Digital Twin-enabled failure prediction for indoor air quality
A CNN-BiLSTM model with multi-head attention mechanism

Hu, Wei; Chang, Cheng; Wu, Han; Lai, Kang; Cai, Yiyu

DOI
10.1016/j.jobe.2025.114755

Publication date
2025

Document Version
Final published version

Published in
Journal of Building Engineering

Citation (APA)
Hu, W., Chang, C., Wu, H., Lai, K., & Cai, Y. (2025). Digital Twin-enabled failure prediction for indoor air quality: A CNN-BiLSTM model with multi-head attention mechanism. *Journal of Building Engineering*, *117*, Article 114755. https://doi.org/10.1016/j.jobe.2025.114755

Journal of Building Engineering 117 (2026) 114755

Contents lists available at ScienceDirect

Journal of Building Engineering

journal homepage: www.elsevier.com/locate/jobe

Digital Twin-enabled failure prediction for indoor air quality: A CNN-BiLSTM model with multi-head attention mechanism

Wei Hu [a,b], Cheng Chang [c], Han Wu [d,e], Kang Lai [b], Yiyu Cai [b,*]

[a] School of Artificial Intelligence (School of Future Technology), Nanjing University of Information Science and Technology, 210044, China
[b] School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798, Singapore
[c] Resource & Recycling, Department of Engineering Structures, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, the Netherlands
[d] School of Electronics and Control Engineering, Chang'an University, Xi'an, 710064, China
[e] School of Automation, Northwestern Polytechnical University, Xi'an, 710129, China

ARTICLE INFO

Keywords:
Predictive maintenance
Digital twin
Failure prediction
Indoor air quality
Neural network
LSTM

ABSTRACT

Accurate failure prediction is critical to achieving Predictive Maintenance (PdM) for Indoor Air Quality (IAQ), which is highly related to resident well-being and operational effectiveness. However, most existing studies emphasise anomaly detection rather than prediction. To develop a precise and robust method for pre-emptive IAQ warning, this article integrated Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory (BiLSTM), and Muti-Head Attention (MHA) mechanism into a novel C-B-M model, synergistically incorporating feature extraction, temporal dependency analysis, and contextual weighting mechanisms. Additionally, a real-world dataset collected from various buildings in Singapore is employed in a detailed comparative experiment with other benchmark models for different prediction periods, dataset selection, and failure severity levels to illustrate the effectiveness and robustness of the proposed method. Finally, a Digital Twin (DT)-oriented failure prediction framework for the indoor climate is introduced and validated through the prototype system demonstrating the 3D building model and IAQ alert information.

1. Introduction

Indoor air pollution is a global environmental concern, consistently associated with a significant health burden for residents. Research has indicated that prolonged exposure to ambient air pollutants increases the susceptibility to airborne diseases, including cardiovascular disease, respiratory ailments, and lung cancer [1]. Recently, the implementation of advanced maintenance solutions within the building industry has extended its scope from Facility Management (FM) to indoor climate control [2]. Intelligent monitoring and maintenance of Indoor Air Quality (IAQ) are highly dependent on the constant collection and communication capabilities of measurable IAQ indicators, which are poised to enhance performance through the integration with Digital Twin (DT) technology [3]. The real-time bidirectional information exchange enabled by DT enhances interactive capabilities and optimises planning scheduling through the incorporation of physical measurements, encompassing factors like temperature, humidity, and PM2.5, as well as decision

* Corresponding author.
E-mail addresses: huwei@nuist.edu.cn, huwe0015@e.ntu.edu.sg (W. Hu), c.chang-1@tudelft.nl (C. Chang), wuhan@mail.nwpu.edu.cn (H. Wu), kang049@e.ntu.edu.sg (K. Lai), myycai@ntu.edu.sg (Y. Cai).

https://doi.org/10.1016/j.jobe.2025.114755
Received 10 August 2025; Received in revised form 12 November 2025; Accepted 21 November 2025
Available online 24 November 2025

support feedback from the virtual realm, including features such as alarms, advisory inputs, and what-if analyses [4]. The development of such technologies promotes the paradigm transformation from reactive to Predictive Maintenance (PdM), signifying a more proactive approach facilitated by applying real-time data communication and Artificial Intelligence (AI) algorithms to detect and foresee failure events in equipment and environmental contexts [5].

As the mainstream maintenance approach, the preventive method follows periodic time-based or usage-based paradigms, exposing inherent drawbacks such as fixed schedule inefficiency, over-maintenance risk, and limited adaptability for dynamic conditions [6]. Consequently, the emergence of PdM provides more proactive and effective strategies. Fault detection and failure prediction are critical and fundamental problems during the PdM implementation. Prediction typically necessitates higher data analysis and algorithm development requirements than detection, resulting in limited industrial application and research in the field. Effective and acceptable prediction performance is essential to demonstrate the superiority of PdM over preventive approaches, specifically in terms of enhancing operational efficiency and reducing resource consumption. Additionally, the precision, interactivity, and adaptability offered by enabling technologies of DT-enabled PdM also contribute significantly to other advanced modelling-driven functions, such as pedestrian localisation [7], intelligent robot control [8,9], and indoor localisation [10], showcasing the broad applicability and transformative potential of this study across various domains.

For indoor climate maintenance, advancement has been achieved by prior studies in many aspects, including real-time monitoring, comfort improvement, and even energy management optimisation [11]. However, serval limitations still exist during the implementation of PdM solutions within the building industry. Firstly, current research concentrates on operation monitoring instead of failure prediction, requiring increased data exchange and algorithm requirements. Secondly, DT-enabled PdM methodology remains a largely unexplored area, promising significant improvements in operational performance and decision support process. Thirdly, the scarcity of real-world datasets impedes developing and evaluating advanced algorithms. To address the above gap, the author first introduced an ensemble Deep Learning (DL) model that incorporates a Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory (BiLSTM), and Muti-Head Attention (MHA) to develop an accurate and reliable IAQ warning system. Additionally, a unified IAQ-PdM framework is proposed from the DT perspective and verified by a prototype platform designed for visualization and interaction. Lastly, accuracy, precision, recall, and F1 scores are utilised as the evaluation metric to assess the model in a comparative experiment with other DL models based on a real-world dataset collected from Singapore, encompassing different scenario configurations of the predictive period, severity level, and dataset selection.

The subsequent sections of this paper are structured as follows: Section 2 provides an overview of existing studies within the field, and Section 3 presents the DT framework and elaborates on the system architecture. Section 4 introduces relevant neural network models and the proposed methodology. Section 5 illustrates the experiment setup and outlines the dataset information, while Section 6 discusses the experiment results and showcases the DT platform. Conclusions and limitations are discussed in Section 7.

2. Related Work

This section provides an overview of existing academic works that have contributed to the advancement of PdM implementation, including (1) Digital Twin in Facility Management, (2) Failure Detection and Prediction, and (3) Machine Learning (ML)-driven Approach.

2.1. Digital Twin in Facility Management

Integrating the PdM strategy in the FM field is increasingly recognised as a pivotal aspect for improving operational efficiency and minimizing downtime [12], with the DT approach greatly facilitating this progression. Serval researchers have contributed to the deployments and implementations of DT-enabled FM solutions.

2.1.1. Theoretical Research

Existing theoretical research establishes the foundation and offers the framework for integrating the DT paradigm into FM systems, primarily centred on literature analysis and conceptual frameworks. For literature analysis, a comprehensive examination of over seventy articles was conducted by Hosamo et al. [13] to investigate the advancements afforded by DT adoption, highlighting the significance of intelligent information management, Building Information Modelling (BIM) -FM interoperability, and information exchange. Additionally, Hodavand et al. [14] conducted a systematic review focusing on the integration of DT in FM, emphasising the optimisation of building lifecycle management and maintenance strategies. Piras et al. [15] contributed to this area by reviewing vital enabling technologies and digital tools critical to the successful deployment of DT. On the conceptual framework front, Zhao et al. [16] introduced a conceptual framework of DT-enabled FM to systematically integrate the industrial demands and technologies adoption, which includes components such as data stream processing and information model cooperation. Moreover, Yitmen et al. [17] developed an adapted model considering the building lifecycle management for incorporating the cognitive DT system, which has been examined across various life stages of building assets. Quirk et al. [18] employed BIM as the foundational platform for the DT-driven FM to investigate the implementation feasibility, covering deployment specifics regarding built environments, Mechanical, Electrical, and Plumbing (MEP) infrastructures, information hardware, etc.

