean of these two metrics. This makes it particularly valuable in scenarios where there is a trade-off between precision and recall, such as in imbalanced datasets or when both false positives and false negatives incur significant costs. By integrating the strengths of precision and recall, the F1 score provides a comprehensive assessment, ensuring that neither metric outweighs the other. This balance is crucial for developing a reliable failure prediction model, where both minimising false alarms and accurately identifying failures are equally important for optimal performance.

$$F1 - Score = \frac{2*TP}{2*TP+FP+FN} \quad (27)$$

By incorporating these four evaluation metrics, the author ensures that the validation of the proposed model reflects a well-rounded and thorough analysis. Each metric serves a specific purpose in assessing various performance aspects of the model, from overall accuracy to its ability to handle critical maintenance decisions. This multi-faceted evaluation framework contributes to the development of accurate and reliable failure prediction models, supporting informed decision-making and ensuring that the model can be effectively applied to diverse failure prediction and PdM scenarios.

Scenario Setting:

The experiment across diverse scenarios containing a range of anomaly severities (100,150, 200) and warning time (10m, 20m, 30m, 1h, 2h) has been conducted to provide a comprehensive evaluation. By incorporating hyperparameter tuning, performance metrics, and diverse experiment scenarios, the experimental design not only ensured a meticulous assessment of the proposed method's effectiveness but also contributed to a robust and insightful analysis of its performance under varying conditions.

6.7 Result and Discussion

This section presents the exposition and discourse on the experiment results of the proposed models. We aim to validate its ability to provide accurate alarms for potential failure and assess its performance on a broader scale through different scenario settings. Furthermore, we explore the practical challenges encountered during the experimentation process, aiming to provide insights into refining failure prediction strategies.

6.7.1 AE for Failure Prediction

AE has been proven to be an effective technique for anomaly detection processes among various domains [375], where instances that deviate significantly from normal patterns are identified and labelled as anomalies. By integrating TS techniques, AE can be trained using this augmented dataset to learn normal patterns, thereby achieving the failure prediction function by recognising deviations when the associated errors exceed the predefined threshold. Setting an appropriate threshold is pivotal for the task implementation, and reconstruction loss is commonly utilised [123]. In this study, the threshold is established by analysing the distribution of reconstruction loss, as shown in Figure 37, which is imperative to consider and balance the values of the evaluation metric alongside the proportion of the target outlier. Upon employing the threshold-based failure prediction methodology, it becomes essential to adjust the threshold according to distinct severity settings, owing to the differing magnitudes of deviations within anomalous features. The proposed model is trained with the "Overall index" dataset, and Figure 38 presents the failure distribution of three different severities in the test set.

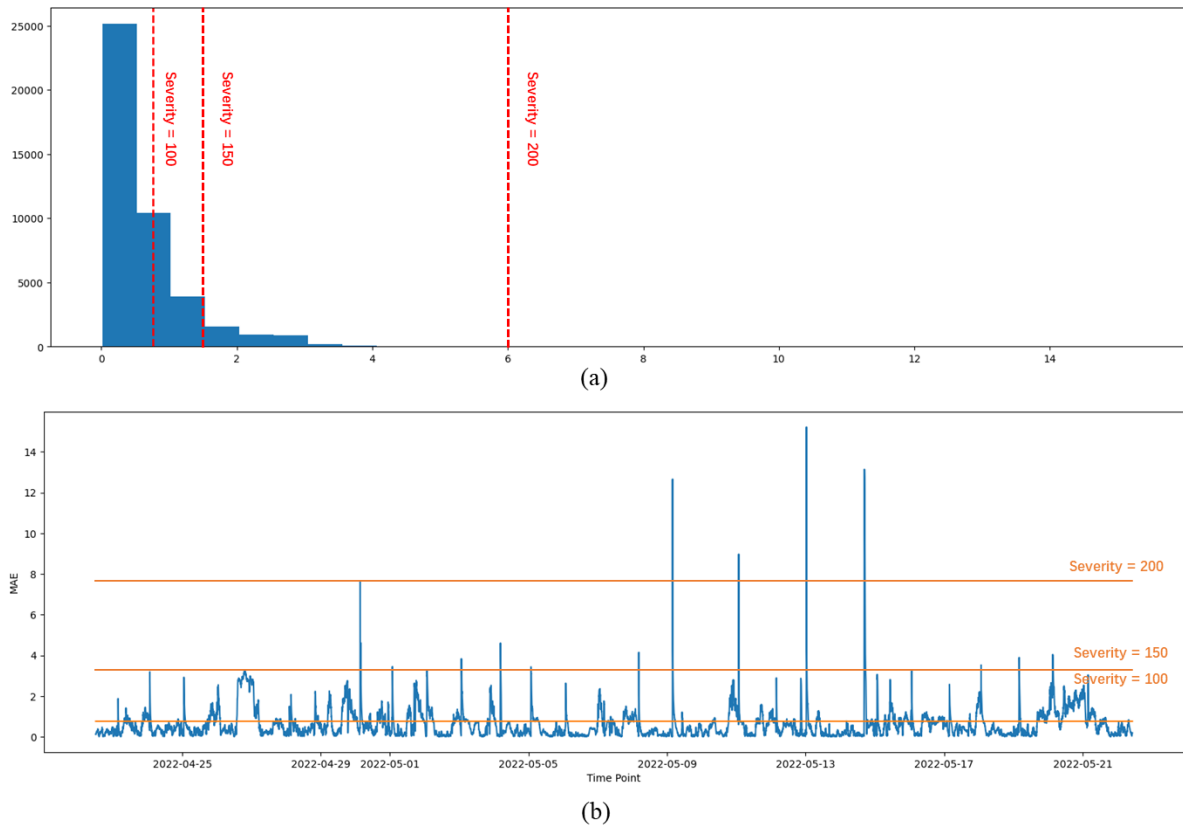


Figure 37. Distribution of reconstruction loss (a), Benchmark line in overall index (b)

Through the observation of outcomes, it is evident that a discernible relationship exists between the magnitude of the reconstruction error and the occurrence of failures in the 'overall index', particularly pronounced in severe cases (200), nearly all failure events in this category correspond to notably elevated reconstruction errors. Conversely, an increased number of false alarms is observed in situations characterised by relatively minor severities (100). This can be attributed to instances where the overall index experiences fluctuations, leading to varying reconstruction errors, and it is also affected by the baseline air quality of the practical environment where we obtained the dataset. Typically, the overall index in the experiment dataset tends to stabilise at approximately 70, while the reconstruction error is primarily induced by the scale of alteration in the overall index, rather than its absolute magnitude. Therefore, the AE component can trigger fake alerts due to sudden improvements in air quality in low-severity cases, while in the cases of high severity, the overall index is often accompanied by a single-direction drastic change that makes the air quality worse, which renders misjudgements less likely. Hence, two distinct methodologies are employed to optimise the ultimate predictions in the integration component for diverse scenarios, aiming to further improve the prediction performance and generalizability of the model.

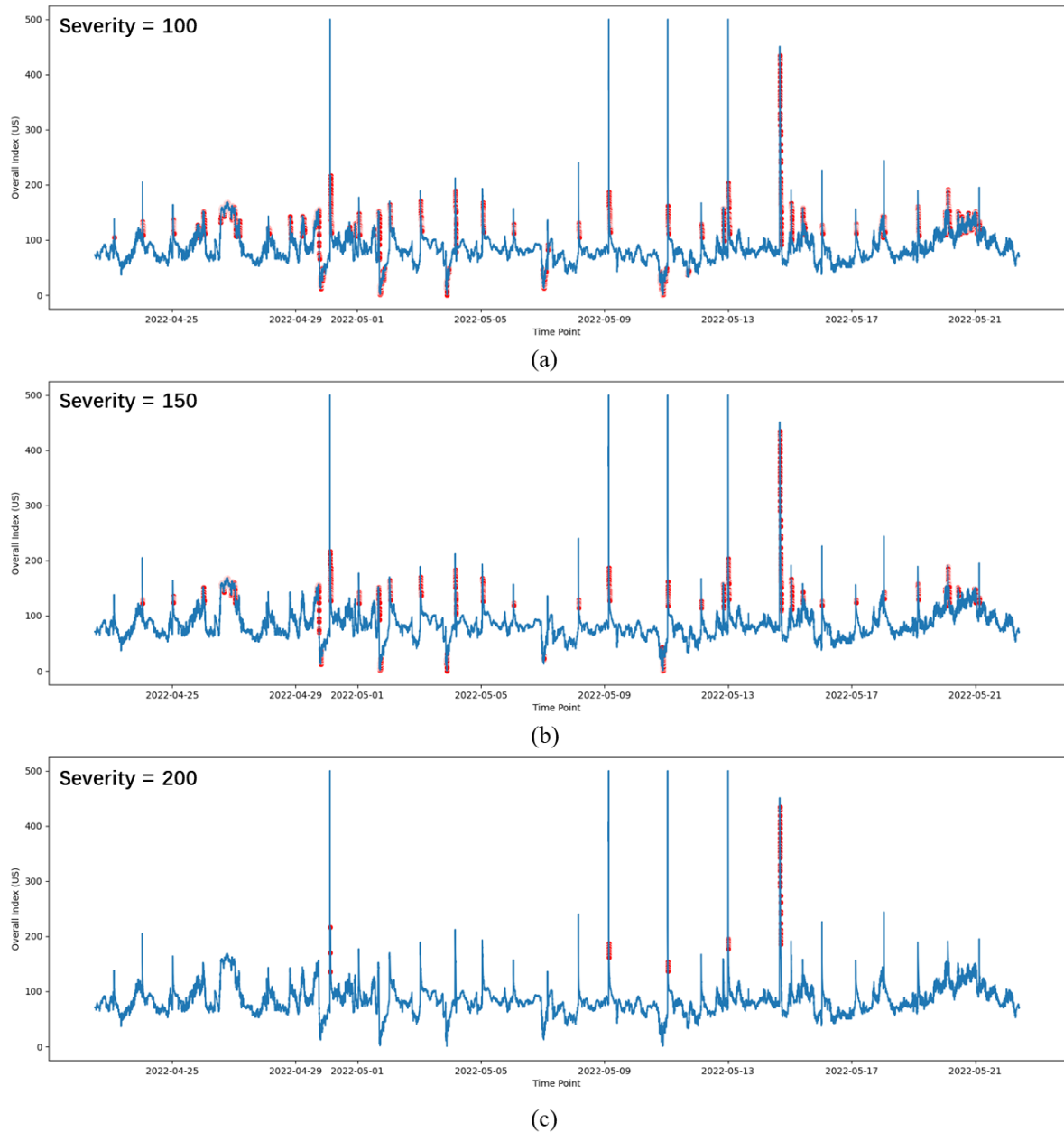


Figure 38. Failure distribution on the test set by AE component, (a) severity = 100, (b) severity =150, (c) severity = 200

6.7.2 Effect of Warning Time and Severity

In this section, a systematic investigation into the impact of failure severity and warning time configuration has been conducted to comprehensively assess the performance of the proposed methodology. The failure events have been categorised into three classes based on the obtained 'overall index' and USAQI standard [372]. Specifically, failure events with an overall index exceeding 100 were classified as low-level severity, those surpassing 150 were categorised as medium-level severity, and those exceeding 200

were designated as high-level severity. Additionally, the influence of warning time aligned with the RUL and TS methods has been systematically investigated. The range of warning time configurations varied from 10 minutes to 2 hours. Lastly, it should be emphasised that the model is trained ten times for each scenario, which contributes to a more robust evaluation and accounts for potential influence from diverse initial conditions and other pertinent factors.

The experiment results are shown in Figure 39, revealing discernible trends and patterns in the performance of the proposed methodology. For low severity, the algorithm initially exhibits an ascending performance trajectory with an ensuing decline as warning time increases. Remarkably, at a warning time of 30 minutes, the algorithm attains its best performance (Accuracy = 0.87, Precision = 0.96, Recall = 0.77, F1 score = 0.86). Under medium and high levels, the performance experiences degradation as the warning time increases. Notably, the accuracy is decreased to approximately 0.5 for high-severity cases under 2 hours warning time, signifying a performance akin to random guessing, and the algorithm is almost ineffective. Furthermore, an observation of the varying performance within evaluation metrics across different severity levels is found. Specifically, for low-severity cases, precision consistently has the highest value, followed sequentially by accuracy, F1-score, and recall. In contrast, for medium- and high-severity scenarios, recall stands as the best performance indicator, and it is succeeded by the F1 score, accuracy, and precision. This discernible pattern suggests distinct behaviours of the model for differentiating severity levels. The model demonstrates accuracy for low-severity events despite some cases evading its identification. Simultaneously, it is effective at capturing most actual medium and high-severity cases even though the prediction is not infallible.

The F1-score, being an amalgamation of precision and recall, is widely acknowledged as a well-balanced metric for assessing model performance. Our observations indicate that the model's performance is notably influenced by both the warning time and the severity degree. The model's configuration should be tailored to specific situations and requirements to attain optimal and satisfactory outcomes during the implementation. Furthermore, the categorisation of severity is subject to influence not only from established industry standards but also from the distribution patterns within the original

dataset. The severity categories threshold should be adjusted to align with different air quality levels in various environments.

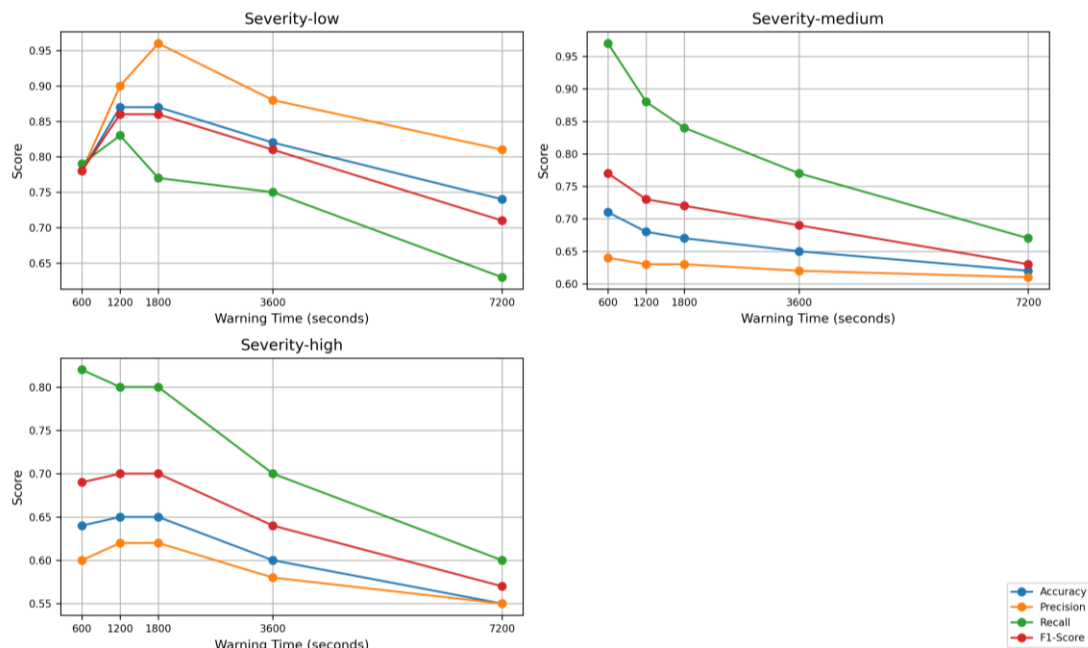


Figure 39. Model performance for different severity levels under different warning times

6.7.3 Comparison with Different Models

Prior PdM-related studies in the construction industry were mainly focused on anomaly detection of building facilities, with a limited emphasis on the prediction of failure events for IAQ. Notably, the present investigation has not found comparable studies in the existing literature, particularly for the failure prediction of indoor climate maintenance. Consequently, 4 classic ML models, including RF, SVM, Neural Network (NN), and LSTM, are employed as the benchmark models to assess the performance of the proposed model. RF and SVM are typical ML techniques proficiently utilised for classification and prediction tasks, with notable effectiveness in high-dimensional datasets. NN and LSTM are DL models consisting of interconnected layers of artificial neurons, learning complex patterns from datasets.

A comparative analysis of benchmark models and the proposed across three distinct severity levels has been conducted in this study. The evaluation metrics encompass accuracy, precision, recall, and F1 Score, which are utilised to analyse predictive

performance. We adopt the performance of the warning time with 30 minutes to represent the general performance of the proposed method. The predictive results of the proposed method for the test set under low severity with 30-minute warning time are illustrated in Figure 40, while Figure 41 graphically depicts the outcomes, revealing that the proposed method consistently outperforms other comparison models across low, medium, and high severity levels. Notably, as the severity level increases, a reduction in the performance gap between the proposed methodology and other approaches is observed in terms of accuracy and precision. At the highest severity, the disparity to LSTM in accuracy shrinks to 0.06, while the accuracy of the NN approach is even higher than the proposed model. Nonetheless, both the proposed and LSTM models excel in overall performance, primarily because of their remarkable performance in the recall domain. F1 provides a balanced perspective of the performance, considering both precision and recall, serving as a comprehensive comparative standard. Moreover, the LSTM exhibits comparable recall performance to the proposed method, except in cases of high severity, where the LSTM significantly lags behind the proposed model.

Overall, the performance of traditional ML methods (RF, SVM) exhibits suboptimal performance, primarily attributable to their limited recall rates, resulting in many overlooked alerts, whereas DL methods (NN, LSTM) show notably superior performance, with multiplicative improvements in F1 scores. The proposed algorithm consistently surpasses others by more than 10% across all severity scenarios. This success can be attributed to the proposed algorithm's enhanced utilisation of the dataset containing predictor and target data through the incorporation of RUL and TS techniques. Besides, the AE component captures meaningful representations from prior target inputs and detects the anomaly in future warning periods through an unsupervised approach, contributing to the final predictions through the GB or LD approach from the integration component. During the experimentation, a notable enhancement was observed in predicting high-severity failure events through the AE model, particularly regarding recall performance. This improvement can be attributed to AE's ability to effectively detect irregular and abrupt degradation in the overall index, which might be ignored by the LSTM model.