Although preceding studies have established foundations and pathways for practical implementation, limitations and concerns are also identified. Notably, the lack of a unified interaction platform and data synchronisation system for multisource data and

stakeholder requirements is currently apparent, which is crucial for continuous real-time decision support services for complex buildings, potentially surpassing the capabilities of BIM [16]. Besides, exploring legal and financial dimensions throughout the building lifecycle is anticipated to be instrumental in shaping future research and industry implementation [17].

2.1.2. Practical studies

Expanding on theoretical foundations, researchers increasingly deployed real-world solutions to bridge the gap between academia and practical application, advancing the field's practicality in industry settings. Table 1 summarises the deployment of DT-enabled FM in related literature, highlighting the system functions and target assets. For independent buildings, university and residential buildings emerge as predominant venues for implementing DT-enabled FM, concentrating on managing indoor environments and energy consumption. Additionally, some studies have ventured into advanced areas such as evacuation guidance and failure detection. Complex structures, such as hospitals, airports, and communities, present multifaceted environments that necessitate beyond-basic functionalities. Alongside the functions mentioned above, sophisticated capabilities, including life cycle management, fire safety measures, and financial efficiency improvements, are developed to enhance asset performance. Furthermore, DT technology is notably advancing the inspection and maintenance processes within the transportation infrastructure.

Despite the significant progress in bridging theory and practice, several limitations and concerns have been highlighted in existing research. A notable deficiency in current DT solutions is the absence of autonomous control over malfunctioning equipment [19]. Besides, the financial risks of deploying DT solutions across diverse building environments warrant careful consideration [20]. Finally, incorporating maintenance schedules and budget constraints will significantly enhance the practical applicability of the DT system [21].

2.2. Failure Detection and Prediction

The detection and prediction of failure events are pivotal to the PdM strategies, which are instrumental in ensuring the reliability and efficiency of building facilities and environments [38]. This section categorises existing research into theoretical foundational research and practical case studies, providing an overview of conceptual advancements and industrial applicability.

2.2.1. Theoretical Research

Theoretical studies focus on analysing the existing studies and developing the underlying frameworks or models that support PdM. A detailed literature review on anomaly detection is provided by Nassif et al. [39], covering application scenarios, methodologies, and evaluation metrics. Himeur et al. [40] examined current AI-based frameworks for fault detection in building energy utilization. They presented a detailed classification system focused on the modules and criteria used in analysing these algorithms. For original framework research, Hosamo et al. [41] developed a DT framework for occupant comfort improvement using an ontology graphs-based method, integrating multisource data from BIM, sensor data, and resident feedback, while Roy et al. [42] developed a taxonomy combining anomaly and building measurement typologies, offering a systematic framework for anomaly detection methodologies. Furthermore, Chandra et al. [43] explored the principles of fault events and the statistical information for building facilities management, and Kayan et al. [44] refined the implementation of IoT and ML techniques for anomaly detection via the AnoML pipeline, underscoring the integration of AI methods, communication systems, and real-time deployment.

Table 1
Practical studies of the digital twin in facility management.

Building asset	Ref	Function
Independent building		
University building	[22]	Building evacuation guidance
	[23]	Indoor air quality and thermal comfort optimisation; Real-time indoor environment control
	[24]	Facility failure detection; Electricity & Water, & Gas consumption reduction
	[25]	Energy management;
Residential building	[26]	Thermal comfort optimisation; Energy management
	[27]	Energy consumption reduction based on activity recognition
	[28]	Real-time monitoring of thermal comfort
Heritage building	[29]	Long-term indoor environment monitoring; Preventive maintenance
	[30]	Relative humidity optimisation; Ventilation system real-time control
Public toilet	[20]	Public health assessment; Indoor air quality evaluation; Energy management; Carbon emission reduction
Building complex		
Hospital	[19]	Energy management; Facility faults reduction
Airport	[31]	Life cycle information demonstration; Digital management; Fire protection; Structure & MEP monitoring
Community	[32]	House stock energy management; Urban spatial distribution
	[33]	Energy management; Emissions reduction; Financial cost saving
	[34]	Regional energy management; Financial cost saving
	[35]	Predictive maintenance; Financial cost saving
Transportation infrastructure		
Tunnel	[36]	Operations management; Traffic planning; Accident-prevention; Emergency response
Bridge	[21]	Digital maintenance; Real-time inspection
	[37]	Vulnerability assessment; Risk-based maintenance planning

2.2.2. Practical Studies

Practical studies emphasise the application of PdM theories in real-world scenarios, demonstrating their practicality and effectiveness in maintaining operational standards for different building-related targets.

Building facilities have garnered considerable attention for their parallels with manufacturing industry equipment in maintenance and efficiency optimisation. Chen et al. [45] enhanced building facilities' administration through dependability assessment and scheduling improvements. Lu et al. [46] innovatively introduced an Industry Foundation Classes (IFC) - extension data structure into the facility failure detection process and developed a DT-enabled asset management system for daily operation management. Concurrently, Villa et al. [38] explored the synergistic use of IoT systems and BIM for effective fault detection and maintenance scheduling. Malki et al. [47] utilised an ML-based methodology to identify emerging patterns and trends to improve energy system maintenance. Besides, Santiago et al. [33] designed a significant data service architecture focused on analysing and predicting malfunctions in Heating, Ventilation, and Air Conditioning (HVAC) systems. Alongside advanced digital technologies, many researchers have employed DL methods to advance this field. Gálvez et al. [48] presented a hybrid methodology that combines model-based and data-driven strategies to address data scarcity in the PdM of HVAC systems. Additionally, Xia et al. [49] introduced a multi-layer self-aware model utilising LSTM networks to improve the accuracy and reliability of Remaining Useful Life estimations, contributing to cost savings and reducing unforeseen breakdowns.

The indoor environment represents a significant application scenario where failure detection and prediction are pivotal in enhancing safety, comfort, and energy efficiency. Li and Cai [50] developed a dynamic demand-responsive ventilation strategy to reduce the transmission of COVID-19 and improve energy efficiency, while Wei et al. [51] introduced an innovative detection system combining Autoencoders with LSTM networks to identify anomalies in indoor climates. Recognised as eco-friendly innovation equipment, Vertical Green Living Walls (VGW) have gained significant interest within the building industry. Liu et al. [52] explored ML-based techniques for monitoring the condition of VGW systems, thereby facilitating automated and intelligent PdM in indoor environments. Moreover, Jung et al. [53] developed an abnormal detection model for indoor environments based on temperature, humidity, and light intensity sensor systems to address the specific requirements of communal and open-plan offices.

2.3. Comparative Analysis of Existing Methods

The advancement of PdM in the construction industry is closely tied to the progress of modern information technologies. ML has greatly enhanced PdM efficiency in fault detection, prediction, and decision-support functions. As shown in Table 2, a comparative review of existing PdM methods illustrates the transition from traditional ML algorithms to advanced DL and DT–based frameworks, each with distinct advantages and constraints. Conventional methods such as Random Forests, Support Vector Machines, and Bayesian networks are robust and interpretable but insufficient for modelling nonlinear temporal and spatial dependencies in complex environments. DL models such as LSTM and Autoencoders enhance temporal learning and anomaly detection, yet their performance remains limited when addressing heterogeneous data and cross-system interactions. The proposed ensemble model integrates spatial, temporal, and attention-based learning to capture dynamic variations and interdependencies among indoor air-quality parameters,

Table 2
Comparative analysis of existing methods.

Ref.	Method	Key strengths	Reference limitations
[54]	Random Forest Classifier	Interpretable, suitable for tabular sensor data	Limited in capturing nonlinear temporal dynamics
[19]	Clustering, LSTM, Chinese NLP	Handles unstructured text with LSTM	Lacks multi-sensor fusion and spatial context
[49]	LSTM–based Multi-Layer Self-Attention	Captures long-term dependencies with attention	High computation cost and limited generalisation
[48]	SVM, Logistic Regression, Decision Trees, KNN,	Good baseline classification performance	Insufficient for complex and non-stationary patterns
[45]	Particle Filter, Evolutionary Algorithm	Robust to noise, suitable for stochastic estimation	Requires extensive parameter tuning
[46]	Bayesian Change Point Detection	Effective for change detection	Sensitive to parameter priors
[41]	Bayesian Network, ANN, SVM, Decision Tree	Integrates multiple ML paradigms	Limited scalability for real-time PdM
[51]	LSTM–AE	Learns nonlinear time-series representations	Overfitting risk under small datasets
[53]	LSTM	Sequential modelling capability	Unable to model spatial dependencies
[52]	LSTM–AE	Detects anomalies effectively	Requires careful threshold tuning
[44]	CNN, RNN, AE, SVM, Isolation Forest	Extracts local and sequential features	Fragmented architectures and limited interpretability
[42]	Supervised Learning Method	Baseline predictive framework	Limited adaptability to dynamic IAQ data
[27]	Hidden Markov Model, Naïve Bayes, SVM	Probabilistic modelling capability	Assumes stationary transition probabilities
[26]	KNN Regression, SVR, Decision Tree, Linear Regression	Simplicity, interpretability	Poor temporal feature learning
[47]	Prophet, LightGBM, VAR	Good for short-term trend prediction	Ignores deep temporal dependencies
[50]	Computer Vision Algorithm	Visual detection of equipment anomalies	Requires largely labelled datasets
[28]	Computer Vision Algorithm	Effective for visual defect recognition	Limited temporal forecasting capability
[22]	Bellman–Ford, Dual Sub-Gradient Algorithm	Optimisation for maintenance scheduling	Not designed for real-time PdM
Ours	CNN + BiLSTM + Multi-Head Attention	Integrates spatial, temporal, and contextual learning;	High computational demand

thereby providing a more comprehensive and adaptive PdM solution than existing single-network or conventional methods.

3. Digital Twin framework

This section details the proposed DT-enabled PdM framework for indoor climate, including (1) the Digital Twin Concept, (2) the Predictive Maintenance Concept, and (3) the Digital Twin-enabled Predictive Maintenance Framework.

3.1. Digital Twin concept

DT is regarded as an innovative methodology for generating digital replicas for physical targets, allowing for real-time bidirectional information exchange [55]. As shown in Fig. 1, the information from physical assets and operations are mapped to the virtual side, formulating informed decisions on the digital platform based on real-world situations.

3.2. Predictive Maintenance Concept

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 [56], while effective maintenance plays an important 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 has been outlined in Fig. 2. Currently, 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.

3.3. Digital Twin enabled Predictive Maintenance Framework

Current maintenance solutions for indoor climate primarily emphasise real-time monitoring and anomaly detection. However, predicting potential failure events is crucial in implementing PdM solutions, facilitating data-driven decision-making processes and improving residential satisfaction. As the representative enablers for Industry 4.0, the DT paradigm is a dynamic and virtual representation of physical systems, processes, or services, integrating real-time data, advanced simulation techniques, and predictive analytics to create a virtual mirror of the physical world [55]. It offers an efficient platform for AI engine integration and enhancing user engagement and solution delivery. As shown in Fig. 3 A layered failure prediction framework has been introduced to facilitate the adoption of DT-enabled solutions in indoor climate control.

The physical layer initiates with data collection and ends with the information transfer to the cloud database, encompassing the target environment and the devices used for data acquisition. Specific research objectives or practical task requirements guide the selection and placement of devices. IoT systems are predominantly utilised for data collection, serving as a cornerstone in this domain. Moreover, the landscape of information gathering is evolving with the introduction of advanced technologies such as infrared scanners and cameras [57]. These emerging tools are progressively supplementing traditional IoT systems, broadening the scope and efficiency of data acquisition. In this research, the building environment was meticulously scanned using a Faro scanner and a commercial IoT system was employed to gather IAQ data.

The virtual layer is designed to achieve accurate mapping and visualization of the components present in the physical layer. Also, this layer plays a crucial role in demonstrating the decision support information to stakeholders and operators in an understandable and interactive format. This study employed the IScan2BIM scheme [58] to generate a 3D visualization model of the building

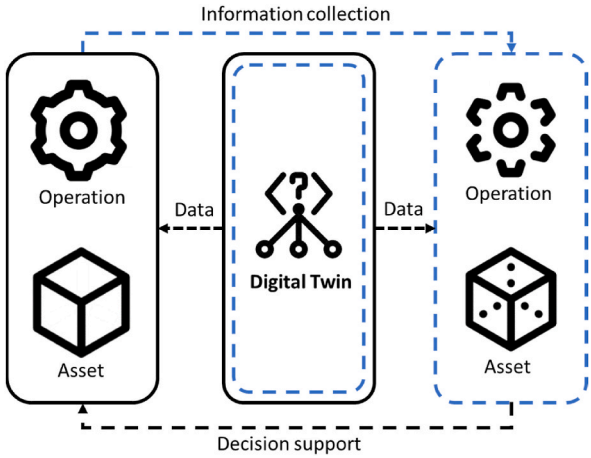

Fig. 1. The schematic Figure of digital twin.

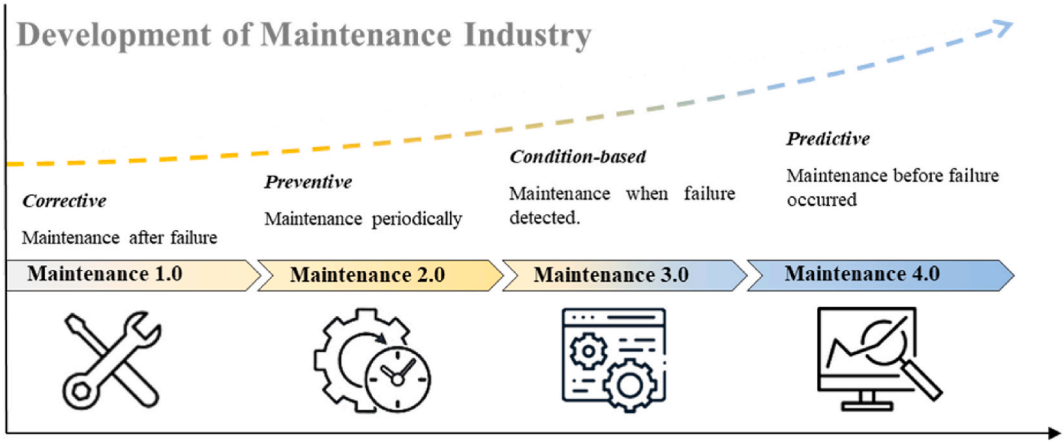

Fig. 2. The evolution of predictive maintenance.