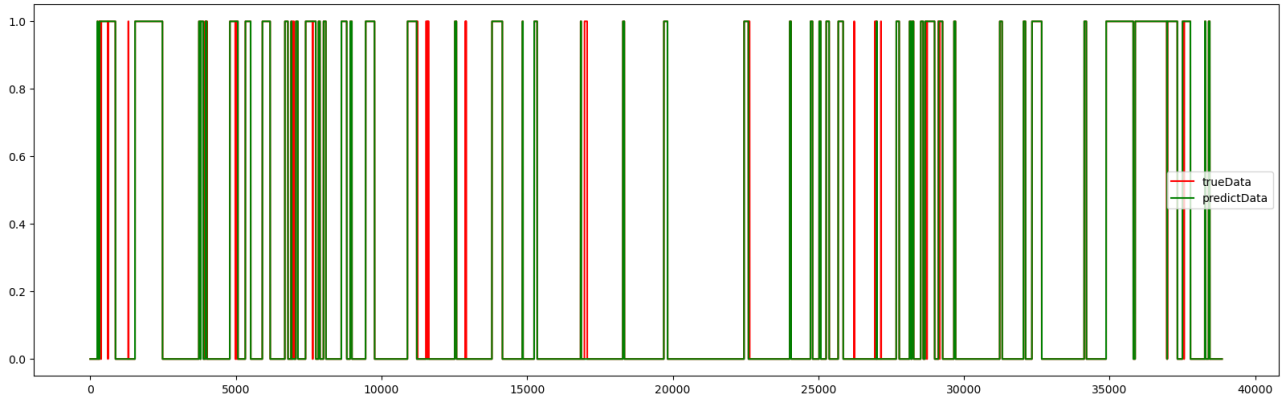


Figure 40. Failure prediction result of the proposed method (warning time = 30 m, severity = low)

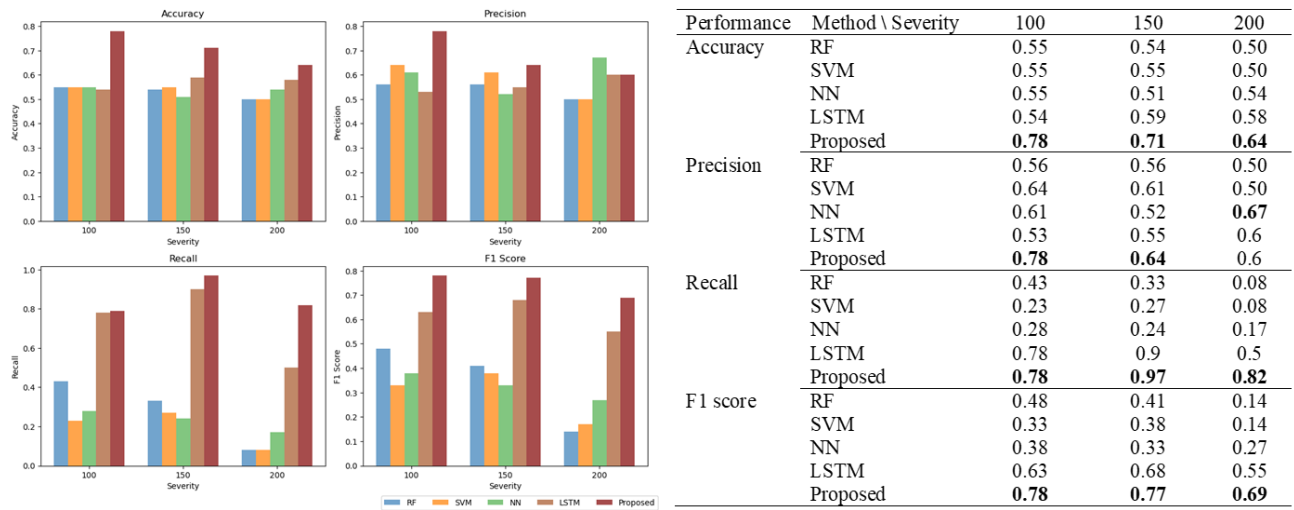


Figure 41. Comparative study of different benchmark models (warning time = 30 m)

6.7.4 Effect of Dataset

In this section, two additional datasets are employed further to investigate the applicability and universality of the model on unfamiliar datasets. We configure the experimental scenarios with a 30-minute warning time for low-severity cases. Moreover, the proposed method utilised the pre-trained model based on the Lakeside dataset in all scenarios, while the comparative models are individually specialised trained based on different target datasets.

Figure 42 demonstrates the model performance and dataset locations, revealing that, although the precision of the comparative model surpasses that of the proposed method in locations (Pioneer and Woodland) without specialised training, the overall performance of the proposed model remains superior to other comparative models

across all datasets. Furthermore, the disparity between the comparison and proposed models diminishes from Pioneer to Woodland. This trend can be attributed to the model's sensitivity to data distribution. Compared to Woodland, Pioneer exhibits greater similarity to Lakeside and shares more analogous environmental characteristics. Woodland, located on the northern border with distinct surroundings, exhibits a more disparate data structure and distribution. As a result, the Pioneer dataset aligns more closely with the training set from Lakeside, yielding superior performance of the proposed method compared to the other datasets.

In summary, the proposed model can provide competitive and satisfactory performance when directly applied to new datasets and scenarios, although specialised training can enhance its performance. For practical implementation, we recommend providing services and functions based on existing models during the initial phase, with data collection conducted concurrently within the proposed DT framework. In the advanced stage, the model should continuously improve by incorporating current data to achieve the best performance.

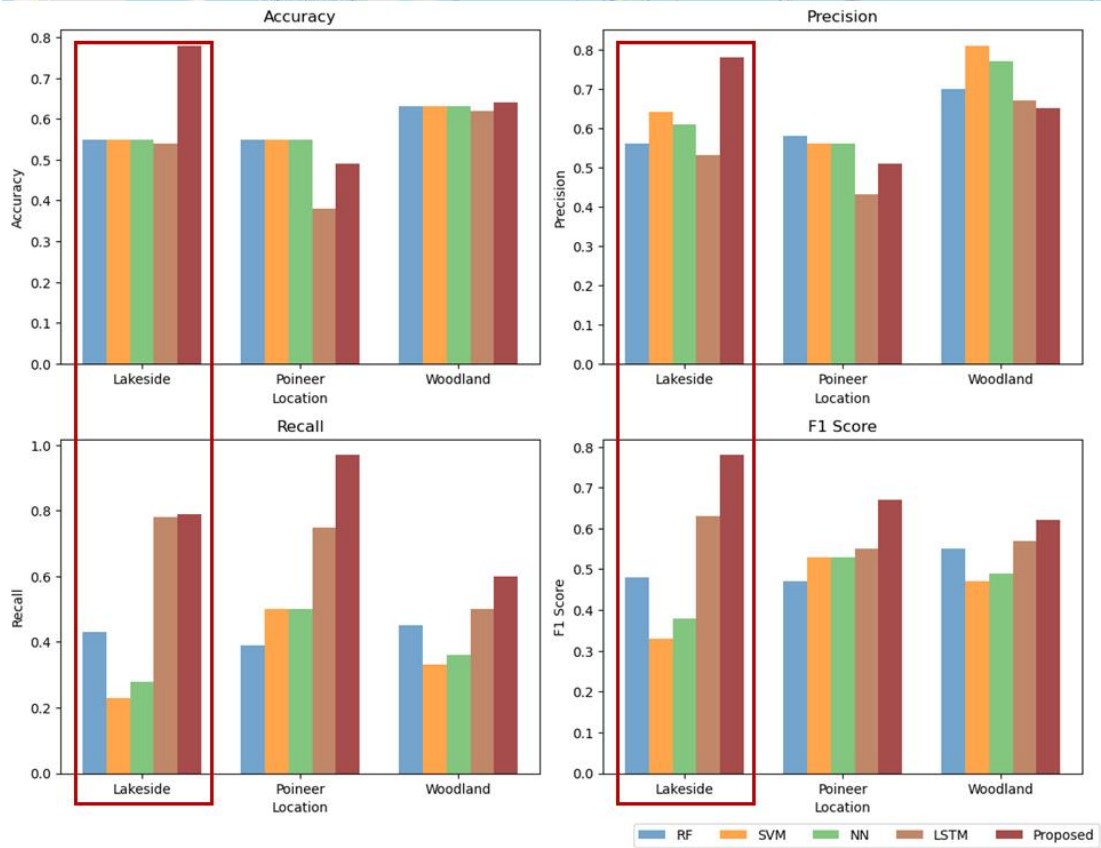
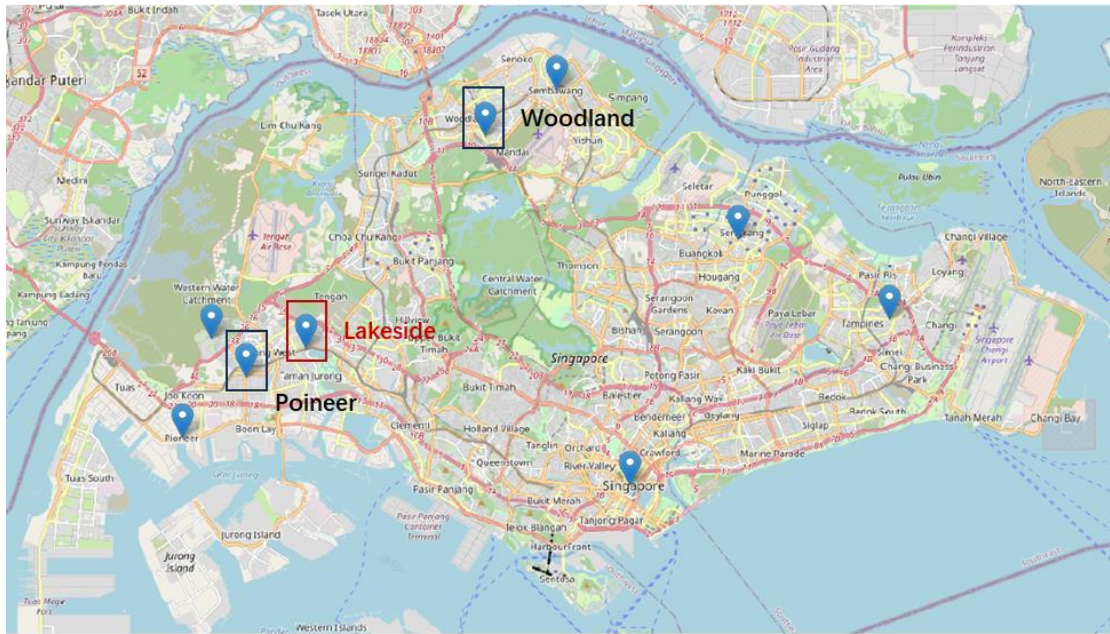


Figure 42. Comparative study of datasets with different locations (warning time = 30 m, severity = low)

6.7.5 Online Deployment

This section presents the online platform developed for DT concretisation and implementation, which consists of two sub-systems. The first is dedicated to visualising the building environment, and the second focuses on the presentation of real-time monitoring information and alarm notifications. The mainstream approach for building visualisation is executed through the implementation of BIM. Nevertheless, it is noteworthy that not all target buildings have BIM models, particularly those built before the occurrence of the BIM. As shown in Figure 43, the IScan2BIM system has been developed to overcome this limitation by generating as-built BIM models for existing buildings, offering building information such as 3D measurements and layout specifics. The system starts with the scanning processing of the target building utilising automated robotics, followed by the reconstruction of building structures and MEP components. The algorithm development is illustrated in BEAcon [370] and PointNet [325], while the practical demonstration can be accessed at https://www.youtube.com/watch?v=RE-vNYs_i_w, showcasing the BIMbot-driven intelligent scanning process, AI-enabled reconstruction procedures, and an Amazon Web Services cloud platform for the B&C industry. This system functions as the main platform for integrating indoor climate information for PdM applications in this study.

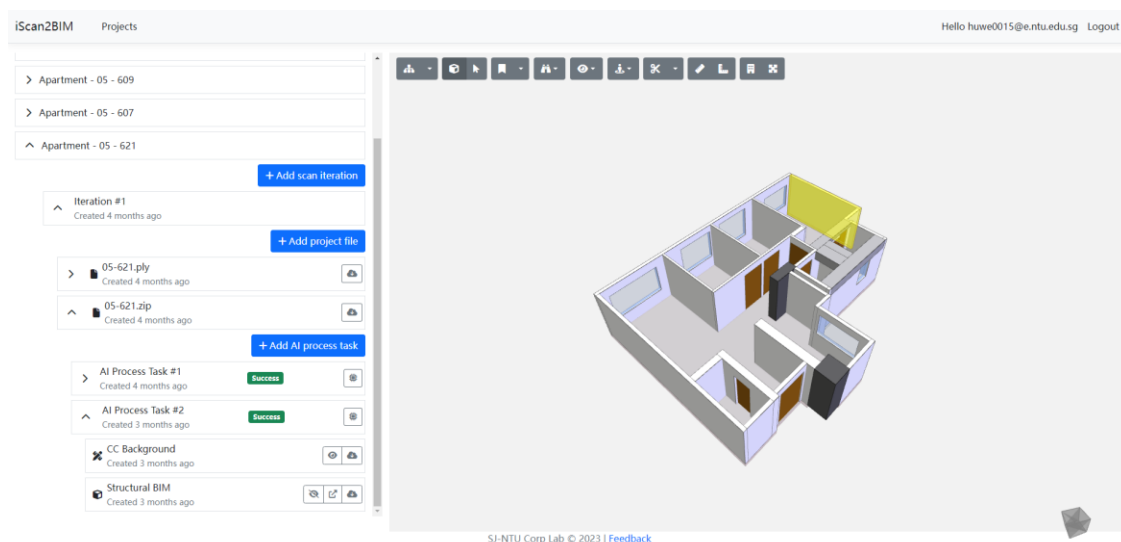


Figure 43. BIM reconstruction from IScan2BIM

To provide real-time monitoring and alerts for indoor climate conditions, a data-centric web application has been developed by the Streamlit library in Python, as depicted

in • ure 44. Streamlit was chosen as the development framework because it is highly accessible for individuals with limited programming or web development knowledge. A standout attribute of Streamlit is its hot-loading implementation, facilitating dynamic updates without requiring manual page refreshes. This feature guarantees that users consistently access the most up-to-date information without disruptions. Furthermore,

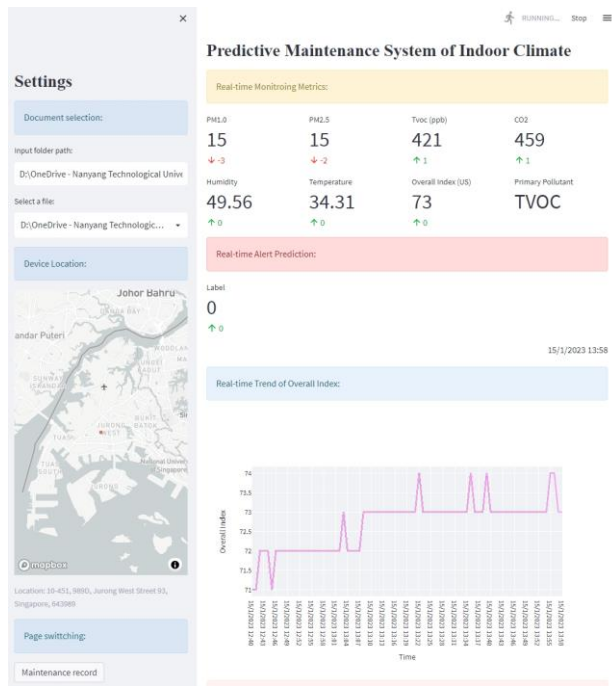


Figure 45. The online platform of PdM for indoor climate

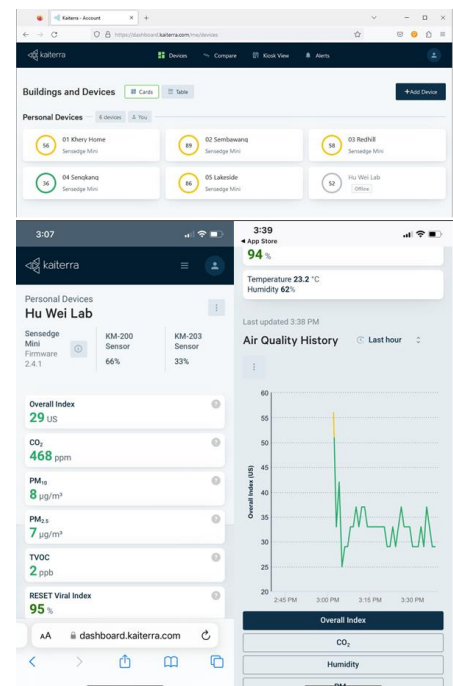


Figure 44. IoT platform (Web & APP)

Streamlit applications can be effortlessly shared by distributing a URL link, and they offer versatile deployment options across a spectrum of platforms and cloud services. This flexibility simplifies the deployment and demonstration processes, making it a practical choice for showcasing the application's functionality. The website presents various indoor climate metrics, including PM1.0, PM2.5, TVOC, CO₂, humidity, temperature, and overall index. It also displays binary warning signals, where 0 indicates normal operating conditions, and 1 signals a potential issue within the subsequent warning period (30 minutes in this experiment). Additionally, the system integrates a maintenance record page, enabling operators to document failure occurrences and maintenance information. Currently, the system derives input data from CSV files and is deployed to a local server for the proof of concept, while future work is planned to leverage APIs for real-time data collection and enable online deployment. Finally, the commercial IoT devices utilised for data acquisition are shown in Figure 45.

6.7.6 Discussion on DT-enhanced PdM Implementation

The implementation of DT technology significantly enhances PdM strategies by introducing improved interactive capabilities and optimised planning and scheduling, as demonstrated by the proposed framework and experimental results. DT enables real-time synchronisation between physical systems and their digital counterparts, offering a dynamic and interactive platform for monitoring and diagnosing indoor air quality. This interactive capability allows facility managers to visualise the current state of equipment, receive early warnings for potential faults, and simulate maintenance actions before they are performed on the actual systems.

In this study, the predictive capabilities of the DT framework are powered by the proposed parallel LSTM-AE method, enhancing maintenance planning by providing accurate and detailed failure predictions. Experiment 1 explored varying warning times and severity levels, demonstrated that the system can flexibly adjust its prediction period from short-term (10 minutes) to long-term (12 hours), and effectively alert users based on the severity of the failure events. This detailed decision-support information significantly improves the selection and deployment of maintenance strategies, reducing equipment downtime and avoiding emergency repairs. Besides, the optimisation of scheduling can also be improved through accurate and actionable predictions. Experiment 2 compared the proposed method to other common benchmark models, confirming the superior performance of the proposed method in accuracy. This enables better planning by forecasting failure more reliably and allowing maintenance actions to be scheduled more precisely, minimising disruptions. Additionally, experiment 3 verified the proposed model on new datasets and highlighted the system's versatility and applicability across different scenarios. The ability to perform well on unfamiliar datasets demonstrates the framework's generalizability, ensuring that PdM strategies can be scaled across diverse environments and equipment types without sacrificing performance.

Overall, DT implementation not only improves real-time monitoring and user interaction but also enhances PdM strategy by enabling more efficient and optimised scheduling. This ensures maintenance activities are carried out based on accurate predictions, reducing costs and improving operational efficiency.

6.8 Remark

PdM in the building industry has emerged as a prominent application in Construction 4.0 [7], encompassing not only the maintenance of building facilities but also extending to indoor climate. This study explores the potential of combining the DT concept with DL models to address PdM challenges in the realm of indoor climate, which showcases the effectiveness of the proposed parallel model for the failure prediction of indoor climate and illustrates how the DT paradigm can expedite the implementation of PdM strategies within the built environment through the proposed framework and practical online application.

The contributions of this study are summarised as follows. First, it integrates the RUL and TS into the data processing pipeline to convert the failure prediction task into a combination of a supervised and an unsupervised learning task. Secondly, a hybrid DL model, fusing LSTM and AE, is developed to predict multi-severity failures, which is implemented in real-world IAQ datasets sourced from diverse buildings in Singapore. Moreover, a meticulous comparison between the proposed model and the other 4 benchmark models substantiates the superior predictive capabilities of the former, signifying the optimised utilisation of the predictors and previous targets in the dataset. For the most common failure for indoor climate, the low severity level, the performance matrix of the proposed model is close to 80%, which is 10 to 15% higher than other comparative algorithms overall and holds a certain practical value in real-world implementations. Lastly, the study provides a depiction of the proposed DT-enabled PdM framework for indoor climate, verified by the online platform that showcases building models, air quality data, and alert information. This practical demonstration can provide a guiding example for future endeavours in the realm of DT-enhanced PdM solutions.

This study can be seen as the initiation stage of DT-enabled failure prediction for the indoor climate within the PdM domain, and there are several limitations and future works worth exploring. The proposed overall index should ideally include additional parameters like SO_2 , NO_2 , CO, PM10, and O_3 for a more holistic assessment. However,

the limitations of our data collection device have constrained us from including the abovementioned parameters in the current dataset, underscoring the need to strengthen data acquisition capabilities for future research. Additionally, variables beyond air quality metrics, including device location, building structure, green space, etc, can substantially affect the model's performance. Therefore, future research should deploy devices in diverse locations within buildings and consider more relevant parameters for enhanced accuracy and reliability of the prediction model. Furthermore, the efficacy of high-severity failure event forecasting remains limited in this study, primarily due to a scarcity of severe air pollution events in the collected dataset, which calls for augmenting data diversity through extended collection efforts or the incorporation of synthetic data generation methods, thereby potentially enhancing the versatility and resilience of predictive models. Finally, DT-enhanced PdM holds great potential; more functions, including asset scheduling, cost reduction, and decision-making, are valuable to develop and integrate into the platform, enabling the building industry to safety, efficiency, sustainability, and predictability, ultimately transforming maintenance practices, and optimising FM.