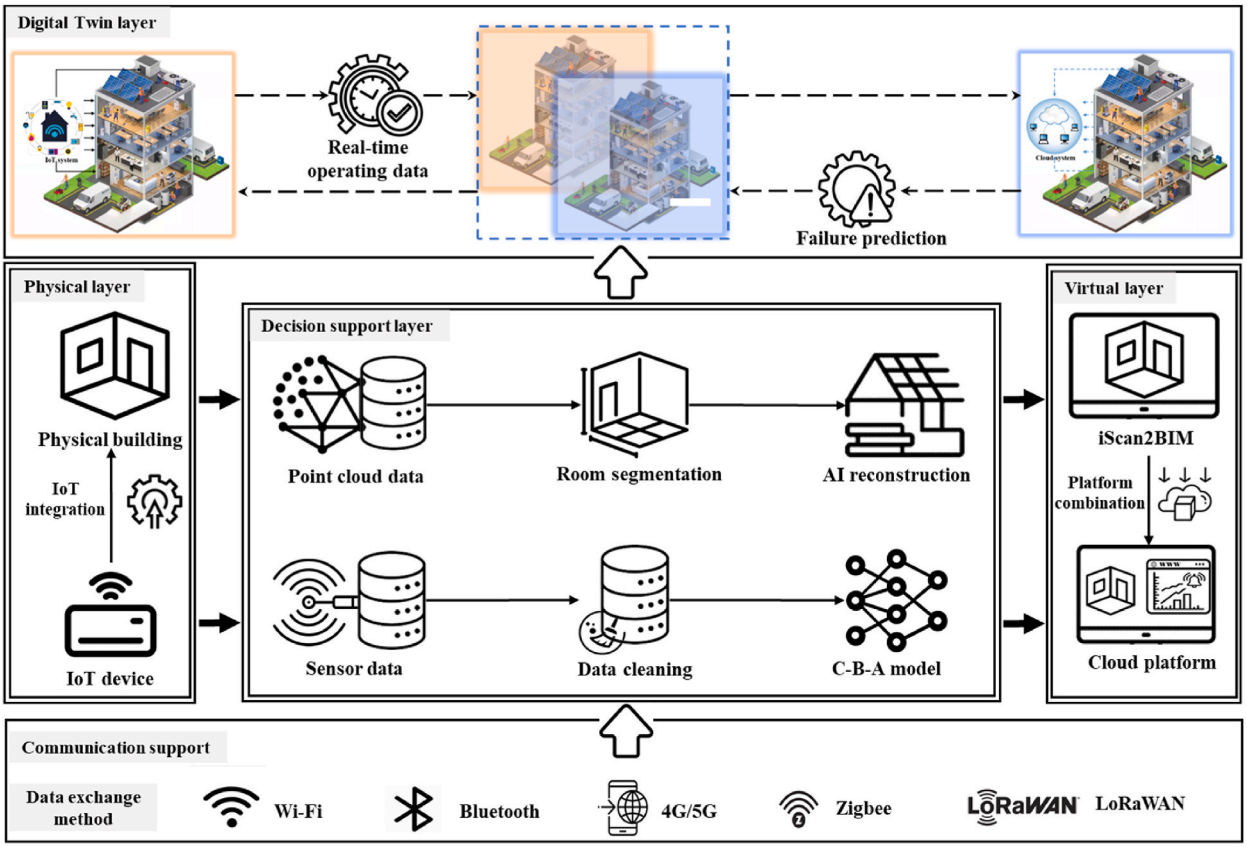

Fig. 3. Digital twin-enabled failure prediction framework for indoor air quality.

environment. Additionally, an indoor climate alarm system was developed to display the air quality index and provide alerts in case of air quality deterioration. This comprehensive approach enhances understanding environmental conditions and facilitates timely responses to potential issues.

The connection between the physical and virtual layers is realized through the communication support module. Data streams from the physical layer are transmitted to the cloud database using multiple communication protocols, including Wi-Fi, Bluetooth, Zigbee, and 4G/5G networks, via secure MQTT and RESTful protocols. The cloud serves as the data management hub, performing integration, transformation, and quality control before distributing the processed data to the virtual layer. Conversely, the virtual layer generates decision-support outputs, such as predicted IAQ indices, failure risks, and maintenance recommendations, which are transmitted back

to the physical layer through the same cloud interface. These outputs guide on-site control actions, such as triggering ventilation adjustments or alerting operators through the indoor climate alarm system. This closed feedback loop ensures synchronous updates and continuous interaction between the physical and virtual representations, forming the core mechanism of the DT-enabled PdM system and facilitating intelligent monitoring, early anomaly detection, and adaptive decision-making.

Subsequently, the decision support layer contains the reconstruction of a 3D model and the development of a failure prediction model. Given the volume and diversity of real-time, heterogeneous data collected, the initial step involves converting raw data into meaningful information for modelling and analysis. Specifically, the point cloud data are captured and generated through the Faro scanner, and the subsequent data processing involves removing noise and irrelevant points. At the same time, registration aligns multiple scans into a unified coordinate system [58]. Then, the suitable dataset is divided into segments and labelled according to real-world features, followed by the regeneration of the BIM model utilising the developed AI engine [59,60]. The processing procedure for sensor data is relatively straightforward, mainly involving cleaning and normalising, complemented by data splitting to prepare it for practical DL model training. Finally, the C-B-M are developed to achieve the failure prediction and generate the alarm information within the specific prediction periods.

The DT layer functions at the service and function level, leveraging real-time data from the physical side to offer capabilities such as condition monitoring and failure prediction [55]. Various services and functions can be integrated into the unified platform, supported by different decision-support models. This solution allows for continuous monitoring and analysis, leading to more informed and effective decision-making processes.

3.4. Scalability Analysis for Large-scale Building Complexes

The proposed framework provides the basic concepts and fundamental basis for PdM in general single-building systems, where data flow, system interactions, and management objectives remain relatively contained. However, extending its application to large-scale building complexes introduces additional analytical considerations arising from the substantial increase in data volume, subsystem heterogeneity, and the need for coordinated decision-making across multiple interconnected buildings. To explore the scalability of the proposed framework comprehensively, Table 3 summarises the key distinctions between single-building systems and large-scale building complexes in terms of data characteristics, system heterogeneity, and coordination requirements.

4. Deep Learning Methodology

This section provides detailed information on the related DL and proposed models.

4.1. Data Preprocess

The data preprocessing pipeline was developed to ensure the reliability, consistency, and readiness of the dataset for DL model training and validation. As illustrated in Fig. 4, the complete workflow comprises four stages: data cleaning, imputation, normalization, and data segmentation for temporal modelling.

(1) Data cleaning

Raw IAQ datasets were imported from four *csv* files. This step eliminates errors, inconsistencies, and irrelevant samples generated during data collection, improving the dataset's reliability. Missing or corrupted entries were identified using a null-value check, while non-essential attributes were excluded to maintain feature uniformity. Outliers and sensor malfunctions were identified through statistical inspection and removed accordingly. Moreover, feature correlation analysis was performed to assess redundancy between variables. For instance, the original dataset collected from the devices contains both $PM_{1.0}$ and $PM_{2.5}$, while $PM_{1.0}$ showed a strong correlation with $PM_{2.5}$, and therefore $PM_{1.0}$ was excluded to avoid multicollinearity and simplify the feature space.

Table 3
Scalability analysis for large-scale building complexes.

Aspect	General Building	Large-Scale Building	Framework Scalability
System Scale	Similar structure with limited IoT systems.	Multiple interconnected structures with diverse IoT systems.	Scalable DT hierarchy supporting federated coordination across multiple buildings.
Data Requirement	Moderate and structured data requirement	Massive and multi-source data requirement	Distributed data management with cloud–edge collaboration and real-time synchronization.
Communication Architecture	Local communication within building-level networks.	Distributed communication among multiple buildings and a central cloud.	Hybrid communication architecture enabling cross-building data exchange.
Maintenance Objective	Equipment- or floor-level service.	Cross-building coordination and system-level optimisation.	Multi-level maintenance strategy combining local and global decision-making for efficiency and resource sharing.
Decision Support Layer	Independent decision-support module for individual building.	Networked decision-support module requiring information fusion and coordination.	Federated decision-support layer and shared decision logic enabling complex-wide optimisation.

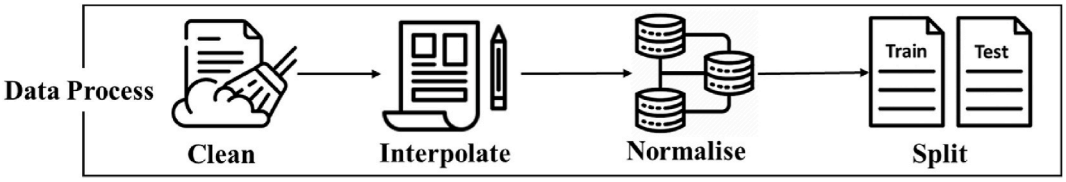

Fig. 4. Data processing pipeline.

(2) Data imputation

Missing values were addressed using an interpolation-based method to ensure temporal continuity and contextual accuracy. Specifically, each missing sample was replaced by the average of its two adjacent observations, as expressed in Eq. (1). This simple yet effective technique preserves the temporal dynamics of the IAQ dataset, minimizing bias introduced by data gaps.

$$y_m^i = \frac{y_m^{i-1} + y_m^{i+1}}{2} \tag{1}$$

(3) Normalization

To standardize feature magnitudes and enhance model convergence, Min-Max normalization was employed to scale all features to the range $[-1,1]$, as expressed in Eq. (2), ensuring equal contribution of each feature to the optimisation process. This normalization step is particularly crucial for neural networks using gradient-based optimisation, as it prevents variable dominance due to differing units or magnitudes.

Where y_m^i stands for the missing value, while $y_m^{i\pm1}$ represents the neighbours' value of the missing sample.

$$Scaled\ Value = \frac{Value - Min}{Max - Min} \times (New\ Max - New\ Min) + New\ Min \tag{2}$$

Where Value is the original value of the feature, Min and Max refer to the minimum and maximum value of the feature in the dataset. New Min and New Max are the target scope of the scaled range, while the range is set to [-1,1] in this study.

(4) Sequence generation and dataset partitioning

To facilitate temporal prediction, the normalized data were restructured into supervised sequences through a sliding-window mechanism. In this process, a fixed window of twelve consecutive observations was defined as the input sequence, and the subsequent time step was designated as the prediction target. This temporal segmentation was executed through the data split function, which systematically converts continuous time-series data into input–output pairs suitable for sequential learning. Following this transformation, the dataset was partitioned into training (80 %) and testing (20 %) subsets using a deterministic split that preserves chronological order, thereby preventing data leakage and ensuring fair evaluation of model generalization.

4.2. Convolutional Neural Network

CNN was initially proposed by Yann et al. [61], and its automatic and adaptive feature learning ability from the high-dimensional dataset laid the foundation for numerous advancements [62]. Concurrently, its effectiveness extends to time series-related problems

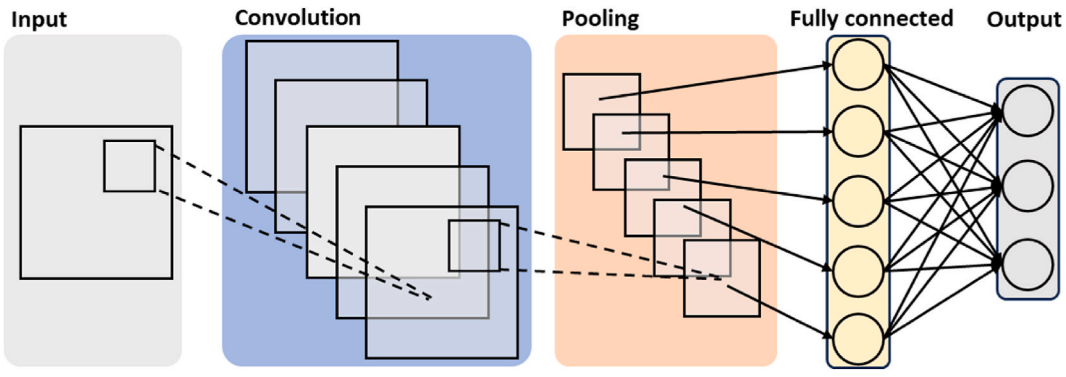

Fig. 5. Convolutional neural network structure.

attributed to its robust pattern and feature detection capabilities [63]. As shown in Fig. 5, it is a feedforward neural network composed of convolution, pooling, and full connection layers.

The CNN is critical in feature extraction primarily through its convolutional layer. For sequential inputs, the one-dimensional convolutional structure is applied for image processing instead of two dimensions (width and height) [64]. Generated features are processed through an activation function, such as Sigmoid, Tanh, and Relu. The convolution calculation is shown in Eq. (3).

$$f_t = \sigma\left(conv1D(x_t * W_t) + b_t\right) \tag{3}$$

Where σ is the activation function, *conv1D* conducts the 1D convolution, x_t is the input vector, W_t refers to the kernel weight, b_t stands for the kernel bias.

Following the convolution operation, the pooling layer is utilised to diminish the features' dimensionality and prevent over-fitting. Mainstream techniques in this domain include average and maximum pooling, which calculate the mean or maximum value of the feature points within a specified neighbourhood. The culmination of the abovementioned procedures results in extracting features with critical information pertinent to future predictive tasks, which is the notable benefit of CNNs for time series problems, enabling consistent feature detection over time and improving temporal data analysis.

4.3. Bidirectional Long Short-Term Memory

The BiLSTM is an advanced variation of the Recurrent Neural Network architecture [65]. Fig. 6 illustrates the structure of the BiLSTM, which integrates two LSTM networks for concurrent data processing in both forward and backward temporal directions. This bidirectional mechanism notably enhances the model's predictive and analytical capacities for sequential datasets, allowing it to capture patterns from past and future contexts [66]. Additional traversing of the input data significantly improves the ability to forecast time series [67], while the integration with the robust feature extraction capabilities of CNN also elevates the model's efficacy in recognition and prediction tasks [68]. This synergistic combination leverages the strengths of CNNs in spatial feature extraction and capitalises on the LSTM's proficiency in capturing temporal dynamics, thus demonstrating considerable potential in temporal-related tasks [69].

The LSTM unit is characterised by three fundamental gate structures: the input gate, the forget gate, and the output gate. Each of these gates serves a crucial function in controlling the information flow. The details are as follows.

4.3.1. Input Sequence

Two distinct LSTM networks process the input sequence. For a given sequence $X = \{x_1, x_2, x_3, \ldots x_T\}$, the sequence for Forward LSTM is from x_1 to x_T, while for Backward LSTM is from x_T to x_1.

where T is the length of the sequence and x_t is the input at time step t.

4.3.2. Details of LSTM unit

For each unit in forward and backward LSTM, the forget gate decides what portions of existing information should be discarded, while the input gate determines the newly added information for the cell state. Lastly, the output gate regulates the information that will be transferred to subsequent layers or emitted as the unit's output. Table 4 outlines the calculation equations and detailed descriptions for each step to elucidate these processes.

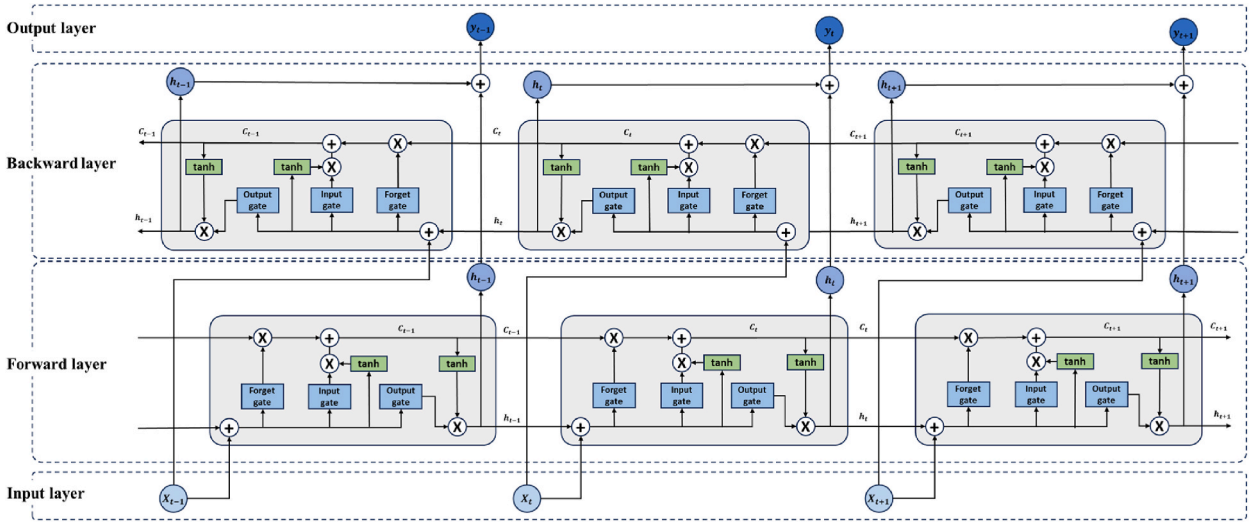

Fig. 6. Bidirectional long short-term memory schematic.