In conclusion, this research provides comprehensive insights into failure prediction, data processing techniques, DL models, and the implementation of DT in PdM for indoor climate. The contributions and insights gained provide a foundation for improved PdM strategies, while limitations acknowledged offer a guideline for future research and development.

Chapter 7 Conclusion and Future Work

In this chapter, we summarise the research that was conducted. This chapter presents a summary of the research and details all key contributions, followed by discussions of potential future research directions.

7.1 Summary

This thesis combines four studies that collectively enhance our understanding of PdM and DT in the building industry. These chapters analyse various aspects of PdM and DT, showcasing their versatility and applicability in improving the building field.

Firstly, the versatility and scalability of DT solutions across industries, including aerospace and healthcare, are discussed. Then, recent related studies within the B&C industry are systematically reviewed, deriving a unified framework that integrates Industry 4.0 technologies as functional modules within the DT structure. It also introduces the innovative Six M methodology for analysing the role of DT-enabled systems within construction functional roles.

Secondly, existing research on PdM in the B&C industry is categorised into four clusters based on the keyword network analysis approach to identify the key enabling factors and challenges during the implementation of PdM. Besides, the functionality and applicability of PdM deployments are highlighted through an in-depth analysis utilising the Three-by-Three matrix M approach, adapted from Ishikawa's diagram. This study provides a foundation for future research and practice.

Thirdly, we emphasise the adaptability and scalability of DT solutions, specifically focusing on their application in PdM for building facilities. A novel failure prediction approach based on Semi-supervised GAN is proposed to improve prediction accuracy and reduce dependency on labelled data. Also, a three-layer DT-enabled PdM framework is developed to underscore the potential for future advancements in PdM solutions through DT paradigms.

Lastly, we explore the combination of DT concepts with DL models for PdM in indoor climate maintenance. We introduce a hybrid DL model and demonstrate its superior predictive capabilities in real-world IAQ datasets. A DT-enabled PdM framework for indoor climate is also presented, accompanied by a practical online platform to show the monitoring and alarm information.

In conclusion, these research sections collectively advance the understanding and application of DT and PdM in the B&C industry. They provide valuable insights, methodologies, and practical applications that contribute to the ongoing transformation of construction practices, building maintenance, and FM. Additionally, future research directions are identified in each study to offer a roadmap for further exploration in this dynamic field.

7.2 Contributions

This thesis provides significant scientific and theoretical advancements in the fields of DT, PdM, and Industry 4.0 technologies, particularly within the building industry. The detailed contributions are outlined as follows:

1) Introduction of a Novel Architecture for DT Integration in Construction: This research introduces an innovative architecture that integrates Industry 4.0 technologies into DT systems for construction, underpinned by the proposed Six M methodology. This architecture serves as a blueprint for enhancing construction operations by leveraging advanced technologies such as multi-functional sensors and AI. Additionally, the research identifies future research directions that focus on enhancing sustainable construction practices through the deployment of AI-driven solutions and multi-functional sensor systems, laying the groundwork for future scientific exploration in these areas.

2) Categorization of PdM Implementations and Analytical Framework: The study categorises PdM implementations into four distinct clusters using a novel keywords network analysis approach. This categorisation provides a structured understanding of PdM deployments across various industries. Moreover, the research introduces the

Three-by-Three Matrix M approach, a comprehensive analytical framework designed to evaluate the functionality and applicability of PdM systems. This methodological contribution offers a new lens for the theoretical exploration of PdM practices, contributing to the ongoing evolution of maintenance strategies in industrial and construction contexts.

3) Emphasis on DT and PdM Integration for Enhanced System Performance: A key scientific contribution of this research is the emphasis on the integration of PdM with DT technologies to improve the performance and accuracy of maintenance solutions. The research proposes a novel three-layer DT-enabled PdM architecture specifically designed for building facilities management. This layered architecture enhances the theoretical understanding of how DT systems can optimise PdM processes and provides a robust framework for the further development of smart building systems.

4) Development of a Semi-supervised GAN-based Failure Prediction Methodology: The study presents a semi-supervised GAN-based approach for failure prediction. This method significantly advances the field by reducing the reliance on labelled datasets, which are often scarce in real-world scenarios. This contribution enhances the theoretical body of knowledge on ML techniques applied to predictive maintenance, particularly in scenarios where labelled data is limited, thereby improving prediction accuracy and system reliability.

5) Development of a Parallel Failure Prediction Methodology: Another critical contribution of this research is the integration of RUL and TS techniques into the PdM data processing pipeline. The development of a hybrid deep learning model for multi-severity failure prediction further contributes to the theoretical foundations of PdM. This hybrid model offers new insights into how multiple failure severities can be predicted more accurately, enhancing the longevity and resilience of industrial systems.

6) Proposal of a DT-enabled PdM Framework for Indoor Climate Management: The research proposes a DT-enabled PdM framework specifically aimed at managing indoor climates. This framework is complemented by the development of an online platform that showcases a comprehensive building model, along with real-time monitoring data and predictive alarm systems. This contribution is both scientific and

practical, providing a novel theoretical framework for the application of DT and PdM in the domain of climate management while also offering a tangible solution for real-world implementation.

7.3 Limitation Discussion

In this section, the limitations of the proposed frameworks and models are discussed in terms of their real-world applicability and generalizability.

The first limitation of the proposed frameworks is their dependency on high-quality and extensive datasets for effective performance. During PdM implementation, the collection of reliable labelled data can be challenging, especially in environments where failure events are rare or difficult to observe. Models trained on imbalanced or noisy datasets may suffer from reduced accuracy and generalisation ability, which limits their effectiveness in real-world settings [366]. Without access to comprehensive datasets, the models may struggle to capture all the relevant features of the system, potentially leading to missed failures or false alarms.

Secondly, while the proposed frameworks perform well in the experimental setup, their scalability to larger and more complex systems poses a challenge. Real-world industrial systems often involve intricate interactions between various subsystems, making it difficult for predictive models to accurately capture all the dependencies. As noted by Yang et al.[133], models may fail to generalise well when applied to systems with more complex operational dynamics than those seen during training. This limitation suggests that additional research is needed to extend the scalability of these models to larger, interconnected environments.

Another limitation relates to the computational complexity of the models, particularly when employing deep learning techniques such as Semi-GANs and AE-LSTMs. Significant computational resources are required for the deployment of these models, especially during the training phase. This can pose challenges in real-time PdM applications, where fast processing and decision-making are critical. Computational

inefficiencies can hinder the deployment of complex models in resource-constrained environments [376].

Finally, although the experiment design demonstrated the model's applicability to new datasets, the framework may still face limitations when applied to completely unseen or evolving operational conditions. As systems change over time due to wear and tear, environmental factors, or upgrades, the model may need continuous retraining to remain effective. As suggested by Liu et al. [377], models without adaptive learning mechanisms may fail to generalise well to changing operational environments, leading to suboptimal maintenance decisions.

In summary, although the proposed frameworks show promise for enhancing PdM strategies, limitations related to data quality, scalability, computational complexity, and adaptability still need to be addressed for widespread practical deployment. Future research should focus on overcoming these challenges to ensure the models are robust, scalable, and applicable to a wide range of industrial contexts.

7.4 Future Direction

1) Enhance the algorithm's performance. New models like Transformers and Informer hold promise for future improvement. Additionally, advanced SMOTE methods warrant future research attention because failure prediction scenarios often suffer from serious data imbalance issues, where erroneous data events are typically much fewer than normal data samples.

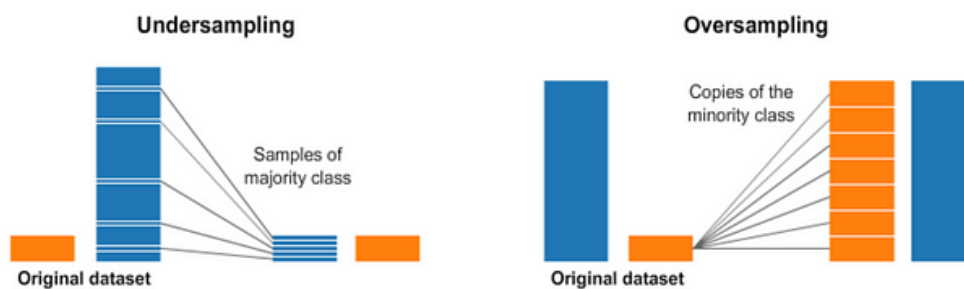


Figure 46. Data imbalance problem

2) Improve the adaptability and functionality of the model. The future method should include the prediction of multiclass faults of the target asset. Besides, concurrent faults in the system should also be explored in the future.



Figure 47. Prediction of multiclass faults & concurrent faults

3) Integrate with the visualisation platform. Future methods should integrate the IoT system and virtual model like BIM to achieve real-time 3D visualisation [378], analysis, and alert information of the target facilities, as well as further improve the performance of PdM.

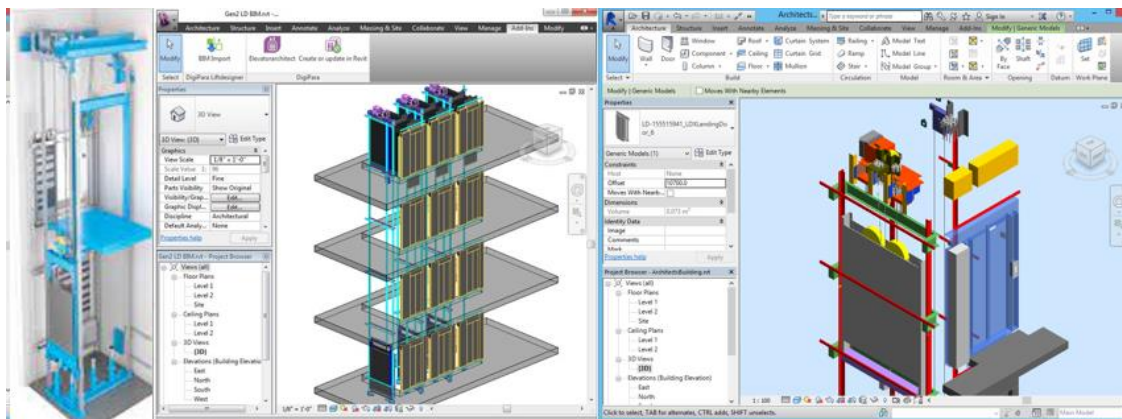


Figure 48. BIM model for building facilities (lift)

4) Develop multi-functional and integrated DT systems. Incorporate functionalities with higher operation performance to include additional project considerations within the same platform [379], such as environment monitoring, safety management, and building energy management.

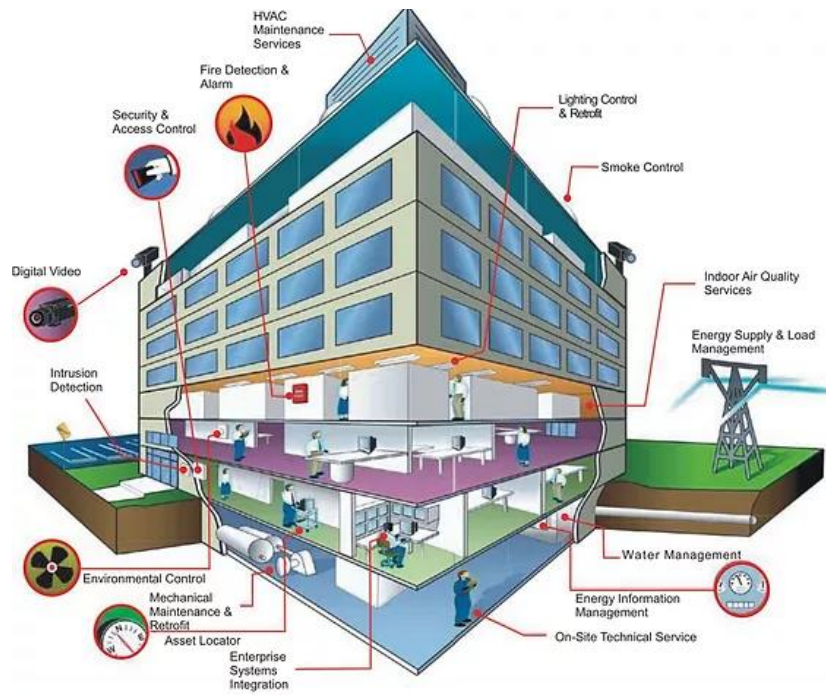


Figure 49. Multi-functional DT platform

Bibliography

- [1] P. Zheng and K. Y. H. Lim, "Product family design and optimisation: A digital twin-enhanced approach," *Procedia CIRP*, vol. 93, pp. 246–250, 2020, doi: 10.1016/j.procir.2020.05.162.
- [2] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital Twin in Industry: State-of-the-Art," *IEEE Trans. Ind. Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019, doi: 10.1109/TII.2018.2873186.
- [3] Y. Zheng, S. Yang, and H. Cheng, "An application framework of digital twin and its case study," *J. Ambient Intell. Humaniz. Comput.*, vol. 10, no. 3, pp. 1141–1153, 2019, doi: 10.1007/s12652-018-0911-3.
- [4] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital Twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018, doi: 10.1016/j.ifacol.2018.08.474.
- [5] Z. Shi and W. O'Brien, "Development and implementation of automated fault detection and diagnostics for building systems: A review," *Autom. Constr.*, vol. 104, no. May, pp. 215–229, 2019, doi: 10.1016/j.autcon.2019.04.002.
- [6] M. Jasiulewicz-Kaczmarek, S. Legutko, and P. Kluk, "Maintenance 4.0 technologies - new opportunities for sustainability driven maintenance," *Manag. Prod. Eng. Rev.*, vol. 11, no. 2, pp. 74–87, 2020, doi: 10.24425/mper.2020.133730.
- [7] A. Sawhney, M. Riley, and J. Irizarry, *Construction 4.0: An Innovation Platform for the Built Environment*. 2020. doi: 10.1201/9780429398100.
- [8] McKinsey & Company and M. G. Institute, "Reinventing Construction: A Route To Higher Productivity," *McKinsey Co.*, no. February, p. 12, 2017, [Online]. Available: <http://www.mckinsey.com/industries/capital-projects-and-infrastructure/our-insights/reinventing-construction-through-a-productivity-revolution%0Awww.revalue.dk>
- [9] Z. You and L. Feng, "Integration of Industry 4.0 Related Technologies in Construction Industry: A Framework of Cyber-Physical System," *IEEE Access*, vol. 8, pp. 122908–122922, 2020, doi: 10.1109/ACCESS.2020.3007206.
- [10] M. Themistocleus, P. R. Da Cunha, and H. Treiblmaier, "Metaverse," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2023-Janua, p. 6034, 2023, doi: 10.4018/ijea.316537.
- [11] British Standards Institution (BSI), "BSI Standards Publication Maintenance — Maintenance terminology," p. 36, 2010.
- [12] D. Mourtzis, V. Zogopoulos, and E. Vlachou, "Augmented reality application to support remote maintenance as a service in the Robotics industry," *Procedia CIRP*, vol. 63, pp. 46–51, 2017, doi: 10.1016/j.procir.2017.03.154.
- [13] M. James *et al.*, "Unlocking the potential of the Internet of Things | McKinsey & Company," *McKinsey*, pp. 1–4, 2015, [Online]. Available: <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world>
- [14] S. J. Emmerich, K. Y. Teichman, and A. K. Persily, "Literature review on field study of ventilation and indoor air quality performance verification in high-performance commercial buildings in North America," *Sci. Technol. Built Environ.*, vol. 23, no. 7, pp. 1159–1166, 2017, doi: 10.1080/23744731.2016.1274627.
- [15] J. Y. Kim, C. H. Chu, and S. M. Shin, "ISSAQ: An integrated sensing systems for real-time indoor air quality monitoring," *IEEE Sens. J.*, vol. 14, no. 12, pp. 4230–4244, 2014, doi: 10.1109/JSEN.2014.2359832.
- [16] N. R. Kapoor, A. Kumar, T. Alam, A. Kumar, K. S. Kulkarni, and P. Blechich, "A review on indoor environment quality of indian school classrooms," *Sustain.*, vol. 13, no. 21, 2021, doi: 10.3390/su132111855.
- [17] Y. Liu, Z. Pang, M. Karlsson, and S. Gong, "Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control," *Build. Environ.*, vol. 183, no. July, p. 107212, 2020, doi: 10.1016/j.buildenv.2020.107212.
- [18] L. Zhao *et al.*, "Digital Twin Evaluation of Environment and Health of Public Toilet Ventilation Design Based on Building Information Modeling," *Buildings*, vol. 12, no. 4, 2022, doi: 10.3390/buildings12040470.
- [19] S. L. Ullo and G. R. Sinha, "Advances in smart environment monitoring systems using iot and sensors," *Sensors (Switzerland)*, vol. 20, no. 11, pp. 1–18, 2020, doi: 10.3390/s20113113.
- [20] S. Shirowzhan, W. Tan, and S. M. E. Sepasgozar, "Digital twin and CyberGIS for improving connectivity and measuring the impact of infrastructure construction planning in smart cities,"