Table 4
Equations and Process of LSTM unit.

Gate	Equation	Process description
Forget Gate:	$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ (4)	Determine discarded information.
Input Gate:	$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ (5)	Decide added information.
	$\widetilde{C}_t = tanh\,(W_C \cdot [h_{t-1}, x_t] + b_C)$ (6)	Create the new candidate cell state.
	$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t$ (7)	Update the cell state.
Output Gate:	$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ (8)	Determine the output information.
	$h_t = o_t * \tanh\,(C_t)$ (9)	Decide the next hidden state.

where σ denotes the sigmoid function. * denotes Hadamard product. W and b are the weights and biases for different gates. h_{t-1} is the previous hidden state. x_t is the input at time step t. C_t is the cell state at time step t.

4.3.3. Bidirectional Integration

The final output for each time step in a BiLSTM is typically the concatenation of the forward and backward hidden states i.e., $\left[h_t^{forward}, h_t^{backward}\right]$, capturing information from both past and future contexts in the sequence. This dual perspective allows us to perform exceptionally well in tasks that require understanding the entire context of the input sequence. Fig. 6 shows the structure of BiLSTM.

4.4. Multi-head Attention

MHA was proposed by Vaswani et al. [70] and introduced as an innovative approach to enhance the model's capability for attending to various segments of the input sequence and extracting relevant features from multiple perspectives. Fig. 7 provides the schematic depiction of the MHA architecture.

4.4.1. Scaled dot-product Attention

The concepts of Query (Q), Key (K), and Value (V) are fundamental components enabling the model to focus on specific parts of the

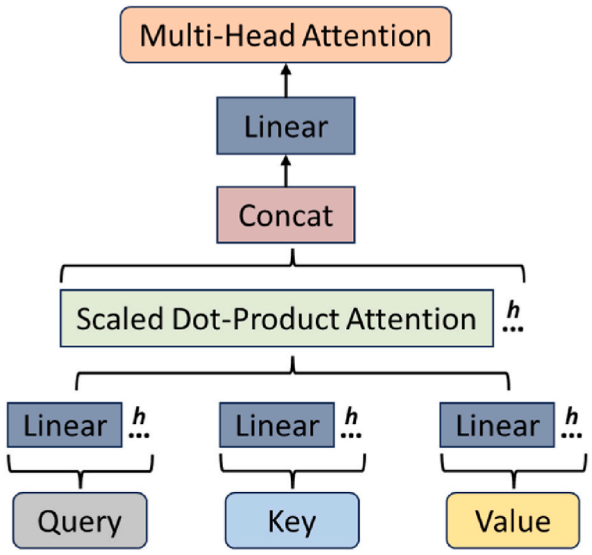

Fig. 7. Multi-head attention mechanism.

input data selectively. The Q vectors are designed to represent the current sequence of interest, serving as focal points for determining the relevance of past data points within the attention framework, while the K vectors correspond to these past data points, enabling the computation of attention scores that reflect their significance relative to the current time step. The V vectors are aligned with the actual data points or features at each timestep in the input sequence.

The core component of MHA is the scaled dot-product attention, which computes the attention scores based on the Q, K, and V vectors. The attention score for each element in a sequence is calculated by taking the dot product of the Q with all Ks, dividing each by $\sqrt{d_k}$, and applying a softmax function to obtain the weights on the values. The equation is given as:

$$Attention\ (Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{10}$$

where d_k is the dimension of the K vectors.

4.4.2. Multi-head Mechanism

The mechanism employs multiple attention heads instead of a single attention operation, enhancing its ability to discern various information aspects from the input sequence. Each attention head independently projects the Q, K, and V vectors into distinct representation subspaces through different and learnable linear projections, which allows the model to focus on different parts of the sequence and capture a diverse range of information. The equation for the MHA mechanism is:

$$MultiHead\ (Q, K, V) = Concat(head_1, \ldots, head_h)W^O \tag{11}$$

where $head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right)$ and W_i^Q, W_i^K, W_i^V, and W^O are parameter matrices.

4.5. The Proposed C-B-M model

This section introduces the details of the model principle and training workflow, comprehensively describing the theoretical foundation and the procedural steps.

4.5.1. The Framework of the Model Principle

The ensemble DL model has emerged as a pivotal strategy in complex data-driven tasks such as fault detection and failure

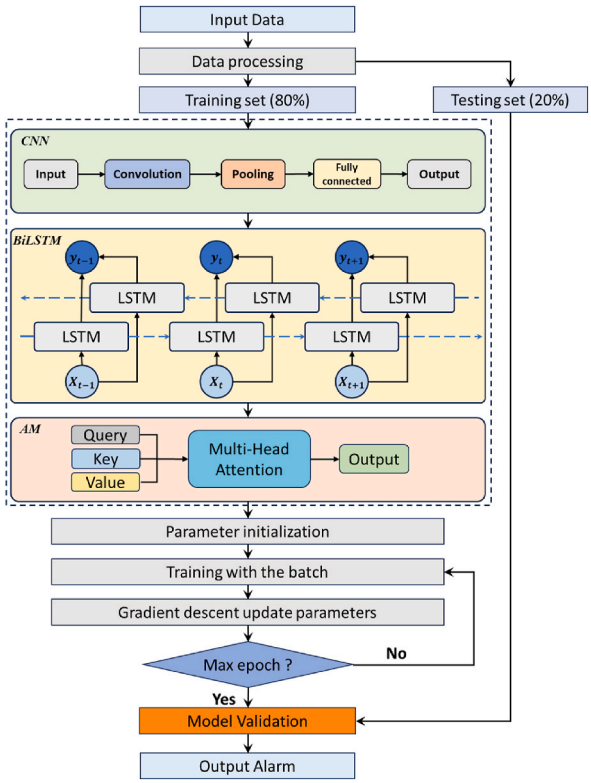

Fig. 8. The Flowchart of the proposed C-B-M model.

prediction [71]. Our proposed C-B-M model sequentially integrates CNN, BiLSTM, and MHA mechanisms to conduct the failure prediction for IAQ based on time-series datasets. The framework of the proposed model is presented in Fig. 8.

The model starts with the CNN component, where the temporal features are learned and extracted from high-dimensional data utilising the convolutional layers. This method is instrumental in identifying nuanced patterns and variations pertinent to environmental factors affecting air quality. After the feature extraction, BiLSTM networks are deployed to effectively comprehend long-term temporal dependencies within sequential data, analysing it from both forward and backward perspectives. The bidirectional framework facilitates a comprehensive analysis of time-series data, which is critical in air quality forecasting, where past and future trends significantly influence prediction accuracy. Finally, integrating the MHA mechanism marks a significant advancement in the model's design, offering improved interpretability of neural networks and focusing attention on the most relevant input segments for informed decision-making.

Overall, this integration leverages the strengths of each method: CNNs for their ability to extract spatial features from time-series data, BiLSTM for capturing long-term dependencies in time-series sequences, and MHA for focusing on the most relevant parts of the data, thereby enhancing the model's predictive performance.

4.5.2. Training Details

Table 5 meticulously outlines the training process for the C-B-M model to explain the model details further. The practical experiment begins with the definition of its components and goes through its initialisation, forward pass processing, and training loop.

The convolutional layer, BiLSTM layer, attention mechanism, and fully connected layer are defined first, and then the model is set up by initialisation. Input size, hidden channels, and kernel sizes are utilised for the convolutional layer, while the BiLSTM layer is optimised for both forward and backward sequence learning with tailored layer counts and hidden sizes. Besides, the Attention mechanism is designed to handle the bidirectional output by doubling its dimension. Finally, the setup is completed by incorporating a fully connected layer, which bridges the final output size.