- ISPRS Int. J. Geo-Information*, vol. 9, no. 4, 2020, doi: 10.3390/ijgi9040240.
- [21] Z. A. M. Bhuiyan, J. Wu, G. Wang, J. Cao, W. Jiang, and M. Atiquzzaman, “Towards cyber-physical systems design for structural health monitoring: Hurdles and opportunities,” *ACM Trans. Cyber-Physical Syst.*, vol. 1, no. 4, 2017, doi: 10.1145/3086508.
- [22] G. Loubet, A. Takacs, E. Gardner, A. De Luca, F. Udrea, and D. Dragomirescu, “LoRaWAN battery-free wireless sensors network designed for structural health monitoring in the construction domain,” *Sensors (Switzerland)*, vol. 19, no. 7, 2019, doi: 10.3390/s19071510.
- [23] L. Yang, C. Fu, Y. Li, and L. Su, “Survey and study on intelligent monitoring and health management for large civil structure,” *Int. J. Intell. Robot. Appl.*, vol. 3, no. 3, pp. 239–254, 2019, doi: 10.1007/s41315-019-00079-2.
- [24] Y. Y. Zhang, K. Kang, J. R. Lin, J. P. Zhang, and Y. Zhang, “Building information modeling-based cyber-physical platform for building performance monitoring,” *Int. J. Distrib. Sens. Networks*, vol. 16, no. 2, 2020, doi: 10.1177/1550147720908170.
- [25] Y.-C. Lin and W.-F. Cheung, “Developing WSN/BIM-Based Environmental Monitoring Management System for Parking Garages in Smart Cities,” *J. Manag. Eng.*, vol. 36, no. 3, p. 04020012, 2020, doi: 10.1061/(asce)me.1943-5479.0000760.
- [26] S. H. Khajavi, N. H. Motlagh, A. Jaribion, L. C. Werner, and J. Holmstrom, “Digital Twin: Vision, benefits, boundaries, and creation for buildings,” *IEEE Access*, vol. 7, pp. 147406–147419, 2019, doi: 10.1109/ACCESS.2019.2946515.
- [27] J. Grübel *et al.*, “Dense Indoor Sensor Networks: Towards passively sensing human presence with LORAWAN,” *Pervasive Mob. Comput.*, vol. 84, 2022, doi: 10.1016/j.pmcj.2022.101640.
- [28] S. Abrol, A. Mehmani, M. Kerman, C. J. Meinrenken, and P. J. Culligan, “Data-Enabled Building Energy Savings (D-E BES),” *Proc. IEEE*, vol. 106, no. 4, pp. 661–679, 2018, doi: 10.1109/JPROC.2018.2791405.
- [29] V. G. González, G. R. Ruiz, and C. F. Bandera, “Empirical and comparative validation for a building energy model calibration methodology,” *Sensors (Switzerland)*, vol. 20, no. 17, pp. 1–29, 2020, doi: 10.3390/s20175003.
- [30] C. Keskin and M. Pinar Mengüç, “An adaptive vent system for localized and customized thermal management in buildings,” *J. Heat Transfer*, vol. 142, no. 5, 2020, doi: 10.1115/1.4045664.
- [31] G. Liu *et al.*, “Cyber-physical system-based real-time monitoring and visualization of greenhouse gas emissions of prefabricated construction,” *J. Clean. Prod.*, vol. 246, 2020, doi: 10.1016/j.jclepro.2019.119059.
- [32] G. Liu, R. Chen, P. Xu, Y. Fu, C. Mao, and J. Hong, “Real-time carbon emission monitoring in prefabricated construction,” *Autom. Constr.*, vol. 110, no. 174, 2020, doi: 10.1016/j.autcon.2019.102945.
- [33] L. Mainetti, L. Patrono, and A. Vilei, “Evolution of wireless sensor networks towards the Internet of Things: A survey,” *2011 Int. Conf. Software, Telecommun. Comput. Networks, SoftCOM 2011*, pp. 16–21, 2011.
- [34] L. C. Tagliabue, F. R. Cecconi, S. Maltese, S. Rinaldi, A. L. C. Ciribini, and A. Flammini, “Leveraging digital twin for sustainability assessment of an educational building,” *Sustain.*, vol. 13, no. 2, pp. 1–16, 2021, doi: 10.3390/su13020480.
- [35] Z. Liu, A. Zhang, and W. Wang, “A framework for an indoor safety management system based on digital twin,” *Sensors (Switzerland)*, vol. 20, no. 20, pp. 1–20, 2020, doi: 10.3390/s20205771.
- [36] Y. Niu, C. Anumba, and W. Lu, “Taxonomy and Deployment Framework for Emerging Pervasive Technologies in Construction Projects,” *J. Constr. Eng. Manag.*, vol. 145, no. 5, p. 04019028, 2019, doi: 10.1061/(asce)co.1943-7862.0001653.
- [37] F. Zonzini *et al.*, “Structural Health Monitoring and Prognostic of Industrial Plants and Civil Structures: A Sensor to Cloud Architecture,” *IEEE Instrum. Meas. Mag.*, vol. 23, no. 9, pp. 21–27, 2020, doi: 10.1109/MIM.2020.9289069.
- [38] Ž. Turk and R. Klinc, “A social-product-process framework for construction,” *Build. Res. Inf.*, vol. 48, no. 7, pp. 747–762, 2020, doi: 10.1080/09613218.2019.1691487.
- [39] C. Boje, A. Guerriero, S. Kubicki, and Y. Rezugui, “Towards a semantic Construction Digital Twin: Directions for future research,” *Autom. Constr.*, vol. 114, no. March, p. 103179, 2020, doi: 10.1016/j.autcon.2020.103179.
- [40] Y. Wei and B. Akinici, “Panorama-to-model registration through integration of image retrieval and semantic reprojection,” *Autom. Constr.*, vol. 140, no. May, p. 104356, 2022, doi: 10.1016/j.autcon.2022.104356.
- [41] H. Wu *et al.*, “Multi-level voxel representations for digital twin models of tunnel geological environment,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 112, no. May, p. 102887, 2022, doi: 10.1016/j.jag.2022.102887.

- [42] P. Tasca, T. Aste, L. Pelizzon, and N. Perony, *Banking Beyond Banks and Money*. 2016. doi: 10.1007/978-3-319-42448-4_1.
- [43] D. Lee, S. H. Lee, N. Masoud, M. S. Krishnan, and V. C. Li, “Integrated digital twin and blockchain framework to support accountable information sharing in construction projects,” *Autom. Constr.*, vol. 127, no. July 2020, p. 103688, 2021, doi: 10.1016/j.autcon.2021.103688.
- [44] J. J. Hunhevciz, M. Motie, and D. M. Hall, “Digital building twins and blockchain for performance-based (smart) contracts,” *Autom. Constr.*, vol. 133, p. 103981, 2022, doi: 10.1016/j.autcon.2021.103981.
- [45] Y. Jiang, X. Liu, K. Kang, Z. Wang, R. Y. Zhong, and G. Q. Huang, “Blockchain-enabled cyber-physical smart modular integrated construction,” *Comput. Ind.*, vol. 133, 2021, doi: 10.1016/j.compind.2021.103553.
- [46] C. Z. Li *et al.*, “A blockchain- and IoT-based smart product-service system for the sustainability of prefabricated housing construction,” *J. Clean. Prod.*, vol. 286, 2021, doi: 10.1016/j.jclepro.2020.125391.
- [47] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012. doi: 10.1016/C2009-0-61819-5.
- [48] M. Schmidt, M. V. Moreno, A. Schülke, K. Macek, K. Mařík, and A. G. Pastor, “Optimizing legacy building operation: The evolution into data-driven predictive cyber-physical systems,” *Energy Build.*, vol. 148, pp. 257–279, 2017, doi: 10.1016/j.enbuild.2017.05.002.
- [49] Y. Pan and L. Zhang, “A BIM-data mining integrated digital twin framework for advanced project management,” *Autom. Constr.*, vol. 124, no. July 2020, p. 103564, 2021, doi: 10.1016/j.autcon.2021.103564.
- [50] S. Dutta, Y. Cai, L. Huang, and J. Zheng, “Automatic re-planning of lifting paths for robotized tower cranes in dynamic BIM environments,” *Autom. Constr.*, vol. 110, no. December 2019, p. 102998, 2020, doi: 10.1016/j.autcon.2019.102998.
- [51] G. Desogus, E. Quaquero, G. Rubiu, G. Gatto, and C. Perra, “Bim and iot sensors integration: A framework for consumption and indoor conditions data monitoring of existing buildings,” *Sustain.*, vol. 13, no. 8, 2021, doi: 10.3390/su13084496.
- [52] E. Wemerová, S. Endel, and D. Kutá, “Implementation of the BIM method at the VŠB - technical university of Ostrava,” *Int. J. Eng. Res. Africa*, vol. 47, pp. 133–138, 2020, doi: 10.4028/www.scientific.net/JERA.47.133.
- [53] A. Adibfar and A. M. Costin, “Creation of a Mock-up Bridge Digital Twin by Fusing Intelligent Transportation Systems (ITS) Data into Bridge Information Model (BrIM),” *J. Constr. Eng. Manag.*, vol. 148, no. 9, pp. 1–11, 2022, doi: 10.1061/(asce)co.1943-7862.0002332.
- [54] N. Moretti, X. Xie, J. Merino, J. Brazauskas, and A. K. Parlikad, “An openbim approach to iot integration with incomplete as-built data,” *Appl. Sci.*, vol. 10, no. 22, pp. 1–17, 2020, doi: 10.3390/app10228287.
- [55] C. Rausch and C. Haas, “Automated shape and pose updating of building information model elements from 3D point clouds,” *Autom. Constr.*, vol. 124, no. August 2020, p. 103561, 2021, doi: 10.1016/j.autcon.2021.103561.
- [56] D. Huynh and S. Nguyen-Ky, “Engaging Building Automation Data Visualisation Using Building Information Modelling and Progressive Web Application,” *Open Eng.*, vol. 10, no. 1, pp. 434–442, 2020, doi: 10.1515/eng-2020-0054.
- [57] W. Wang *et al.*, “BIM Information Integration Based VR Modeling in Digital Twins in Industry 5.0,” *J. Ind. Inf. Integr.*, vol. 28, no. April, p. 100351, 2022, doi: 10.1016/j.jii.2022.100351.
- [58] J. A. Torrecilla-García, M. C. Pardo-Ferreira, and J. C. Rubio-Romero, “Overall Introduction to the Framework of BIM-based Digital Twinning in Decision-making in Safety Management in Building Construction Industry,” *Dir. y Organ.*, vol. 76, no. 76, pp. 5–12, 2022, doi: 10.37610/dyo.v0i76.614.
- [59] Y. Lei, Y. Rao, J. Wu, and C. H. Lin, “BIM based cyber-physical systems for intelligent disaster prevention,” *J. Ind. Inf. Integr.*, vol. 20, 2020, doi: 10.1016/j.jii.2020.100171.
- [60] X. Yuan, C. J. Anumba, and M. K. Parfitt, “Cyber-physical systems for temporary structure monitoring,” *Autom. Constr.*, vol. 66, pp. 1–14, 2016, doi: 10.1016/j.autcon.2016.02.005.
- [61] J. Taraben and G. Morgenthal, “Methods for the Automated Assignment and Comparison of Building Damage Geometries,” *Adv. Eng. Informatics*, vol. 47, no. November 2020, p. 101186, 2021, doi: 10.1016/j.aei.2020.101186.
- [62] Y. Al-Saeed, D. J. Edwards, and S. Scaysbrook, “Automating construction manufacturing procedures using BIM digital objects (BDOs): Case study of knowledge transfer partnership project in UK,” *Constr. Innov.*, vol. 20, no. 3, pp. 345–377, 2020, doi: 10.1108/CI-12-2019-0141.
- [63] C. P. Schimanski, G. P. Monizza, C. Marcher, and D. T. Matt, “Pushing digital automation of

- configure-to-order services in small and medium enterprises of the construction equipment industry: A design science research approach,” *Appl. Sci.*, vol. 9, no. 18, pp. 1–22, 2019, doi: 10.3390/app9183780.
- [64] S. Kosse, O. Vogt, M. Wolf, M. König, and D. Gerhard, “Digital Twin Framework for Enabling Serial Construction,” *Front. Built Environ.*, vol. 8, no. April, pp. 1–20, 2022, doi: 10.3389/fbuil.2022.864722.
- [65] S. Agostinelli, F. Cumo, G. Guidi, and C. Tomazzoli, “Cyber-Physical Systems Improving Building Energy Management: Digital Twin and Artificial Intelligence,” *Energies*, vol. 14, no. 8, p. 2338, 2021, doi: 10.3390/en14082338.
- [66] S. Kaewunruen, P. Rungskunroch, and J. Welsh, “A digital-twin evaluation of Net Zero Energy Building for existing buildings,” *Sustain.*, vol. 11, no. 1, pp. 1–22, 2018, doi: 10.3390/su11010159.
- [67] L. Zhao, H. Zhang, Q. Wang, and H. Wang, “Digital-Twin-Based Evaluation of Nearly Zero-Energy Building for Existing Buildings Based on Scan-to-BIM,” *Adv. Civ. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/6638897.
- [68] F. Banfi, R. Brumana, G. Salvalai, and M. Previtali, “Digital Twin and Cloud BIM-XR Platform Development: From Scan-to-BIM-to-DT Process to a 4D Multi-User Live App to Improve Building Comfort, Efficiency and Costs,” *Energies*, vol. 15, no. 12, 2022, doi: 10.3390/en15124497.
- [69] K. Xing, K. P. Kim, and D. Ness, “Cloud-BIM enabled cyber-physical data and service platforms for building component reuse,” *Sustain.*, vol. 12, no. 24, pp. 1–22, 2020, doi: 10.3390/su122410329.
- [70] S. Kaewunruen and Q. Lian, “Digital twin aided sustainability-based lifecycle management for railway turnout systems,” *J. Clean. Prod.*, vol. 228, pp. 1537–1551, 2019, doi: 10.1016/j.jclepro.2019.04.156.
- [71] G. P. Lydon, S. Caranovic, I. Hischer, and A. Schlueter, “Coupled simulation of thermally active building systems to support a digital twin,” *Energy Build.*, vol. 202, p. 109298, 2019, doi: 10.1016/j.enbuild.2019.07.015.
- [72] P. Kyvelou, C. Buchanan, and L. Gardner, “Numerical simulation and evaluation of the world’s first metal additively manufactured bridge,” *Structures*, vol. 42, no. May, pp. 405–416, 2022, doi: 10.1016/j.istruc.2022.06.012.
- [73] G. Lilis, O. Van Cutsem, and M. Kayal, “A High-Speed Integrated building emulation engine based on discrete event simulation,” *J. Syst. Archit.*, vol. 92, no. August 2018, pp. 53–65, 2019, doi: 10.1016/j.sysarc.2018.10.002.
- [74] M. Omer, L. Margetts, M. Hadi Mosleh, S. Hewitt, and M. Parwaiz, “Use of gaming technology to bring bridge inspection to the office,” *Struct. Infrastruct. Eng.*, vol. 15, no. 10, pp. 1292–1307, 2019, doi: 10.1080/15732479.2019.1615962.
- [75] M. F. Funari, A. E. Hajjat, M. G. Masciotta, D. V. Oliveira, and P. B. Lourenço, “A parametric scan-to-FEM framework for the digital twin generation of historic masonry structures,” *Sustain.*, vol. 13, no. 19, 2021, doi: 10.3390/su131911088.
- [76] F. Xue, W. Lu, Z. Chen, and C. J. Webster, “From LiDAR point cloud towards digital twin city: Clustering city objects based on Gestalt principles,” *ISPRS J. Photogramm. Remote Sens.*, vol. 167, no. July, pp. 418–431, 2020, doi: 10.1016/j.isprsjprs.2020.07.020.
- [77] B. G. Pantoja-Rosero, R. Achanta, M. Kozinski, P. Fua, F. Perez-Cruz, and K. Beyer, “Generating LOD3 building models from structure-from-motion and semantic segmentation,” *Autom. Constr.*, vol. 141, no. July 2021, p. 104430, 2022, doi: 10.1016/j.autcon.2022.104430.
- [78] F. Jiang, L. Ma, T. Broyd, W. Chen, and H. Luo, “Building digital twins of existing highways using map data based on engineering expertise,” *Autom. Constr.*, vol. 134, no. November 2021, p. 104081, 2022, doi: 10.1016/j.autcon.2021.104081.
- [79] M. Münzinger, N. Prechtel, and M. Behnisch, “Mapping the urban forest in detail: From LiDAR point clouds to 3D tree models,” *Urban For. Urban Green.*, vol. 74, no. May, p. 127637, 2022, doi: 10.1016/j.ufug.2022.127637.
- [80] X. Wang, C.-J. Liang, C. C. Menassa, and V. R. Kamat, “Interactive and Immersive Process-Level Digital Twin for Collaborative Human–Robot Construction Work,” *J. Comput. Civ. Eng.*, vol. 35, no. 6, p. 04021023, 2021, doi: 10.1061/(asce)cp.1943-5487.0000988.
- [81] N. Kikuchi, T. Fukuda, and N. Yabuki, “Future landscape visualization using a city digital twin: integration of augmented reality and drones with implementation of 3D model-based occlusion handling,” *J. Comput. Des. Eng.*, vol. 9, no. 2, pp. 837–856, 2022, doi: 10.1093/jcde/qwac032.
- [82] C. S. Shim, N. S. Dang, S. Lon, and C. H. Jeon, “Development of a bridge maintenance system for prestressed concrete bridges using 3D digital twin model,” *Struct. Infrastruct. Eng.*, vol. 15,

- no. 10, pp. 1319–1332, 2019, doi: 10.1080/15732479.2019.1620789.
- [83] M. Antonino, M. Nicola, D. M. Claudio, B. Luciano, and R. C. Fulvio, “Office building occupancy monitoring through image recognition sensors,” *Int. J. Saf. Secur. Eng.*, vol. 9, no. 4, pp. 371–380, 2019, doi: 10.2495/SAFE-V9-N4-371-380.
- [84] Q. Lu, L. Chen, S. Li, and M. Pitt, “Semi-automatic geometric digital twinning for existing buildings based on images and CAD drawings,” *Autom. Constr.*, vol. 115, no. December 2019, p. 103183, 2020, doi: 10.1016/j.autcon.2020.103183.
- [85] H. E. Pang and F. Biljecki, “3D building reconstruction from single street view images using deep learning,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 112, no. May, p. 102859, 2022, doi: 10.1016/j.jag.2022.102859.
- [86] ETHEM Alpaydin, *Intorduction to mahcine learning*. MIT Press, 2014.
- [87] J. Döllner, “Geospatial Artificial Intelligence: Potentials of Machine Learning for 3D Point Clouds and Geospatial Digital Twins,” *PFG - J. Photogramm. Remote Sens. Geoinf. Sci.*, vol. 88, no. 1, pp. 15–24, 2020, doi: 10.1007/s41064-020-00102-3.
- [88] K. Alanne and S. Sierla, “An overview of machine learning applications for smart buildings,” *Sustain. Cities Soc.*, vol. 76, no. October 2021, p. 103445, 2022, doi: 10.1016/j.scs.2021.103445.
- [89] I. C. Konstantakopoulos, A. R. Barkan, S. He, T. Veeravalli, H. Liu, and C. Spanos, “A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure,” *Appl. Energy*, vol. 237, no. September 2018, pp. 810–821, 2019, doi: 10.1016/j.apenergy.2018.12.065.
- [90] M. Austin, P. Delgoshaei, M. Coelho, and M. Heidarinejad, “Architecting Smart City Digital Twins: Combined Semantic Model and Machine Learning Approach,” *J. Manag. Eng.*, vol. 36, no. 4, p. 04020026, 2020, doi: 10.1061/(asce)me.1943-5479.0000774.
- [91] M. Kamari and Y. Ham, “AI-based risk assessment for construction site disaster preparedness through deep learning-based digital twinning,” *Autom. Constr.*, vol. 134, no. October 2021, p. 104091, 2022, doi: 10.1016/j.autcon.2021.104091.
- [92] Z. Liu, A. Li, Z. Sun, G. Shi, and X. Meng, “Digital Twin-Based Risk Control during Prefabricated Building Hoisting Operations,” *Sensors*, vol. 22, no. 7, 2022, doi: 10.3390/s22072522.
- [93] M. M. Gichane *et al.*, “Digital triplet approach for real-time monitoring and control of an elevator security system,” *Designs*, vol. 4, no. 2, pp. 1–14, 2020, doi: 10.3390/designs4020009.
- [94] Y. Pan, A. Braun, I. Brilakis, and A. Borrmann, “Enriching geometric digital twins of buildings with small objects by fusing laser scanning and AI-based image recognition,” *Autom. Constr.*, vol. 140, no. May, 2022, doi: 10.1016/j.autcon.2022.104375.
- [95] J. Zhang, L. Zi, Y. Hou, M. Wang, W. Jiang, and D. Deng, “A Deep Learning-Based Approach to Enable Action Recognition for Construction Equipment,” *Adv. Civ. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/8812928.
- [96] Z. Lv, D. Chen, R. Lou, and A. Alazab, “Artificial intelligence for securing industrial-based cyber–physical systems,” *Futur. Gener. Comput. Syst.*, vol. 117, pp. 291–298, 2021, doi: 10.1016/j.future.2020.12.001.
- [97] G. S. Saini, P. Pournazari, P. Ashok, and E. van Oort, “Intelligent Action Planning for Well Construction Operations Demonstrated for Hole Cleaning Optimization and Automation,” *Energies*, vol. 15, no. 15, 2022, doi: 10.3390/en15155749.
- [98] R. Tariq *et al.*, “Digital twin models for optimization and global projection of building-integrated solar chimney,” *Build. Environ.*, vol. 213, no. January, 2022, doi: 10.1016/j.buildenv.2022.108807.
- [99] P. L. Fernández-Cabán, M. L. Whiteman, B. M. Phillips, F. J. Masters, J. R. Davis, and J. A. Bridge, “Cyber-physical design and optimization of tall building dynamics using aeroelastic wind tunnel modeling,” *J. Wind Eng. Ind. Aerodyn.*, vol. 198, no. January, 2020, doi: 10.1016/j.jweia.2020.104092.
- [100] K. M. Abdelaziz and J. D. Hobeck, “Optimum aeroelastic control via iterative neural network training for wind-resistant cyber–physical buildings,” *Appl. Soft Comput.*, vol. 114, p. 108100, 2022, doi: 10.1016/j.asoc.2021.108100.
- [101] H. Sadri, I. Yitmen, L. C. Tagliabue, and F. Westphal, “a Conceptual Framework for Blockchain and Ai-Driven Digital Twins for Predictive Operation and Maintenance,” *Proc. Eur. Conf. Comput. Constr.*, 2023, doi: 10.35490/EC3.2023.219.
- [102] M. B. Maru, Y. Wang, H. Kim, H. Yoon, and S. Park, “Improved building facade segmentation through digital twin-enabled RandLA-Net with empirical intensity correction model,” *J. Build. Eng.*, vol. 78, no. July, p. 107520, 2023, doi: 10.1016/j.job.2023.107520.
- [103] M. Granzner, A. Strauss, M. Reiterer, M. Cao, and D. Novák, “Data-Driven Condition

- Assessment and Life Cycle Analysis Methods for Dynamically and Fatigue-Loaded Railway Infrastructure Components,” *Infrastructures*, vol. 8, no. 11, 2023, doi: 10.3390/infrastructures8110162.
- [104] N. N. M. Cross-functionally, “Track Geometry Prediction Using Three-Dimensional Recurrent Neural Network-Based Models Cross-Functionally Co-Simulated with BIM,” 2023.
- [105] A. Consilvio *et al.*, “Towards a digital twin-based intelligent decision support for road maintenance,” *Transp. Res. Procedia*, vol. 69, pp. 791–798, 2023, doi: 10.1016/j.trpro.2023.02.237.
- [106] H. H. Hosamo, P. R. Svennevig, K. Svidt, D. Han, and H. K. Nielsen, “A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics,” *Energy Build.*, vol. 261, p. 111988, 2022, doi: 10.1016/j.enbuild.2022.111988.
- [107] H. H. Hosamo, H. K. Nielsen, D. Kraniotis, P. R. Svennevig, and K. Svidt, “Improving building occupant comfort through a digital twin approach: A Bayesian network model and predictive maintenance method,” *Energy Build.*, vol. 288, p. 112992, 2023, doi: 10.1016/j.enbuild.2023.112992.
- [108] J. H. Chan and C. Y. Lau, “Enhancement of Jaibot: Developing Safety and Monitoring Features for Jaibot Using IoT Technologies,” *Int. J. Technol.*, vol. 14, no. 6, pp. 1309–1319, 2023, doi: 10.14716/ijtech.v14i6.6627.
- [109] M. Milašinović, D. Ivetić, M. Stojković, and D. Savić, “Failure Conditions Assessment of Complex Water Systems Using Fuzzy Logic,” *Water Resour. Manag.*, vol. 37, no. 3, pp. 1153–1182, 2023, doi: 10.1007/s11269-022-03420-w.
- [110] V. Weerapura, R. Sugathadasa, M. M. De Silva, I. Nielsen, and A. Thibbotuwawa, “Feasibility of Digital Twins to Manage the Operational Risks in the Production of a Ready-Mix Concrete Plant,” *Buildings*, vol. 13, no. 2, 2023, doi: 10.3390/buildings13020447.
- [111] J. Sresakoolchai and S. Kaewunruen, “Railway infrastructure maintenance efficiency improvement using deep reinforcement learning integrated with digital twin based on track geometry and component defects,” *Sci. Rep.*, vol. 13, no. 1, pp. 1–16, 2023, doi: 10.1038/s41598-023-29526-8.
- [112] X. Kestelyn, G. Denis, V. Champany, N. Hascoet, C. Ghnatios, and F. Chinesta, “Towards a Hybrid Twin for Infrastructure Asset Management : Investigation on Power Transformer Asset Maintenance Management,” *2022 7th Int. Adv. Res. Work. Transform.*, no. October, pp. 109–114, 2022, doi: 10.23919/ARWtr54586.2022.9959914.
- [113] C. S. Wickramasinghe, K. Amarasinghe, D. L. Marino, C. Rieger, and M. Manic, “Explainable Unsupervised Machine Learning for Cyber-Physical Systems,” *IEEE Access*, vol. 9, pp. 131824–131843, 2021, doi: 10.1109/ACCESS.2021.3112397.
- [114] H. H. Hosamo, H. K. Nielsen, D. Kraniotis, P. R. Svennevig, and K. Svidt, “Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings,” *Energy Build.*, vol. 281, p. 112732, 2023, doi: 10.1016/j.enbuild.2022.112732.
- [115] S. Agostinelli, “ACTIONABLE FRAMEWORK for CITY DIGITAL TWIN-ENABLED PREDICTIVE MAINTENANCE and SECURITY MANAGEMENT SYSTEMS,” *WIT Trans. Built Environ.*, vol. 205, pp. 223–233, 2021, doi: 10.2495/BIM210191.
- [116] C. C. Lin, D. J. Deng, C. H. Kuo, and L. Chen, “Concept drift detection and adaption in big imbalance industrial IoT data using an ensemble learning method of offline classifiers,” *IEEE Access*, vol. 7, pp. 56198–56207, 2019, doi: 10.1109/ACCESS.2019.2912631.
- [117] W. Hu, X. Wang, K. Tan, and Y. Cai, “Digital twin-enhanced predictive maintenance for indoor climate: A parallel LSTM-autoencoder failure prediction approach,” *Energy Build.*, vol. 301, no. September, p. 113738, 2023, doi: 10.1016/j.enbuild.2023.113738.
- [118] S. Agostinelli and F. Cumo, “Machine Learning Approach for Predictive Maintenance in an Advanced Building Management System,” *WIT Trans. Ecol. Environ.*, vol. 255, pp. 131–138, 2022, doi: 10.2495/EPM220111.
- [119] F. Shalabi and Y. Turkan, “Bim-energy simulation approach for detecting building spaces with faults and problematic behavior,” *J. Inf. Technol. Constr.*, vol. 25, no. May 2019, pp. 342–360, 2020, doi: 10.36680/JITCON.2020.020.
- [120] V. Villa, B. Naticchia, G. Bruno, K. Aliev, P. Piantanida, and D. Antonelli, “Iot open-source architecture for the maintenance of building facilities,” *Appl. Sci.*, vol. 11, no. 12, 2021, doi: 10.3390/app11125374.
- [121] J. C. P. Cheng, W. Chen, K. Chen, and Q. Wang, “Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms,” *Autom. Constr.*, vol. 112, no. December 2019, p. 103087, 2020, doi:

- 10.1016/j.autcon.2020.103087.
- [122] P. Panfilov, “29th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2018,” *Ann. DAAAM Proc. Int. DAAAM Symp.*, vol. 29, no. 1, 2018, doi: 10.2507/29th.daaam.proceedings.xxx.
- [123] Y. Bouabdallaoui, Z. Lafhaj, P. Yim, L. Ducoulombier, and B. Bennadji, “Predictive maintenance in building facilities: A machine learning-based approach,” *Sensors (Switzerland)*, vol. 21, no. 4, pp. 1–15, 2021, doi: 10.3390/s21041044.
- [124] A. R. Santiago, M. Antunes, J. P. Barraca, D. Gomes, and R. L. Aguiar, “Predictive maintenance system for efficiency improvement of heating equipment,” *Proc. - 5th IEEE Int. Conf. Big Data Serv. Appl. BigDataService 2019, Work. Big Data Water Resour. Environ. Hydraul. Eng. Work. Medical, Heal. Using Big Data Technol.*, no. April, pp. 93–98, 2019, doi: 10.1109/BigDataService.2019.00019.
- [125] M. N. Razali *et al.*, “Big data analytics for preventive maintenance management,” *Plan. Malaysia*, vol. 19, no. 3, pp. 423–437, 2021, doi: 10.21837/PM.V19I17.1019.
- [126] R. Chandra *et al.*, “A Survey of Failure Mechanisms and Statistics for Critical Electrical Equipment in Buildings,” *IECON Proc. (Industrial Electron. Conf.)*, vol. 2020-October, pp. 1955–1961, 2020, doi: 10.1109/IECON43393.2020.9254225.
- [127] H. M. Chen, C. C. Hou, and Y. H. Wang, “A 3D visualized expert system for maintenance and management of existing building facilities using reliability-based method,” *Expert Syst. Appl.*, vol. 40, no. 1, pp. 287–299, 2013, doi: 10.1016/j.eswa.2012.07.045.
- [128] R. Ahmed, T. Zayed, and F. Nasiri, “A hybrid genetic algorithm-based fuzzy markovian model for the deterioration modeling of healthcare facilities,” *Algorithms*, vol. 13, no. 9, 2020, doi: 10.3390/a13090210.
- [129] C. Nzukam, D. Sauter, A. Voisin, and E. Levrat, “Performances evaluation in view of Predictive Maintenance-A case study,” *Conf. Control Fault-Tolerant Syst. SysTol*, pp. 226–231, 2019, doi: 10.1109/SYSTOL.2019.8864798.
- [130] A. Staino, R. Abou-Eid, and P. Dersin, “A Monte-Carlo approach for prognostics of clogging process in HVAC filters using a hybrid strategy: A real case study in railway systems,” *2018 IEEE Int. Conf. Progn. Heal. Manag. ICPHM 2018*, pp. 1–8, 2018, doi: 10.1109/ICPHM.2018.8448706.
- [131] C. Yang, W. Shen, Q. Chen, and B. Gunay, “A practical solution for HVAC prognostics: Failure mode and effects analysis in building maintenance,” *J. Build. Eng.*, vol. 15, no. July 2017, pp. 26–32, 2018, doi: 10.1016/j.jobe.2017.10.013.
- [132] A. Gálvez, A. Diez-Olivan, D. Seneviratne, and D. Galar, “Fault detection and RUL estimation for railway HVAC systems using a hybrid model-based approach,” *Sustain.*, vol. 13, no. 12, 2021, doi: 10.3390/su13126828.
- [133] Q. P. Yang, J. J. Mesa-Jiménez, L. Stokes, Q. P. Yang, and V. Livina, “Early warning signals of failures in building management systems,” *Int. J. Metrol. Qual. Eng.*, vol. 12, 2021, doi: 10.1051/ijmqe/2021009.
- [134] I. S. Candanedo, E. H. Nieves, S. R. González, M. T. S. Martín, and A. G. Briones, “Machine learning predictive model for industry 4.0,” *Commun. Comput. Inf. Sci.*, vol. 877, no. July, pp. 501–510, 2018, doi: 10.1007/978-3-319-95204-8_42.
- [135] C. Yang, B. Gunay, Z. Shi, and W. Shen, “Machine Learning-Based Prognostics for Central Heating and Cooling Plant Equipment Health Monitoring,” *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 1, pp. 346–355, 2021, doi: 10.1109/TASE.2020.2998586.
- [136] S. Trivedi, S. Bhola, A. Talegaonkar, P. Gaur, and S. Sharma, “Predictive Maintenance of Air Conditioning Systems Using Supervised Machine Learning,” *2019 20th Int. Conf. Intell. Syst. Appl. to Power Syst. ISAP 2019*, pp. 1–6, 2019, doi: 10.1109/ISAP48318.2019.9065995.
- [137] R. Li, S. Pan, H. Fang, Y. Xiong, and F. Wang, “Fault Prediction Technology of Civil Aircraft Based on Qar Data,” *Proc. - 2017 Int. Conf. Sensing, Diagnostics, Progn. Control. SDPC 2017*, vol. 2017-December, pp. 468–472, 2017, doi: 10.1109/SDPC.2017.94.
- [138] K. M. Mishra and K. Huhtala, “Elevator fault detection using profile extraction and deep autoencoder feature extraction for acceleration and magnetic signals,” *Appl. Sci.*, vol. 9, no. 15, 2019, doi: 10.3390/app9152990.
- [139] X. Ma, L. Chengkai, K. H. Ng, and H. P. Tan, “An Internet of Things-Based Lift Predictive Maintenance System,” *IEEE Potentials*, vol. 40, no. 1, pp. 17–23, 2021, doi: 10.1109/MPOT.2020.2973697.
- [140] L. H. Li, C. Y. Lai, F. H. Kuo, and P. Y. Chai, “Predictive maintenance of vertical lift storage motor based on machine learning,” *Int. J. Appl. Sci. Eng.*, vol. 16, no. 2, pp. 109–118, 2019, doi: 10.6703/IJASE.201909_16(2).109.

- [141] Y. T. Quek, W. L. Woo, and L. Thillainathan, "IoT Load Classification and Anomaly Warning in ELV DC Picogrids Using Hierarchical Extended k -Nearest Neighbors," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 863–873, 2020, doi: 10.1109/JIOT.2019.2945425.
- [142] S. Assaf and I. Srour, "Using a data driven neural network approach to forecast building occupant complaints," *Build. Environ.*, vol. 200, no. May, p. 107972, 2021, doi: 10.1016/j.buildenv.2021.107972.
- [143] F. J. Ferrández-Pastor, J. M. García-Chamizo, S. Gomez-Trillo, R. Valdivieso-Sarabia, and M. Nieto-Hidalgo, "Smart management consumption in renewable energy fed ecosystems†," *Sensors (Switzerland)*, vol. 19, no. 13, pp. 1–28, 2019, doi: 10.3390/s19132967.
- [144] A. Gouda Mohamed and M. Marzouk, "Building condition assessment using artificial neural network and structural equations," *Expert Syst. Appl.*, vol. 186, no. May, p. 115743, 2021, doi: 10.1016/j.eswa.2021.115743.
- [145] R. Ahmed, F. Nasiri, and T. Zayed, "Two-Stage Predictive Maintenance Planning for Hospital Buildings: A Multiple-Objective Optimization-Based Clustering Approach," *J. Perform. Constr. Facil.*, vol. 36, no. 1, pp. 1–11, 2022, doi: 10.1061/(asce)cf.1943-5509.0001691.
- [146] R. Ahmed, F. Nasiri, and T. Zayed, "A novel Neutrosophic-based machine learning approach for maintenance prioritization in healthcare facilities," *J. Build. Eng.*, vol. 42, no. November 2020, p. 102480, 2021, doi: 10.1016/j.jobe.2021.102480.
- [147] S. Peiris and N. De Silva, "A fuzzy-logic model for benchmarking concrete roof maintenance," *Facilities*, vol. 39, no. 7–8, pp. 568–583, 2021, doi: 10.1108/F-03-2020-0026.
- [148] J. Xia, Y. Feng, C. Lu, C. Fei, and X. Xue, "LSTM-based multi-layer self-attention method for remaining useful life estimation of mechanical systems," *Eng. Fail. Anal.*, vol. 125, no. December 2020, p. 105385, 2021, doi: 10.1016/j.engfailanal.2021.105385.
- [149] Y. Wei, J. Jang-Jaccard, W. Xu, F. Sabrina, S. Camtepe, and M. Boulic, "LSTM-Autoencoder-Based Anomaly Detection for Indoor Air Quality Time-Series Data," *IEEE Sens. J.*, vol. 23, no. 4, pp. 3787–3800, 2023, doi: 10.1109/JSEN.2022.3230361.
- [150] B. Li and W. Cai, "A novel CO₂-based demand-controlled ventilation strategy to limit the spread of COVID-19 in the indoor environment," *Build. Environ.*, vol. 219, no. February, p. 109232, 2022, doi: 10.1016/j.buildenv.2022.109232.
- [151] Y. S. Jung, T. W. Kang, and C. Chun, "Anomaly analysis on indoor office spaces for facility management using deep learning methods," *J. Build. Eng.*, vol. 43, p. 103139, 2021, doi: 10.1016/j.jobe.2021.103139.
- [152] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, and A. Amira, "Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives," *Appl. Energy*, vol. 287, no. November 2020, p. 116601, 2021, doi: 10.1016/j.apenergy.2021.116601.
- [153] A. Malki, E. S. Atlam, and I. Gad, "Machine learning approach of detecting anomalies and forecasting time-series of IoT devices," *Alexandria Eng. J.*, vol. 61, no. 11, pp. 8973–8986, 2022, doi: 10.1016/j.aej.2022.02.038.
- [154] A. Copiaco *et al.*, "An innovative deep anomaly detection of building energy consumption using energy time-series images," *Eng. Appl. Artif. Intell.*, vol. 119, no. January, p. 105775, 2023, doi: 10.1016/j.engappai.2022.105775.
- [155] D. Helman, Y. Yungstein, G. Mulero, and Y. Michael, "High-Throughput Remote Sensing of Vertical Green Living Walls (VGWs) in Workplaces," *Remote Sens.*, vol. 14, no. 14, 2022, doi: 10.3390/rs14143485.
- [156] C. J. Turner, J. Oyekan, L. Stergioulas, and D. Griffin, "Utilizing Industry 4.0 on the Construction Site: Challenges and Opportunities," *IEEE Trans. Ind. Informatics*, vol. 17, no. 2, pp. 746–756, 2021, doi: 10.1109/TII.2020.3002197.
- [157] E. Bradley, *Reliability engineering : a life cycle approach*. Boca Raton : CRC Press, 2016, 2016. [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1768608>
- [158] P. Vähä, T. Heikkilä, P. Kilpeläinen, M. Järviluoma, and E. Gambao, "Extending automation of building construction - Survey on potential sensor technologies and robotic applications," *Autom. Constr.*, vol. 36, pp. 168–178, 2013, doi: 10.1016/j.autcon.2013.08.002.
- [159] Y. Zhang, G. Gong, H. Yang, W. Li, and J. Liu, "Precision versus intelligence: Autonomous supporting pressure balance control for slurry shield tunnel boring machines," *Autom. Constr.*, vol. 114, no. March, pp. 1–11, 2020, doi: 10.1016/j.autcon.2020.103173.
- [160] M. Li, Q. Lu, S. Bai, M. Zhang, H. Tian, and L. Qin, "Digital twin-driven virtual sensor approach for safe construction operations of trailing suction hopper dredger," *Autom. Constr.*, vol. 132, no.