The forward pass begins with the sequential application of convolutional layers coupled with ReLU activation functions for non-linearity. The output is then adjusted for LSTM compatibility, followed by the initialisation of hidden and cell states of the LSTM. After LSTM processing, the data is reshaped for the application attention mechanism, where the MHA mechanism refines the LSTM output by focusing on salient features. Finally, a residual connection combines the attention-enhanced output with the original LSTM output, transferring to the fully connected layer for final predictions.

The training loop encapsulates the process within an iterative epoch structure, where the training data is shuffled and batch-processed. Each batch undergoes a forward pass, loss computation against true labels, and an optimisation step via back-propagation, iterating until the model parameters are tuned to achieve an accurate prediction based on the given input. This

Table 5
Training process for the proposed C-B-M model.

Definition
- Convolutional layer - LSTM layer (Bidirectional) - Attention mechanism (MultiheadAttention) - Fully connected layer (Linear)
Initialisation:
- Initialize the convolutional layers based on the number of convolutional layers, input size, hidden channels, and kernel sizes. - Initialize a bidirectional LSTM layer with the number of layers and hidden size. - Initialize an attention mechanism with the LSTM output dimension doubled (to account for bidirectionality) and a single head. - Initialize a fully connected layer to produce the final output size from the LSTM output size.
Forward pass:
- Apply the convolutional layers sequentially with ReLU activation. - Permute the output to ensure compatibility with the LSTM input. - Initialize the hidden and cell states of the LSTM. - Pass the data through the LSTM. - Reshape the output for MultiheadAttention - Apply the attention mechanism on the LSTM output. - Add the attention output to the original LSTM (residual connection) output. - Apply the fully connected layer on the last output of the LSTM to get the final prediction.
Training loop:
- For each epoch: a. Shuffle the training data. b. Split the training data into batches. - For each batch: a. Perform a forward pass through the model. b. Compute the loss between the model output and the true labels. c. Perform backpropagation and an optimisation step.
Output the final model's predictions

comprehensive description illustrates the C-B-M model's design and operational flow, highlighting its integration of convolutional, recurrent, and attention-based mechanisms to enhance predictive performance on complex data sets.

5. Experiment

This section presents a detailed exposition of the target problem, dataset information, and experimental design.

5.1. Problem Formulation

Effective strategies for improving IAQ are essential in promoting a healthy and productive living and working environment, while poor air quality can exacerbate respiratory conditions like asthma and allergies and has been linked to decreased productivity and impaired cognitive function [11]. The United States Air Quality Index (USAQI) categorisation [72] is employed to define the failure events for indoor climate [73]. As detailed in Table 6, the failure event is identified when the measured IAQ deteriorates beyond the acceptable threshold set by the USAQI, indicating a potential health hazard for the building's occupants. During the experiment, events are defined as normal, failure with low severity, and failure with high-severity when the value of the Overall Index belongs to [0,100], [100, 200], and [200, 500], respectively. This systematic approach provides a precise and objective assessment of indoor air quality, facilitating prompt interventions to protect the health and comfort of individuals within these environments.

The Overall Index computation approach is a segmented linear function applied to pollutant concentrations. The index is computed for every environmental parameter, with the highest resulting index presented as the Overall Index. The below equation is utilised to convert the parameter concentration into index categories.

$$I = \frac{I_{high} - I_{low}}{C_{high} - C_{low}} (C - C_{low}) + I_{low} \tag{12}$$

where I represents the resulting index value, C is the pollutant concentration, C_{low} and C_{high} stand for the concentration breakpoint below and above C, respectively, I_{low} and I_{high} refer to the index breakpoint corresponding to C_{low} and C_{high}.

5.2. Dataset Acquisition

This section contains the sensor deployment and the dataset introduction.

5.2.1. Sensor Deployment

The experimental IAQ data were collected from different residential buildings at different locations in Singapore through commercial IoT devices. Fig. 9 showcases the device deployment during the data collection. Each site was continuously monitored between Nov 2021–Nov 2022, with devices configured at a 1-min logging interval. The Kaiterra Sensedge Mini devices were factory-calibrated prior to deployment, and additional data-level calibration procedures were implemented in this study, including exclusion of the initial 30-min warm-up period, removal of implausible outliers, and inter-sensor consistency validation across dwellings. The details of sensor specifications and deployment protocol are shown in Tables 7 and 8, respectively. These additions address device capability, placement, calibration, and data quality control, which are fundamental for ensuring the validity of PdM experiments.

5.2.2. Dataset Introduction

Table 9 provides the details of the experimental dataset, Fig. 10 illustrates the geographic locations of the monitored sites, while Fig. 11 presents the distribution of the recorded parameters. To preserve clarity of the figure scale, outliers of particulate matter, CO_2, humidity, and temperature were excluded from the boxplots, while the Overall IAQ Index is retained in its entirety. It should be noted that the IAQ index provided by the Kaiterra Sensedge Mini devices is computed internally from multiple pollutants ($PM_{2.5}$, CO_2, TVOCs, O_3, temperature, and humidity). Although the raw measurements of O_3 and TVOCs were not exported due to device output settings, they were included in the internal index calculation.

Four datasets are acquired from Singapore's east, west, south, and north regions to ensure a representation of environmental diversity, enabling a thorough evaluation of the algorithm's robustness. The analysis reveals consistent trends across different datasets. Firstly, the concentration of particulate matter and CO_2 predominantly exhibits low baseline levels with intermittent spikes, which is also discernible in the distribution of the Overall Index, reflecting the correlation among these variables. Additionally, humidity consistently fluctuates around its mean value, while temperature exhibits a steady upward trend over the observed period. Beyond common trends, variations in air quality across different regions contribute to the distinct characteristics observed in the datasets. For example, the Sengkang records the highest PM 2.5 level, while the Redhill is characterised by its highest humidity. Meanwhile, the Woodland stands out for the highest carbon dioxide level and temperature. The expansive temporal and spatial scope of the dataset provides a valuable resource for researching and investigating complex patterns and variations in indoor climate fields.

5.3. Experiment Design

The experiment details, including the experiment setup, experiment design, and evaluation metrics, are presented in this section.

Table 6
Breakpoints Table for overall index values [72].

Classification	*Index Category*	*Index Value*	$PM_{2.5}(\mu g/m^3)$	$CO_2(ppm)$	$TVOC(ppb)$	$TVOC\ (\mu g/m^3)$	$O_3(ppb)$	$O_3(\mu g/m^3)$
Normal	**Good**	**0–50**	**0–12**	**400–1000**	**0–220**	**0–1005.5**	**20–50**	**39.2–98.0**
	Moderate	**51–100**	**12.1–35.4**	**1001–1500**	**221–660**	**1005.6–3061.7**	**51–100**	**98.1–196.2**
Failure-low	**High**	**101–150**	**35.5–55.4**	**1501–2000**	**661–1430**	**3016.8–10056.1**	**101–165**	**196.3–323.7**
	Very High	**151–200**	**55.5–150.4**	**2001–2500**	**1431–2200**	**10056.2–15084.2**	**166–205**	**323.8–402.3**
Failure-high	**Very High**	**201–300**	**150.5–250.4**	**2501–5000**	**2201–5000**	**15084.3–20112.3**	**206–405**	**402.4–794.9**
	Very High	**300–500**	**250.5–500.4**	**5001–10000**	**5001–10000**	**20112.4–25140.5**	**406–605**	**795.0–1187.6**

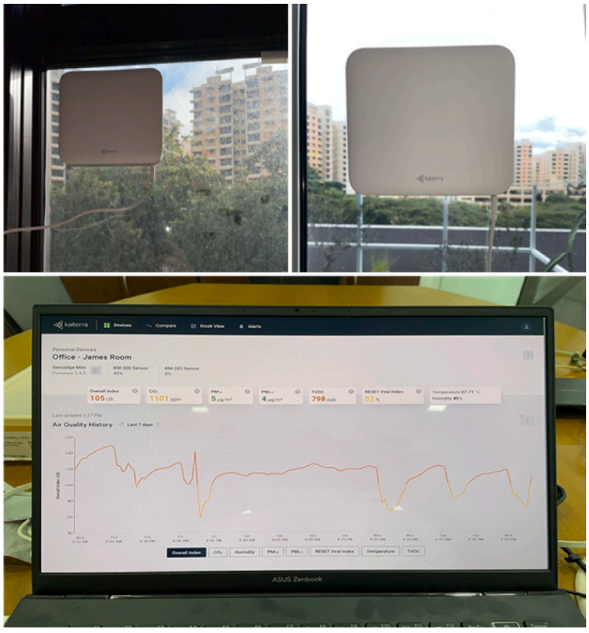

Fig. 9. Data collection & monitoring device.