- September, p. 103961, 2021, doi: 10.1016/j.autcon.2021.103961.
- [161] J. Zhang, H. Luo, and J. Xu, “Towards fully BIM-enabled building automation and robotics: A perspective of lifecycle information flow,” *Comput. Ind.*, vol. 135, p. 103570, 2022, doi: 10.1016/j.compind.2021.103570.
- [162] C. J. Liang, V. R. Kamat, and C. C. Menassa, “Teaching robots to perform quasi-repetitive construction tasks through human demonstration,” *Autom. Constr.*, vol. 120, no. July, 2020, doi: 10.1016/j.autcon.2020.103370.
- [163] Y. Zhang, A. Meina, X. Lin, K. Zhang, and Z. Xu, “Digital twin in computational design and robotic construction of wooden architecture,” *Adv. Civ. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/8898997.
- [164] D. Lee, S. H. Lee, N. Masoud, M. S. Krishnan, and V. C. Li, “Digital twin-driven deep reinforcement learning for adaptive task allocation in robotic construction,” *Adv. Eng. Informatics*, vol. 53, no. November 2020, p. 101710, 2022, doi: 10.1016/j.aei.2022.101710.
- [165] W. Jiang, L. Ding, and C. Zhou, “Digital twin: Stability analysis for tower crane hoisting safety with a scale model,” *Autom. Constr.*, vol. 138, no. April 2021, 2022, doi: 10.1016/j.autcon.2022.104257.
- [166] Z. S. Liu, X. T. Meng, Z. Z. Xing, C. F. Cao, Y. Y. Jiao, and A. X. Li, “Digital Twin-Based Intelligent Safety Risks Prediction of Prefabricated Construction Hoisting,” *Sustain.*, vol. 14, no. 9, 2022, doi: 10.3390/su14095179.
- [167] C. Kan, Y. Fang, C. J. Anumba, and J. I. Messner, “A cyber-physical system (CPS) for planning and monitoring mobile cranes on construction sites,” *Proc. Inst. Civ. Eng. Manag. Procure. Law*, vol. 171, no. 6, pp. 240–250, 2018, doi: 10.1680/jmapl.17.00042.
- [168] S. Wu, L. Hou, G. (Kevin) Zhang, and H. Chen, “Real-time mixed reality-based visual warning for construction workforce safety,” *Autom. Constr.*, vol. 139, no. December 2021, p. 104252, 2022, doi: 10.1016/j.autcon.2022.104252.
- [169] A. A. Akanmu, J. Olayiwola, O. Ogunsejju, and D. McFeeters, “Cyber-physical postural training system for construction workers,” *Autom. Constr.*, vol. 117, no. November 2019, p. 103272, 2020, doi: 10.1016/j.autcon.2020.103272.
- [170] S. M. E. Sepasgozar, “Digital twin and web-based virtual gaming technologies for online education: A case of construction management and engineering,” *Appl. Sci.*, vol. 10, no. 13, 2020, doi: 10.3390/app10134678.
- [171] R. He, M. Li, V. J. L. Gan, and J. Ma, “BIM-enabled computerized design and digital fabrication of industrialized buildings: A case study,” *J. Clean. Prod.*, vol. 278, p. 123505, 2021, doi: 10.1016/j.jclepro.2020.123505.
- [172] J. Orozco-Messana, V. Lopez-Mateu, and T. M. Pellicer, “City Regeneration through Modular Phase Change Materials (PCM) Envelopes for Climate Neutral Buildings,” *Sustain.*, vol. 14, no. 14, 2022, doi: 10.3390/su14148902.
- [173] M. Marini *et al.*, “An easily integrable industrial system for gamma spectroscopic analysis and traceability of stones and building materials,” *Sensors (Switzerland)*, vol. 21, no. 2, pp. 1–16, 2021, doi: 10.3390/s21020352.
- [174] S. Meža, A. Mauko Pranjić, R. Vežočanik, I. Osmokrović, and S. Lenart, “Digital Twins and Road Construction Using Secondary Raw Materials,” *J. Adv. Transp.*, vol. 2021, 2021, doi: 10.1155/2021/8833058.
- [175] S. Züst, R. Züst, V. Züst, S. West, O. Stoll, and C. Minonne, “A Graph Based Monte Carlo Simulation supporting a Digital Twin for the Curatorial Management of Excavation and Demolition Material Flows,” *J. Clean. Prod.*, vol. 310, no. May, p. 127453, 2021, doi: 10.1016/j.jclepro.2021.127453.
- [176] J. L. Hao *et al.*, “Carbon emission reduction in prefabrication construction during materialization stage: A BIM-based life-cycle assessment approach,” *Sci. Total Environ.*, vol. 723, p. 137870, 2020, doi: 10.1016/j.scitotenv.2020.137870.
- [177] M. Chiachío, M. Megía, J. Chiachío, J. Fernandez, and M. L. Jalón, “Structural digital twin framework: Formulation and technology integration,” *Autom. Constr.*, vol. 140, no. May, 2022, doi: 10.1016/j.autcon.2022.104333.
- [178] N. Moussallam, R. Ziegler, J. Rudolph, and S. Bergholz, “A Method for Monitoring Vibrational Fatigue of Structure and Components,” *J. Press. Vessel Technol. Trans. ASME*, vol. 144, no. 3, 2022, doi: 10.1115/1.4053380.
- [179] H. Zhu and Y. Wang, “Key Component Capture and Safety Intelligent Analysis of Beam String Structure Based on Digital Twins,” *Symmetry (Basel)*, vol. 14, no. 6, 2022, doi: 10.3390/sym14061152.
- [180] E. García-Macías and F. Ubertini, “Real-time Bayesian damage identification enabled by sparse

- PCE-Kriging meta-modelling for continuous SHM of large-scale civil engineering structures,” *J. Build. Eng.*, vol. 59, no. August, 2022, doi: 10.1016/j.jobbe.2022.105004.
- [181] M. Riggio, M. Mrissa, M. Krész, J. Včelák, J. Sandak, and A. Sandak, “Leveraging Structural Health Monitoring Data Through Avatars to Extend the Service Life of Mass Timber Buildings,” *Front. Built Environ.*, vol. 8, no. June, pp. 1–7, 2022, doi: 10.3389/fbuil.2022.887593.
- [182] I. John Samuel, O. Salem, and S. He, “Defect-oriented supportive bridge inspection system featuring building information modeling and augmented reality,” *Innov. Infrastruct. Solut.*, vol. 7, no. 4, 2022, doi: 10.1007/s41062-022-00847-3.
- [183] W. Flores-Fuentes, “Bridge Load Classifier Based on Deep Learning for Structural Displacement Correlation,” *Program. Comput. Softw.*, vol. 46, no. 8, pp. 526–535, 2020, doi: 10.1134/S0361768820080101.
- [184] F. Jiang, L. Ma, T. Broyd, K. Chen, and H. Luo, “Underpass clearance checking in highway widening projects using digital twins,” *Autom. Constr.*, vol. 141, no. May, p. 104406, 2022, doi: 10.1016/j.autcon.2022.104406.
- [185] T. Han, T. Ma, Z. Fang, Y. Zhang, and C. Han, “A BIM-IoT and intelligent compaction integrated framework for advanced road compaction quality monitoring and management,” *Comput. Electr. Eng.*, vol. 100, no. March, pp. 1–12, 2022, doi: 10.1016/j.compeleceng.2022.107981.
- [186] G. Angjeliu, D. Coronelli, and G. Cardani, “Development of the simulation model for Digital Twin applications in historical masonry buildings: The integration between numerical and experimental reality,” *Comput. Struct.*, vol. 238, p. 106282, 2020, doi: 10.1016/j.compstruc.2020.106282.
- [187] Q. Kong, J. Zhu, S. C. M. Ho, and G. Song, “Tapping and listening: A new approach to bolt looseness monitoring,” *Smart Mater. Struct.*, vol. 27, no. 7, p. 07LT02, 2018, doi: 10.1088/1361-665X/aac962.
- [188] M. Mahmoodian, F. Shahrivar, S. Setunge, and S. Mazaheri, “Development of Digital Twin for Intelligent Maintenance of Civil Infrastructure,” *Sustain.*, vol. 14, no. 14, 2022, doi: 10.3390/su14148664.
- [189] J. Zhang, J. C. P. Cheng, W. Chen, and K. Chen, “Digital Twins for Construction Sites: Concepts, LoD Definition, and Applications,” *J. Manag. Eng.*, vol. 38, no. 2, pp. 1–16, 2022, doi: 10.1061/(asce)me.1943-5479.0000948.
- [190] Y. Jiang, M. Li, D. Guo, W. Wu, R. Y. Zhong, and G. Q. Huang, “Digital twin-enabled smart modular integrated construction system for on-site assembly,” *Comput. Ind.*, vol. 136, p. 103594, 2022, doi: 10.1016/j.compind.2021.103594.
- [191] Y. Jiang *et al.*, “Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly,” *Autom. Constr.*, vol. 141, no. May, 2022, doi: 10.1016/j.autcon.2022.104397.
- [192] Y. Zhou, Y. Su, Z. Xu, X. Wang, J. Wu, and X. Guan, “A hybrid physics-based/data-driven model for personalized dynamic thermal comfort in ordinary office environment,” *Energy Build.*, vol. 238, p. 110790, 2021, doi: 10.1016/j.enbuild.2021.110790.
- [193] A. Karbasi and A. Farhadi, “A cyber-physical system for building automation and control based on a distributed MPC with an efficient method for communication,” *Eur. J. Control.*, vol. 61, pp. 151–170, 2021, doi: 10.1016/j.ejcon.2021.04.008.
- [194] A. Clausen *et al.*, “A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings,” *Energy Informatics*, vol. 4, no. Suppl 2, 2021, doi: 10.1186/s42162-021-00153-9.
- [195] N. Zhou, A. Chang-Richards, K. I. K. Wang, and K. N. Dirks, “Prototyping an IoT-based system for monitoring building indoor environment,” *J. Eng. Des. Technol.*, vol. 20, no. 4, pp. 965–992, 2022, doi: 10.1108/JEDT-06-2021-0320.
- [196] M. Shahinmoghadam, W. Natephra, and A. Motamedi, “BIM- and IoT-based virtual reality tool for real-time thermal comfort assessment in building enclosures,” *Build. Environ.*, vol. 199, no. April, p. 107905, 2021, doi: 10.1016/j.buildenv.2021.107905.
- [197] J. Zhang, H. H. L. Kwok, H. Luo, J. C. K. Tong, and J. C. P. Cheng, “Automatic relative humidity optimization in underground heritage sites through ventilation system based on digital twins,” *Build. Environ.*, vol. 216, no. January, 2022, doi: 10.1016/j.buildenv.2022.108999.
- [198] A. Bonci, A. Carbonari, A. Cucchiarelli, L. Messi, M. Pirani, and M. Vaccarini, “A cyber-physical system approach for building efficiency monitoring,” *Autom. Constr.*, vol. 102, no. August 2018, pp. 68–85, 2019, doi: 10.1016/j.autcon.2019.02.010.
- [199] L. Gomes, C. Almeida, and Z. Vale, “Recommendation of workplaces in a coworking building: A cyber-physical approach supported by a context-aware multi-agent system,” *Sensors*

- (Switzerland), vol. 20, no. 12, pp. 1–13, 2020, doi: 10.3390/s20123597.
- [200] E. Seghezzi *et al.*, “Towards an occupancy-oriented digital twin for facility management: Test campaign and sensors assessment,” *Appl. Sci.*, vol. 11, no. 7, 2021, doi: 10.3390/app11073108.
- [201] G. Schrotter and C. Hürzeler, “The Digital Twin of the City of Zurich for Urban Planning,” *PFG - J. Photogramm. Remote Sens. Geoinf. Sci.*, vol. 88, no. 1, pp. 99–112, 2020, doi: 10.1007/s41064-020-00092-2.
- [202] G. White, A. Zink, L. Codecá, and S. Clarke, “A digital twin smart city for citizen feedback,” *Cities*, vol. 110, no. November 2020, 2021, doi: 10.1016/j.cities.2020.103064.
- [203] Y. Cho and J. Kim, “A study on setting the direction of digital twin implementation for urban regeneration business,” *Int. J. Adv. Appl. Sci.*, vol. 9, no. 4, pp. 147–154, 2022, doi: 10.21833/ijaas.2022.04.018.
- [204] J. Wu, X. Wang, Y. Dang, and Z. Lv, “Digital twins and artificial intelligence in transportation infrastructure: Classification, application, and future research directions,” *Comput. Electr. Eng.*, vol. 101, no. March, p. 107983, 2022, doi: 10.1016/j.compeleceng.2022.107983.
- [205] Y. Wei, Z. Lei, and S. Altaf, “An Off-Site Construction Digital Twin Assessment Framework Using Wood Panelized Construction as a Case Study,” *Buildings*, vol. 12, no. 5, 2022, doi: 10.3390/buildings12050566.
- [206] J. Böke, U. Knaack, and M. Hemmerling, “Automated adaptive façade functions in practice - Case studies on office buildings,” *Autom. Constr.*, vol. 113, no. February, 2020, doi: 10.1016/j.autcon.2020.103113.
- [207] F. Ding and A. Kareem, “Tall Buildings with Dynamic Facade Under Winds,” *Engineering*, vol. 6, no. 12, pp. 1443–1453, 2020, doi: 10.1016/j.eng.2020.07.020.
- [208] G. B. Porsani, K. D. V. de Lersundi, A. S. O. Gutiérrez, and C. F. Bandera, “Interoperability between building information modelling (Bim) and building energy model (bem),” *Appl. Sci.*, vol. 11, no. 5, pp. 1–20, 2021, doi: 10.3390/app11052167.
- [209] M. Demianenko and C. I. De Gaetani, “A Procedure for Automating Energy Analyses in the BIM Context Exploiting Artificial Neural Networks and Transfer Learning Technique,” 2021, doi: <https://doi.org/10.3390/en14102956>.
- [210] R. Jia *et al.*, “Design Automation for Smart Building Systems,” *Proc. IEEE*, vol. 106, no. 9, pp. 1680–1699, 2018, doi: 10.1109/JPROC.2018.2856932.
- [211] W. Jiang, L. Ding, and C. Zhou, “Cyber physical system for safety management in smart construction site,” *Eng. Constr. Archit. Manag.*, vol. 28, no. 3, pp. 788–808, 2021, doi: 10.1108/ECAM-10-2019-0578.
- [212] B. Hoisting, “Digital Twin-Based Safety Risk Coupling of Prefabricated Building Hoisting,” 2021.
- [213] T. Greif, N. Stein, and C. M. Flath, “Peeking into the void: Digital twins for construction site logistics,” *Comput. Ind.*, vol. 121, p. 103264, 2020, doi: 10.1016/j.compind.2020.103264.
- [214] H. Tran, T. N. Nguyen, P. Christopher, D. K. Bui, K. Khoshelham, and T. D. Ngo, “A digital twin approach for geometric quality assessment of as-built prefabricated façades,” *J. Build. Eng.*, vol. 41, no. December 2020, p. 102377, 2021, doi: 10.1016/j.jobbe.2021.102377.
- [215] B. Yang, Z. Lv, and F. Wang, “Digital Twins for Intelligent Green Buildings,” *Buildings*, vol. 12, no. 6, 2022, doi: 10.3390/buildings12060856.
- [216] S. Çetin, V. Gruis, and A. Straub, “Digitalization for a circular economy in the building industry: Multiple-case study of Dutch social housing organizations,” *Resour. Conserv. Recycl. Adv.*, vol. 15, no. August, p. 200110, 2022, doi: 10.1016/j.rcreadv.2022.200110.
- [217] RIBA, “RIBA Plan of Work 2020 overview,” 2020. doi: 10.4324/9780429347177-2.
- [218] J. Zhao, H. Feng, Q. Chen, and B. Garcia de Soto, “Developing a conceptual framework for the application of digital twin technologies to revamp building operation and maintenance processes,” *J. Build. Eng.*, vol. 49, no. October 2021, pp. 1–12, 2022, doi: 10.1016/j.jobbe.2022.104028.
- [219] D. Quirk, J. Lanni, and N. Chauhan, “Digital twins: Details of implementation,” *ASHRAE J.*, vol. 62, no. 10, pp. 20–24, 2020.
- [220] Q. Lu, X. Xie, A. K. Parlikad, and J. M. Schooling, “Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance,” *Autom. Constr.*, vol. 118, no. March, p. 103277, 2020, doi: 10.1016/j.autcon.2020.103277.
- [221] Z. Wu, Y. Chang, Q. Li, and R. Cai, “A Novel Method for Tunnel Digital Twin Construction and Virtual-Real Fusion Application,” *Electron.*, vol. 11, no. 9, 2022, doi: 10.3390/electronics11091413.
- [222] M. Pregnolato *et al.*, “Towards Civil Engineering 4.0: Concept, workflow and application of Digital Twins for existing infrastructure,” *Autom. Constr.*, vol. 141, no. July, p. 104421, 2022, doi: 10.1016/j.autcon.2022.104421.

- [223] B. Keskin, B. Salman, and O. Koseoglu, "Architecting a BIM-Based Digital Twin Platform for Airport Asset Management: A Model-Based System Engineering with SysML Approach," *J. Constr. Eng. Manag.*, vol. 148, no. 5, pp. 1–15, 2022, doi: 10.1061/(asce)co.1943-7862.0002271.
- [224] P. P. Oliveira, "Digital twin development for airport management," *J. Airt. Manag.*, vol. 14, no. 3, pp. 246–259, 2020.
- [225] A. Francisco, N. Mohammadi, and J. E. Taylor, "Smart City Digital Twin–Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking," *J. Manag. Eng.*, vol. 36, no. 2, p. 04019045, 2020, doi: 10.1061/(asce)me.1943-5479.0000741.
- [226] B. L. Thomas and D. J. Cook, "Activity-aware energy-efficient automation of smart buildings," *Energies*, vol. 9, no. 8, 2016, doi: 10.3390/en9080624.
- [227] J. Henzel, Ł. Wróbel, M. Fice, and M. Sikora, "Energy Consumption Forecasting for the Digital-Twin Model of the Building," *Energies*, vol. 15, no. 12, pp. 1–21, 2022, doi: 10.3390/en15124318.
- [228] B. Montalbán Pozas, B. Muriel Holgado, M. Lucas Bonilla, S. Barroso Ramírez, and P. Bustos García de Castro, "Iterative Optimization of a Social Inmotics-Based Method in Order to Make Buildings Smart and Resilient," *Sustain. Cities Soc.*, vol. 82, no. January, 2022, doi: 10.1016/j.scs.2022.103876.
- [229] U. Alibrandi, "Risk-Informed Digital Twin of Buildings and Infrastructures for Sustainable and Resilient Urban Communities," *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.*, vol. 8, no. 3, pp. 1–28, 2022, doi: 10.1061/ajrua6.0001238.
- [230] B. Bass, J. New, and W. Copeland, "Potential Energy, Demand, Emissions, and Cost Savings Distributions for Buildings in a Utility 's Service Area," 2021.
- [231] W. Huang, Y. Zhang, and W. Zeng, "Development and application of digital twin technology for integrated regional energy systems in smart cities," *Sustain. Comput. Informatics Syst.*, vol. 36, no. August, p. 100781, 2022, doi: 10.1016/j.suscom.2022.100781.
- [232] RIBA, "RIBA Sustainable outcomes guide," 2019. [Online]. Available: <https://www.architecture.com/-/media/GatherContent/Test-resources-page/Additional-Documents/RIBASustainableOutcomesGuide2019pdf.pdf>
- [233] S. Hosseini Haghghi, P. M. Á. de Urbarri, R. Padsala, and U. Eicker, "Characterizing and structuring urban GIS data for housing stock energy modelling and retrofitting," *Energy Build.*, vol. 256, p. 111706, 2022, doi: 10.1016/j.enbuild.2021.111706.
- [234] I. Yitmen, S. Alizadehsalehi, İ. Akiner, and M. E. Akiner, "An Adapted Model of Cognitive Digital Twins for Building Lifecycle Management," *Appl. Sci.*, vol. 11, no. 9, p. 4276, 2021, doi: 10.3390/app11094276.
- [235] Y. Peng, M. Zhang, F. Yu, J. Xu, and S. Gao, "Digital Twin Hospital Buildings: An Exemplary Case Study through Continuous Lifecycle Integration," *Adv. Civ. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/8846667.
- [236] S. Kaewunruen, S. Peng, and O. Phil-Ebosie, "Digital twin aided sustainability and vulnerability audit for subway stations," *Sustain.*, vol. 12, no. 19, pp. 1–17, 2020, doi: 10.3390/SU12197873.
- [237] S. Kaewunruen, J. Sresakoolchai, W. Ma, and O. Phil-Ebosie, "Digital twin aided vulnerability assessment and risk-based maintenance planning of bridge infrastructures exposed to extreme conditions," *Sustain.*, vol. 13, no. 4, pp. 1–19, 2021, doi: 10.3390/su13042051.
- [238] Y. Zu and R. Dai, "Distributed path planning for building evacuation guidance," *Cyber-Physical Syst.*, vol. 3, no. 1–4, pp. 1–21, 2017, doi: 10.1080/23335777.2017.1326983.
- [239] P. Cai, Y. Cai, I. Chandrasekaran, and J. Zheng, "Parallel genetic algorithm based automatic path planning for crane lifting in complex environments," *Autom. Constr.*, vol. 62, pp. 133–147, 2016, doi: 10.1016/j.autcon.2015.09.007.
- [240] Y. Pan and L. Zhang, "Roles of artificial intelligence in construction engineering and management: A critical review and future trends," *Autom. Constr.*, vol. 122, no. December 2020, p. 103517, 2021, doi: 10.1016/j.autcon.2020.103517.
- [241] P. Li, Y. Lu, D. Yan, J. Xiao, and H. Wu, "Scientometric mapping of smart building research: Towards a framework of human-cyber-physical system (HCPS)," *Autom. Constr.*, vol. 129, no. 1239, p. 103776, 2021, doi: 10.1016/j.autcon.2021.103776.
- [242] R. Sacks, I. Brilakis, E. Piskas, H. S. Xie, and M. Girolami, "Construction with digital twin information systems," *Data-Centric Eng.*, vol. 1, 2020, doi: 10.1017/dce.2020.16.
- [243] Z. Wang *et al.*, "Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase," *J. Comput. Civ. Eng.*, vol. 35, no. 1, p. 04020056, 2021, doi: 10.1061/(asce)cp.1943-5487.0000933.
- [244] M. Jia, A. Komeily, Y. Wang, and R. S. Srinivasan, "Adopting Internet of Things for the development of smart buildings: A review of enabling technologies and applications," *Autom.*

- Constr.*, vol. 101, no. January, pp. 111–126, 2019, doi: 10.1016/j.autcon.2019.01.023.
- [245] Q. Lu *et al.*, “Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus,” *J. Manag. Eng.*, vol. 36, no. 3, p. 05020004, 2020, doi: 10.1061/(asce)me.1943-5479.0000763.
- [246] F. Khan, R. Lakshmana Kumar, S. Kadry, Y. Nam, and M. N. Meqdad, “Cyberphysical systems: A smart city perspective,” *Int. J. Electr. Comput. Eng.*, vol. 11, no. 4, pp. 3609–3616, 2021, doi: 10.11591/ijece.v11i4.pp3609-3616.
- [247] S. Kaewunruen and N. Xu, “Digital twin for sustainability evaluation of railway station buildings,” *Front. Built Environ.*, vol. 4, no. December, pp. 1–10, 2018, doi: 10.3389/fbuil.2018.00077.
- [248] S. M. E. Sepasgozar, F. K. P. Hui, S. Shirowzhan, M. Foroozanfar, L. Yang, and L. Aye, “Lean practices using building information modeling (Bim) and digital twinning for sustainable construction,” *Sustain.*, vol. 13, no. 1, pp. 1–22, 2021, doi: 10.3390/su13010161.
- [249] S. M. E. Sepasgozar, R. Karimi, S. Shirowzhan, M. Mojtahedi, S. Ebrahimzadeh, and D. McCarthy, *Delay causes and emerging digital tools: A novel model of delay analysis, including integrated project delivery and PMBOK*, vol. 9, no. 9, 2019. doi: 10.3390/buildings9090191.
- [250] N. A. Simchenko, S. Y. Tsohla, and P. P. Chuvatkin, “Effects of supply chain digital twins in the development of digital industry,” *Int. J. Supply Chain Manag.*, vol. 9, no. 3, pp. 799–805, 2020.
- [251] M. Tetik, A. Peltokorpi, O. Seppänen, and J. Holmström, “Direct digital construction: Technology-based operations management practice for continuous improvement of construction industry performance,” *Autom. Constr.*, vol. 107, no. July, p. 102910, 2019, doi: 10.1016/j.autcon.2019.102910.
- [252] P. E. D. Love and J. Matthews, “The ‘how’ of benefits management for digital technology: From engineering to asset management,” *Autom. Constr.*, vol. 107, no. July, 2019, doi: 10.1016/j.autcon.2019.102930.
- [253] J. Moubray, *Reliability Centered Maintenance*, Second Edi. 1997.
- [254] “The PRISMA2020 statement: An updated guideline for reporting systematic reviews.” [Online]. Available: <https://www.equator-network.org/reporting-guidelines/prisma/>
- [255] D.-G. Joe Opoku, S. Perera, R. Osei-Kyei, and M. Rashidi, “Digital twin application in the construction industry: A literature review,” *J. Build. Eng.*, vol. 40, no. April, p. 102726, 2021, doi: 10.1016/j.job.2021.102726.
- [256] G. B. Ozturk, “Digital Twin Research in the AECO-FM Industry,” *J. Build. Eng.*, vol. 40, no. April, p. 102730, 2021, doi: 10.1016/j.job.2021.102730.
- [257] F. Jiang, L. Ma, T. Broyd, and K. Chen, “Digital twin and its implementations in the civil engineering sector,” *Autom. Constr.*, vol. 130, no. July, p. 103838, 2021, doi: 10.1016/j.autcon.2021.103838.
- [258] and Sevda R. L. Milad Baghalzadeh Shishehgarckhaneh, Afram Keivani, Robert C. Moehler, Nasim Jelodari, “Internet of Things (IoT), Building Information Modeling (BIM), and Digital Twin (DT) in Construction Industry: A Review, Bibliometric, and Network Analysis,” 2022.
- [259] W. HU, K. Y. H. Lim, and Y. Cai, “Digital Twin and Industry 4.0 Enablers in Building and Construction: A Survey,” *Buildings*, pp. 1–27, 2022.
- [260] Q. He, G. Wang, L. Luo, Q. Shi, J. Xie, and X. Meng, “Mapping the managerial areas of Building Information Modeling (BIM) using scientometric analysis,” *Int. J. Proj. Manag.*, vol. 35, no. 4, pp. 670–685, 2017, doi: 10.1016/j.ijproman.2016.08.001.
- [261] D. W. Ying Ding, Ronald Rousseau, *Measuring Scholarly Impact*, 1st ed. Springer Cham, 2014. [Online]. Available: <https://doi.org/10.1007/978-3-319-10377-8>
- [262] J. van den B. Nees Jan van Eck, Ludo Waltman, Rommert Dekker, “A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS,” *J. Am. Soc. Inf. Sci. Technol.*, 2010, doi: 10.1002/asi.
- [263] K. Ishikawa, “Guide to Quality Control,” *Asian Product. Organ.*, 1976, [Online]. Available: ISBN 92-833-1036-5
- [264] M. Elnour and N. Meskin, “Novel Actuator Fault Diagnosis Framework for Multizone HVAC Systems Using 2-D Convolutional Neural Networks,” *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 3, pp. 1985–1996, 2022, doi: 10.1109/TASE.2021.3067866.
- [265] R. Langone, A. Cuzzocrea, and N. Skantzos, “Interpretable Anomaly Prediction: Predicting anomalous behavior in industry 4.0 settings via regularized logistic regression tools,” *Data Knowl. Eng.*, vol. 130, no. December 2018, p. 101850, 2020, doi: 10.1016/j.datak.2020.101850.
- [266] A. Rosato, F. Guarino, V. Filomena, S. Sibilio, and L. Maffei, “Experimental calibration and validation of a simulation model for fault detection of HVAC systems and application to a case study,” *Energies*, vol. 13, no. 15, 2020, doi: 10.3390/en13153948.

- [267] X. Zhu, Z. Du, X. Jin, and Z. Chen, "Fault diagnosis based operation risk evaluation for air conditioning systems in data centers," *Build. Environ.*, vol. 163, no. May, p. 106319, 2019, doi: 10.1016/j.buildenv.2019.106319.
- [268] E. Gulski, G. J. Anders, R. Jongen, and J. Parciak, "Damped AC Testing and Diagnosis of Wind Farm HVAC Long-Length Cable Circuits," *Energies*, vol. 15, no. 22, pp. 1–17, 2022, doi: 10.3390/en15228426.
- [269] Y. Chen, G. Lin, Z. Chen, J. Wen, and J. Granderson, "A simulation-based evaluation of fan coil unit fault effects," *Energy Build.*, vol. 263, p. 112041, 2022, doi: 10.1016/j.enbuild.2022.112041.
- [270] Z. Chen, X. Zhu, X. Jin, and Z. Du, "Machine learning enhanced inverse modeling method for variable speed air conditioning systems," *Int. J. Refrig.*, vol. 118, pp. 311–324, 2020, doi: 10.1016/j.ijrefrig.2020.06.020.
- [271] Y. Guo and H. Chen, "Fault diagnosis of VRF air-conditioning system based on improved Gaussian mixture model with PCA approach," *Int. J. Refrig.*, vol. 118, pp. 1–11, 2020, doi: 10.1016/j.ijrefrig.2020.06.009.
- [272] J. Zhang, C. Liu, and R. X. Gao, "Physics-guided Gaussian process for HVAC system performance prognosis," *Mech. Syst. Signal Process.*, vol. 179, no. May, p. 109336, 2022, doi: 10.1016/j.ymsp.2022.109336.
- [273] Y. Yan, J. Cai, T. Li, W. Zhang, and L. Sun, "Fault prognosis of HVAC air handling unit and its components using hidden-semi Markov model and statistical process control," *Energy Build.*, vol. 240, p. 110875, 2021, doi: 10.1016/j.enbuild.2021.110875.
- [274] Y. Yan, P. B. Luh, and K. R. Pattipati, "Fault Prognosis of Key Components in HVAC Air-Handling Systems at Component and System Levels," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 4, pp. 2145–2153, 2020, doi: 10.1109/TASE.2020.2979166.
- [275] D. T. De Silva, S. Setunge, and H. Tran, "Effect of Runtime on the Deterioration of HVAC Components in Building Services," *J. Infrastruct. Syst.*, vol. 28, no. 1, pp. 1–10, 2022, doi: 10.1061/(asce)is.1943-555x.0000656.
- [276] Y. Wu, C. T. Maravelias, M. J. Wenzel, M. N. ElBsat, and R. T. Turney, "Predictive maintenance scheduling optimization of building heating, ventilation, and air conditioning systems," *Energy Build.*, vol. 231, p. 110487, 2021, doi: 10.1016/j.enbuild.2020.110487.
- [277] G. Sánchez-Barroso and J. G. Sanz-Calcedo, "Application of predictive maintenance in hospital heating, ventilation and air conditioning facilities," *Emerg. Sci. J.*, vol. 3, no. 5, pp. 337–343, 2019, doi: 10.28991/esj-2019-01196.
- [278] D. Besiktepe, M. E. Ozbek, and R. A. Atadero, "Condition assessment framework for facility management based on fuzzy sets theory," *Buildings*, vol. 11, no. 4, 2021, doi: 10.3390/buildings11040156.
- [279] J. C. P. Cheng, W. Chen, K. Chen, and Q. Wang, "Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms," *Autom. Constr.*, vol. 112, no. August 2018, p. 103087, 2020, doi: 10.1016/j.autcon.2020.103087.
- [280] S. M. Aljadani, S. M. Almutairi, S. S. Ghaleb, and L. Al Khuzayem, "Elevint: A Cloud-based Internet of Elevators," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 1, pp. 505–513, 2022, doi: 10.14569/IJACSA.2022.0130163.
- [281] L. H. Li, C. Y. Lai, F. H. Kuo, and P. Y. Chai, "Predictive maintenance of vertical lift storage motor based on machine learning," *Int. J. Appl. Sci. Eng.*, vol. 16, no. 2, pp. 109–118, 2019, doi: 10.6703/IJASE.201909.
- [282] J. Gholami, A. Razavi, and R. Ghaffarpour, "Decision-making regarding the best maintenance strategy for electrical equipment of buildings based on fuzzy analytic hierarchy process; case study: elevator," *J. Qual. Maint. Eng.*, vol. 28, no. 3, pp. 653–668, 2022, doi: 10.1108/JQME-03-2020-0015.
- [283] C. T. A. Lai, W. Jiang, and P. R. Jackson, "Internet of Things enabling condition-based maintenance in elevators service," *J. Qual. Maint. Eng.*, vol. 25, no. 4, pp. 563–588, 2019, doi: 10.1108/JQME-06-2018-0049.
- [284] M. N. Razali, A. F. Jamaluddin, R. Abdul Jalil, and T. K. Nguyen, "Big data analytics for predictive maintenance in maintenance management," *Prop. Manag.*, vol. 38, no. 4, pp. 513–529, 2020, doi: 10.1108/PM-12-2019-0070.
- [285] D. Niu, L. Guo, W. Zhao, and H. Li, "Operation performance evaluation of elevators based on condition monitoring and combination weighting method," *Meas. J. Int. Meas. Confed.*, vol. 194, no. December 2021, p. 111091, 2022, doi: 10.1016/j.measurement.2022.111091.
- [286] B. C. Fialho *et al.*, "Development of a BIM and IoT-Based Smart Lighting Maintenance System Prototype for Universities' FM Sector," *Buildings*, vol. 12, no. 2, 2022, doi: 10.3390/buildings12020099.

- [287] P. Pałasz and R. Przynowa, "Using different ML algorithms and hyperparameter optimization to predict heat meters' failures," *Appl. Sci.*, vol. 9, no. 18, 2019, doi: 10.3390/app9183719.
- [288] Y. J. Chen, Y. S. Lai, and Y. H. Lin, "BIM-based augmented reality inspection and maintenance of fire safety equipment," *Autom. Constr.*, vol. 110, no. November 2019, p. 103041, 2020, doi: 10.1016/j.autcon.2019.103041.
- [289] G. H. Erharder and T. F. Hansen, "Towards optimized TBM cutter changing policies with reinforcement learning," *Geomech. und Tunnelbau*, vol. 15, no. 5, pp. 665–670, 2022, doi: 10.1002/geot.202200032.
- [290] A. Kumar Agrawal, S. Chattopadhyaya, V. M. S. R. Murthy, and A. Kumar Adesh, "Assessment of reliability and maintainability of earth pressure balance tunnel boring machine (EPBTBM) – An approach," *Tunn. Undergr. Sp. Technol.*, vol. 121, no. December 2021, 2022, doi: 10.1016/j.tust.2021.104337.
- [291] T. Moi, A. Cibicik, and T. Rølvåg, "Digital twin based condition monitoring of a knuckle boom crane: An experimental study," *Eng. Fail. Anal.*, vol. 112, no. October 2019, p. 104517, 2020, doi: 10.1016/j.engfailanal.2020.104517.
- [292] A. R. Sahu and S. K. Palei, "Fault analysis of dragline subsystem using Bayesian network model" *Reliab. Eng. Syst. Saf.*, vol. 225, no. May, p. 108579, 2022, doi: 10.1016/j.res.2022.108579.
- [293] Y. L. Tse, M. E. Cholette, and P. W. Tse, "A multi-sensor approach to remaining useful life estimation for a slurry pump," *Meas. J. Int. Meas. Confed.*, vol. 139, pp. 140–151, 2019, doi: 10.1016/j.measurement.2019.02.079.
- [294] Y. Zheng, W. Chen, Y. Zhang, and D. Bai, "Prediction of the Remaining Useful Life of a Switch Machine, Based on Multi-Source Data," *Sustain.*, vol. 14, no. 21, 2022, doi: 10.3390/su142114517.
- [295] E. De Santis, F. Arnò, and A. Rizzi, "Estimation of fault probability in medium voltage feeders through calibration techniques in classification models," *Soft Comput.*, vol. 26, no. 15, pp. 7175–7193, 2022, doi: 10.1007/s00500-022-07194-6.
- [296] Encalada-Dávila *et al.*, "Main Bearing Fault Prognosis in Wind Turbines based on Gated Recurrent Unit Neural Networks," *Mech. Syst. Signal Process.*, vol. 20, no. 20, pp. 419–424, 2022, doi: 10.1016/j.ymsp.2020.106830.
- [297] V. Wickramasinghe, W. P. S. Dias, H. Tran, and S. Setunge, "Using Snapshot Data of Deficiency and Generic Deterioration to Predict the Degradation of Building Elements," *J. Perform. Constr. Facil.*, vol. 36, no. 5, pp. 1–13, 2022, doi: 10.1061/(asce)cf.1943-5509.0001743.
- [298] G. Ma, M. Wu, J. Jia, and W. Yang, "Two-Level Quality Decision Support System for Building Structural Damage Prediction and Maintenance Solution Recommendation in the Operation and Maintenance Phase," *Journal of Construction Engineering and Management*, vol. 147, no. 6, 2021. doi: 10.1061/(asce)co.1943-7862.0002051.
- [299] J. C. Wang, D. P. Wang, W. H. Hou, and X. Niu, "Life Assessment of Railway Tunnel Lining Structure Based on Reliability Theory," *Teh. Vjesn.*, vol. 29, no. 6, pp. 1975–1982, 2022, doi: 10.17559/TV-20221014070502.
- [300] J. Souza, A. Silva, J. de Brito, and E. Bauer, "Service life prediction of ceramic tiling systems in Brasília-Brazil using the factormethod," *Constr. Build. Mater.*, vol. 192, pp. 38–49, 2018, doi: 10.1016/j.conbuildmat.2018.10.084.
- [301] L. B. Bartels, L. Y. Liu, K. El-Rayes, N. El-Gohary, M. Golparvar, and M. N. Grussing, "Work Optimization with Association Rule Mining of Accelerated Deterioration in Building Components," *J. Perform. Constr. Facil.*, vol. 34, no. 3, pp. 1–11, 2020, doi: 10.1061/(asce)cf.1943-5509.0001441.
- [302] J. González-Domínguez, G. Sánchez-Barroso, J. García-Sanz-Calcedo, and M. Sokol, "Condition-based maintenance of ceramic curved tiles roof in Primary Healthcare buildings using Markov chains," *J. Build. Eng.*, vol. 43, no. December 2020, 2021, doi: 10.1016/j.job.2021.102517.
- [303] T. H. Nguyen, D. T. Nguyen, D. H. Nguyen, and D. H. Tran, "Evaluation of Residual Strength of Corroded Reinforced Concrete Beams Using Machine Learning Models," *Arab. J. Sci. Eng.*, vol. 47, no. 8, pp. 9985–10002, 2022, doi: 10.1007/s13369-021-06493-8.
- [304] W. Wannenburg, H. M. Inglis, J. Wannenburg, and C. Roth, "Development of a remaining useful life (RUL) model for reinforced concrete beams subjected to high-cycle fatigue," *J. Qual. Maint. Eng.*, vol. 28, no. 3, pp. 607–633, 2022, doi: 10.1108/JQME-05-2020-0046.
- [305] N. A. Mohamed, A. M. Abdel-Alim, H. H. Ghith, and A. G. Sherif, "Assessment and Prediction Planning of R.C Structures Using BIM Technology," *J. Eng. Res.*, vol. 167, no. 2, pp. 48–56, 2020, doi: 10.21608/erj.2020.145845.
- [306] C. Ferreira, I. S. Dias, A. Silva, J. de Brito, and I. Flores-Colen, "Criteria for selection of cladding

- systems based on their maintainability,” *J. Build. Eng.*, vol. 39, no. December 2020, 2021, doi: 10.1016/j.jobe.2021.102260.
- [307] A. Farahani, H. Wallbaum, and J. O. Dalenbäck, “Optimized maintenance and renovation scheduling in multifamily buildings—a systematic approach based on condition state and life cycle cost of building components,” *Constr. Manag. Econ.*, vol. 37, no. 3, pp. 139–155, 2019, doi: 10.1080/01446193.2018.1512750.
- [308] G. Ruggiero, R. Marmo, and M. Nicoletta, “A methodological approach for assessing the safety of historic buildings’ Façades,” *Sustain.*, vol. 13, no. 5, pp. 1–17, 2021, doi: 10.3390/su13052812.
- [309] X. Yang, M. Kahouadji, O. Lakhali, and R. Merzouki, “Integrated design of an aerial soft-continuum manipulator for predictive maintenance,” *Front. Robot. AI*, vol. 9, no. September, pp. 1–17, 2022, doi: 10.3389/frobt.2022.980800.
- [310] Y. Tan, G. Li, R. Cai, J. Ma, and M. Wang, “Mapping and modelling defect data from UAV captured images to BIM for building external wall inspection,” *Autom. Constr.*, vol. 139, no. December 2021, p. 104284, 2022, doi: 10.1016/j.autcon.2022.104284.
- [311] H. Perez, J. H. M. Tah, and A. Mosavi, “Deep learning for detecting building defects using convolutional neural networks,” *Sensors (Switzerland)*, vol. 19, no. 16, 2019, doi: 10.3390/s19163556.
- [312] C. Ferreira, A. Silva, J. de Brito, I. S. Dias, and I. Flores-Colen, “Condition-based maintenance strategies to enhance the durability of etics,” *Sustain.*, vol. 13, no. 12, pp. 1–18, 2021, doi: 10.3390/su13126677.
- [313] I. S. Dias, A. Silva, C. O. Cruz, C. Ferreira, I. Flores-Colen, and J. de Brito, “Insurance Policies for Condition-Based Maintenance Plans of ETICS,” *Buildings*, vol. 12, no. 6, 2022, doi: 10.3390/buildings12060707.
- [314] A. Silva and J. de Brito, “Do we need a buildings’ inspection, diagnosis and service life prediction software?,” *J. Build. Eng.*, vol. 22, no. December 2018, pp. 335–348, 2019, doi: 10.1016/j.jobe.2018.12.019.
- [315] C. Oreto, L. Massotti, S. A. Biancardo, R. Veropalumbo, N. Viscione, and F. Russo, “Bim-based pavement management tool for scheduling urban road maintenance,” *Infrastructures*, vol. 6, no. 11, 2021, doi: 10.3390/infrastructures6110148.
- [316] S. Shrestha, S. W. Katicha, G. W. Flintsch, and B. K. Diefenderfer, “Pavement Deterioration Modeling and Network-Level Pavement Management Using Continuous Deflection Measurements,” *J. Infrastruct. Syst.*, vol. 27, no. 3, pp. 1–11, 2021, doi: 10.1061/(asce)is.1943-555x.0000626.
- [317] X. Luo, F. Wang, S. Bhandari, N. Wang, and X. Qiu, “Effectiveness evaluation and influencing factor analysis of pavement seal coat treatments using random forests,” *Constr. Build. Mater.*, vol. 282, p. 122688, 2021, doi: 10.1016/j.conbuildmat.2021.122688.
- [318] S. Amarasiri and B. Muhunthan, “Evaluating the Effectiveness of Pavement Preventive-Maintenance Treatments in Mitigating Longitudinal Cracks in Wet-Freeze Climatic Zones,” *J. Transp. Eng. Part B Pavements*, vol. 146, no. 2, p. 04020014, 2020, doi: 10.1061/jpeodx.0000158.
- [319] P. Babashamsi *et al.*, “Perspective of Life-Cycle Cost Analysis and Risk Assessment for Airport Pavement in Delaying Preventive Maintenance,” *Sustain.*, vol. 14, no. 5, pp. 1–14, 2022, doi: 10.3390/su14052905.
- [320] S. H. Khahro, Y. Javed, and Z. A. Memon, “Low cost road health monitoring system: A case of flexible pavements,” *Sustain.*, vol. 13, no. 18, pp. 1–14, 2021, doi: 10.3390/su131810272.
- [321] E. García-Macías, A. Kita, and F. Ubertini, “Synergistic application of operational modal analysis and ambient noise deconvolution interferometry for structural and damage identification in historic masonry structures: three case studies of Italian architectural heritage,” *Structural Health Monitoring*, vol. 19, no. 4, pp. 1250–1272, 2020, doi: 10.1177/1475921719881450.
- [322] E. García-Macías and F. Ubertini, “MOVA/MOSS: Two integrated software solutions for comprehensive Structural Health Monitoring of structures,” *Mech. Syst. Signal Process.*, vol. 143, pp. 1–26, 2020, doi: 10.1016/j.ymsp.2020.106830.
- [323] T. Liu, S. Zhang, and C. Wang, “A BIM-Based Safety Management Framework for Operation and Maintenance in Water Diversion Projects,” *Water Resour. Manag.*, vol. 35, no. 5, pp. 1619–1635, 2021, doi: 10.1007/s11269-021-02813-7.
- [324] X. Yin, A. Bouferguene, and M. Al-Hussein, “Data-Driven Sewer Pipe Data Random Generation and Validation,” *J. Constr. Eng. Manag.*, vol. 146, no. 12, pp. 1–14, 2020, doi: 10.1061/(asce)co.1943-7862.0001937.
- [325] Y. Xie, S. Li, T. Liu, and Y. Cai, “As-built BIM reconstruction of piping systems using PipeNet,” *Autom. Constr.*, vol. 147, no. December 2022, p. 104735, 2023, doi:

- 10.1016/j.autcon.2022.104735.
- [326] H. Mouradi, A. El Barkany, and A. El Biyaali, “Steel wire ropes failure analysis: Experimental study,” *Eng. Fail. Anal.*, vol. 91, no. April, pp. 234–242, 2018, doi: 10.1016/j.engfailanal.2018.04.019.
- [327] C. Wokem, T. Joseph, and M. Curley, “Fatigue life prediction for cables in cyclic tension,” *J. Strain Anal. Eng. Des.*, vol. 53, no. 3, pp. 141–155, 2018, doi: 10.1177/0309324717753211.
- [328] F. Ghofrani, H. Sun, and Q. He, “Analyzing Risk of Service Failures in Heavy Haul Rail Lines: A Hybrid Approach for Imbalanced Data,” *Risk Anal.*, vol. 42, no. 8, pp. 1852–1871, 2022, doi: 10.1111/risa.13694.
- [329] Z. Liu, Z. Zhang, G. Xu, W. Jin, and J. Lee, “Design of cyber-physical systems architecture for prognostics and health management of high-speed railway transportation systems,” *Int. J. Progn. Heal. Manag.*, vol. 9, no. 3, pp. 2–6, 2018, doi: 10.36001/IJPHM.2018.V9I3.2740.
- [330] D. Li, S. Kaewunruen, and R. You, “Time-dependent behaviours of railway prestressed concrete sleepers in a track system,” *Eng. Fail. Anal.*, vol. 127, no. February, p. 105500, 2021, doi: 10.1016/j.engfailanal.2021.105500.
- [331] X. Chen, G. Tian, S. Ding, J. Ahmed, and W. L. Woo, “Tomographic Reconstruction of Rolling Contact Fatigues in Rails Using 3D Eddy Current Pulsed Thermography,” *IEEE Sens. J.*, vol. 21, no. 17, pp. 18488–18496, 2021, doi: 10.1109/JSEN.2021.3086307.
- [332] M. Valinejadshoubi, O. Moselhi, and A. Bagchi, “Integrating BIM into sensor-based facilities management operations,” *J. Facil. Manag.*, vol. 20, no. 3, pp. 385–400, 2022, doi: 10.1108/JFM-08-2020-0055.
- [333] D. E. Ighravwe and S. A. Oke, “A multi-criteria decision-making framework for selecting a suitable maintenance strategy for public buildings using sustainability criteria,” *J. Build. Eng.*, vol. 24, no. April, p. 100753, 2019, doi: 10.1016/j.jobe.2019.100753.
- [334] J. Cárcel-Carrasco and J. A. Cárcel-Carrasco, “Analysis for the knowledge management application in maintenance engineering: Perception from maintenance technicians,” *Appl. Sci.*, vol. 11, no. 2, pp. 1–16, 2021, doi: 10.3390/app11020703.
- [335] M. Ghafoori and M. Abdallah, “Innovative Optimization Model for Planning Upgrade and Maintenance Interventions for Buildings,” *J. Perform. Constr. Facil.*, vol. 36, no. 6, 2022, doi: 10.1061/(asce)cf.1943-5509.0001757.
- [336] M. D’Orazio, E. Di Giuseppe, and G. Bernardini, “Occupant density impact on building maintenance: Data-driven approach for university buildings,” *Autom. Constr.*, vol. 141, no. June, p. 104451, 2022, doi: 10.1016/j.autcon.2022.104451.
- [337] S. Assaf and I. Srour, “Using a data driven neural network approach to forecast building occupant complaints,” *Build. Environ.*, vol. 200, no. March, p. 107972, 2021, doi: 10.1016/j.buildenv.2021.107972.
- [338] F. Faqih and T. Zayed, “Defect-based building condition assessment,” *Build. Environ.*, vol. 191, no. December 2020, p. 107575, 2021, doi: 10.1016/j.buildenv.2020.107575.
- [339] K. Kang, J. Lin, and J. Zhang, “BIM- and IoT-based monitoring framework for building performance management,” *J. Struct. Integr. Maint.*, vol. 3, no. 4, pp. 254–261, 2018, doi: 10.1080/24705314.2018.1536318.
- [340] H. Alavi, R. Bortolini, and N. Forcada, “BIM-based decision support for building condition assessment,” *Autom. Constr.*, vol. 135, no. December 2021, 2022, doi: 10.1016/j.autcon.2021.104117.
- [341] A. Nguyen, K. T. L. Kodikara, T. H. T. Chan, and D. P. Thambiratnam, “Deterioration assessment of buildings using an improved hybrid model updating approach and long-term health monitoring data,” *Struct. Heal. Monit.*, vol. 18, no. 1, pp. 5–19, 2019, doi: 10.1177/1475921718799984.
- [342] J. J. McArthur, N. Shahbazi, R. Fok, C. Raghubar, B. Bortoluzzi, and A. An, “Machine learning and BIM visualization for maintenance issue classification and enhanced data collection,” *Adv. Eng. Informatics*, vol. 38, no. October 2017, pp. 101–112, 2018, doi: 10.1016/j.aei.2018.06.007.
- [343] A. J. Prieto, G. Saelzer, D. Palma, and M. Riveros, “Digital Management System in Heritage Timber Buildings: A Detailed Application in Purránque, Chile,” *J. Perform. Constr. Facil.*, vol. 36, no. 2, pp. 1–16, 2022, doi: 10.1061/(asce)cf.1943-5509.0001710.
- [344] C. M. Prieto A.J., Silva A., Tori F., “Estimating Functional and Physical Service Life of Timber Buildings Concerning Thermal Performance Simulations Estimating Functional and Physical Service Life of Timber Buildings Concerning Thermal Performance Simulations,” pp. 1–21, 2022.
- [345] B. Narendra Kumar Rao, B. Bhaskar Kumar Rao, and N. P. Challa, “Predictive maintenance for monitoring heritage buildings and digitization of structural information,” *Int. J. Innov. Technol.*

- Explor. Eng.*, vol. 8, no. 8, pp. 1463–1468, 2019.
- [346] R. Marmo, M. Nicoletta, F. Polverino, and A. Tibaut, “A methodology for a performance information model to support facility management,” *Sustain.*, vol. 11, no. 24, pp. 36–37, 2019, doi: 10.3390/su11247007.
- [347] N. Wanigarathna, K. Jones, A. Bell, and G. Kapogiannis, “Building information modelling to support maintenance management of healthcare built assets,” *Facilities*, vol. 37, no. 7–8, pp. 415–434, 2019, doi: 10.1108/F-01-2018-0012.
- [348] C. Sassanelli *et al.*, “Industry 4.0 Driven Result-oriented PSS: A n Assessment in the Energy Management,” vol. 12, no. 4, pp. 186–203, 2022.
- [349] K. Liu and N. El-Gohary, “Bridge Deterioration Knowledge Ontology for Supporting Bridge Document Analytics,” *J. Constr. Eng. Manag.*, vol. 148, no. 6, pp. 1–14, 2022, doi: 10.1061/(asce)co.1943-7862.0002210.
- [350] R. Arismendi, A. Barros, and A. Grall, “Piecewise deterministic Markov process for condition-based maintenance models — Application to critical infrastructures with discrete-state deterioration,” *Reliability Engineering and System Safety*, vol. 212. 2021. doi: 10.1016/j.res.2021.107540.
- [351] Y. Gunawardena, F. Aslani, J. Li, and H. Hao, “In Situ Data Analysis for Condition Assessment of an Existing Prestressed Concrete Bridge,” *J. Aerosp. Eng.*, vol. 31, no. 6, pp. 1–13, 2018, doi: 10.1061/(asce)as.1943-5525.0000935.
- [352] D. C. Nguyen, T. Q. Nguyen, R. Jin, C. H. Jeon, and C. S. Shim, “BIM-based mixed-reality application for bridge inspection and maintenance,” *Constr. Innov.*, vol. 22, no. 3, pp. 487–503, 2022, doi: 10.1108/CI-04-2021-0069.
- [353] C. Wan *et al.*, “Development of a bridge management system based on the building information modeling technology,” *Sustain.*, vol. 11, no. 17, pp. 1–17, 2019, doi: 10.3390/su11174583.
- [354] V. Shahsavari, M. Mashayekhi, M. Mehrkash, and E. Santini-Bell, “Diagnostic testing of a vertical lift truss bridge for model verification and decision-making support,” *Front. Built Environ.*, vol. 5, no. July, pp. 1–19, 2019, doi: 10.3389/fbuil.2019.00092.
- [355] J. Lee, S. Jeong, H. Kim, K. C. Lee, and S. H. Sim, “Comparative study of long-term displacement measurement methods – Focusing on a Pre-stressed concrete bridge under construction,” *Meas. J. Int. Meas. Confed.*, vol. 201, no. July, p. 111691, 2022, doi: 10.1016/j.measurement.2022.111691.
- [356] C. Jofré-briceño, F. M. La Rivera, E. Atencio, and R. F. Herrera, “Implementation of facility management for port infrastructure through the use of uavs, photogrammetry and bim,” *Sensors*, vol. 21, no. 19, 2021, doi: 10.3390/s21196686.
- [357] A. Turnbull and J. Carroll, “Cost benefit of implementing advanced monitoring and predictive maintenance strategies for offshore wind farms,” *Energies*, vol. 14, no. 16, 2021, doi: 10.3390/en14164922.
- [358] J. V. Taboada, V. Diaz-Casas, and X. Yu, “CBM challenges and opportunities for O&M of the Johan Sverdrup Oil and Gas Field,” *J. Pet. Sci. Eng.*, vol. 205, 2021, doi: 10.1016/j.petro.2021.108890.
- [359] A. Gascón, R. Casas, D. Buldain, and Á. Marco, “Providing Fault Detection from Sensor Data in Complex Machines That Build the Smart City,” *Sensors*, vol. 22, no. 2, 2022, doi: 10.3390/s22020586.
- [360] Z. Ma, Y. Ren, X. Xiang, and Z. Turk, “Data-driven decision-making for equipment maintenance,” *Autom. Constr.*, vol. 112, no. January, p. 103103, 2020, doi: 10.1016/j.autcon.2020.103103.
- [361] Y. Liu, Z. Pang, M. Karlsson, and S. Gong, “Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control,” *Build. Environ.*, vol. 183, no. May, p. 107212, 2020, doi: 10.1016/j.buildenv.2020.107212.
- [362] E. A. Pam, D. Edwards, Z. Riaz, F. Mehmood, and J. Lai, “Engineering-out hazards: digitising the management working safety in confined spaces,” *Facilities*, vol. 37, no. 3–4, pp. 196–215, 2019, doi: 10.1108/F-03-2018-0039.
- [363] B. Liu, Y. Zhang, G. Zhang, and P. Zheng, “Edge-cloud orchestration driven industrial smart product-service systems solution design based on CPS and IIoT,” *Adv. Eng. Informatics*, vol. 42, no. August, p. 100984, 2019, doi: 10.1016/j.aei.2019.100984.
- [364] K. Yan, J. Huang, W. Shen, and Z. Ji, “Unsupervised learning for fault detection and diagnosis of air handling units,” *Energy Build.*, vol. 210, p. 109689, 2020, doi: 10.1016/j.enbuild.2019.109689.
- [365] J. Granderson, G. Lin, A. Harding, P. Im, and Y. Chen, “Building fault detection data to aid diagnostic algorithm creation and performance testing,” *Sci. Data*, vol. 7, no. 1, pp. 1–14, 2020,

- doi: 10.1038/s41597-020-0398-6.
- [366] B. Li, F. Cheng, H. Cai, X. Zhang, and W. Cai, “A semi-supervised approach to fault detection and diagnosis for building HVAC systems based on the modified generative adversarial network,” *Energy Build.*, vol. 246, p. 111044, 2021, doi: 10.1016/j.enbuild.2021.111044.
- [367] Y. Gamil, M. A. Abdullah, I. Abd Rahman, and M. M. Asad, “Internet of things in construction industry revolution 4.0: Recent trends and challenges in the Malaysian context,” *J. Eng. Des. Technol.*, vol. 18, no. 5, pp. 1091–1102, 2020, doi: 10.1108/JEDT-06-2019-0164.
- [368] J. Saini, M. Dutta, and M. M. N. Halgamuge, *Integrating IoT and AI for Indoor Air Quality Assessment*. 2022. [Online]. Available: https://doi.org/10.1007/978-3-030-96486-3_11
- [369] P. V Dudhe, N. V Kadam, R. M. Hushangabade, and M. S. Deshmukh, “Internet of Things (IOT): An Overview and its Applications,” *2017 Int. Conf. Energy, Commun. Data Anal. Soft Comput.*, pp. 2650–2653, 2017.
- [370] T. Liu, Y. Cai, J. Zheng, and N. M. Thalmann, “BEACon: a boundary embedded attentional convolution network for point cloud instance segmentation,” *Vis. Comput.*, 2021, doi: 10.1007/s00371-021-02112-7.
- [371] Y. Choi and S. Yoon, “Autoencoder-driven fault detection and diagnosis in building automation systems: Residual-based and latent space-based approaches,” *Build. Environ.*, vol. 203, no. May, p. 108066, 2021, doi: 10.1016/j.buildenv.2021.108066.
- [372] Kanchan, A. K. Gorai, and P. Goyal, “A review on air quality indexing system,” *Asian J. Atmos. Environ.*, vol. 9, no. 2, pp. 101–113, 2015, doi: 10.5572/ajae.2015.9.2.101.
- [373] USEPA, *Revised Air Quality Standards for Particle Pollution and Updates To the Air Quality Index (AQI)*. 2012, pp. 1–5. [Online]. Available: https://www.epa.gov/sites/default/files/2016-04/documents/2012_aqi_factsheet.pdf
- [374] E. de Rautlin de la Roy, T. Recht, A. Zemhari, P. Bourreau, and L. Mora, “Deep learning models for building window-openings detection in heating season,” *Build. Environ.*, vol. 231, no. January, p. 110019, 2023, doi: 10.1016/j.buildenv.2023.110019.
- [375] Z. Cheng, S. Wang, P. Zhang, S. Wang, X. Liu, and E. Zhu, “Improved autoencoder for unsupervised anomaly detection,” *Int. J. Intell. Syst.*, vol. 36, no. 12, pp. 7103–7125, 2021, doi: 10.1002/int.22582.
- [376] J. Xia, Y. Feng, C. Lu, C. Fei, and X. Xue, “LSTM-based multi-layer self-attention method for remaining useful life estimation of mechanical systems,” *Eng. Fail. Anal.*, vol. 125, no. March, p. 105385, 2021, doi: 10.1016/j.engfailanal.2021.105385.
- [377] B. Liu, C. H. Chen, P. Zheng, and G. Zhang, “An Adaptive Parallel Feature Learning and Hybrid Feature Fusion-Based Deep Learning Approach for Machining Condition Monitoring,” *IEEE Trans. Cybern.*, vol. PP, pp. 1–12, 2022, doi: 10.1109/TCYB.2022.3178116.
- [378] “Elevator BIM experts.” [Online]. Available: <http://www.elevatorbimexperts.com/elevator-drawings-bim-models-for-thyssenkrupp-cene-elevators/>
- [379] “What is BMS|Definition of Building Management Systemtlt.” [Online]. Available: <https://www.youtube.com/watch?app=desktop&v=wjWtYKpvp2k>