Table 7
Specifications of kaiterra sensedege mini (SE-200/SE-200P).

Parameter	Sensor Technology	Range	Resolution	Accuracy
$PM_{2.5}/PM_{10}$	Laser particle sensor	0–1000 μg/m^3	1 μg/m^3	±3 μg/m^3 (0–30 μg/m^3); ±10 % ($PM_{2.5}$, 30–1000 μg/m^3); ±15 % (PM_{10}, 30–1000 μg/m^3)
TVOC	Multi-pixel MOx	0–60,000 ppb	1 ppb	±15 % ± 8 ppb
CO_2	Non-dispersive infrared	400–2000 ppm	1 ppm	±40 ppm ± 3 %
Ozone (O_3)	Electrochemical	20–2000 ppb	1 ppb	±10 %
Temperature	Digital sensor	−40 to 125 °C	0.01 °C	±0.3 °C
Relative Humidity	Digital sensor	0–100 % RH	0.01 % RH	±5 % RH

Resource: https://www.kaiterra.com/sensedge-mini-indoor-air-quality-monitor

Table 8
Sensor deployment protocol.

Aspect	Implementation situation	Rationale
Rooms type	Living room	Captures main occupant exposure zones
Mounting height	1.2 m above floor	Approximates human breathing zone
Ventilation proximity	Installed on closed windows	Provides consistent placement and avoids airflow disturbance
Sensor count	3 per dwelling	Provides adequate coverage across unit layouts
Warm-up exclusion	First 30 min of readings removed	Eliminates sensor stabilization artifacts
Data QA	>5 % error days removed ≤5 min gaps interpolated	Ensures data reliability and continuity

Table 9
Details of dataset.

Location:	Woodland	Redhill	Sengkang	Lakeside
Building type:	HDB-5 room	HDB-3 room	HDB 4-room	Condo 3-room
Samples number:	525600	525600	525600	525600
Collection period:	Nov 2021 to Nov 2022			
Time interval:	One sample per minute			
IoT device:	Kaiterra			
Parameters:	PM 2.5, CO_2, Humidity, Temperature, Overall Index			

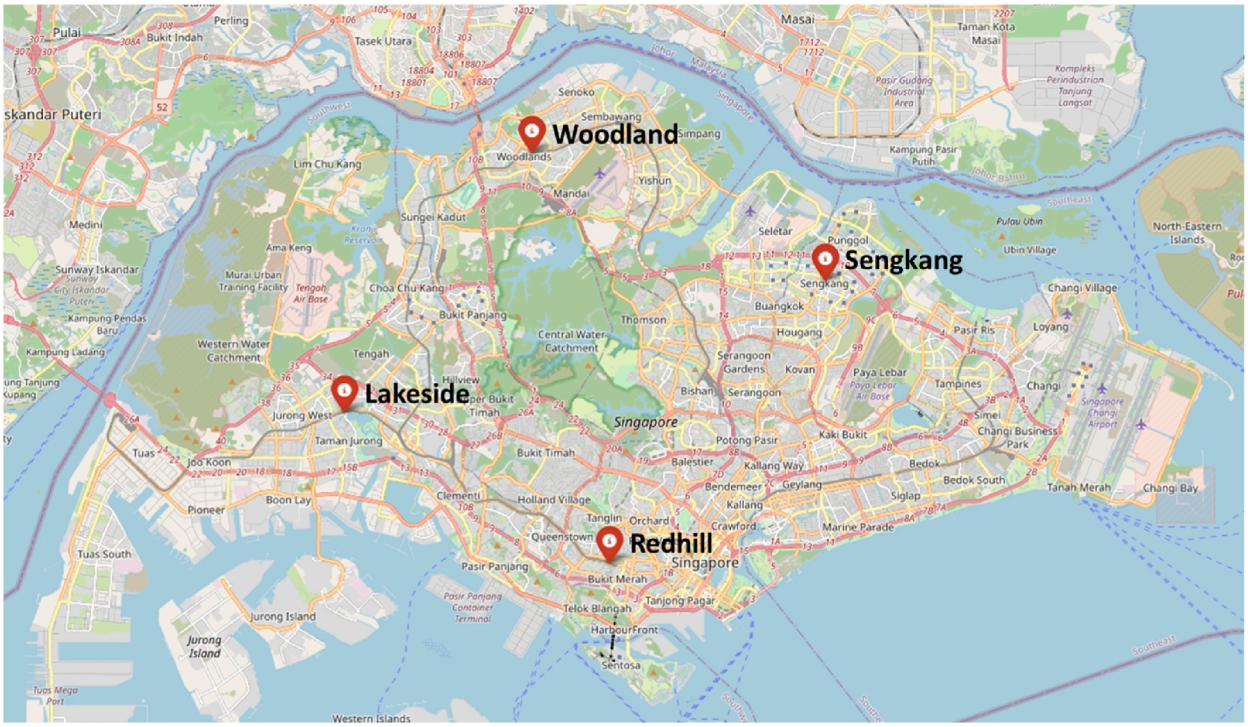

Fig. 10. The distribution of dataset location.

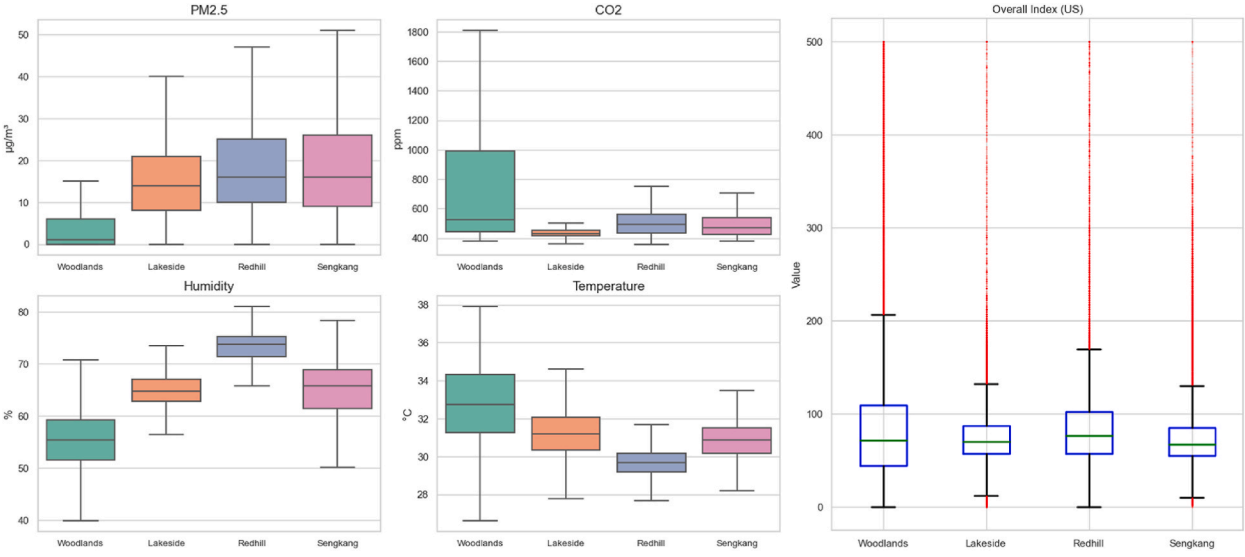

Fig. 11. Distribution Boxplot of parameter and overall index.

Table 10
System Configuration.

Unit	Description
Central Processing Unit	13the 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	Pytorch 2.0.1

5.3.1. Experiment Setup

The experimental environment, including hardware configuration and software framework, is shown in Table 10, all models were implemented in the PyTorch 2.0.1 framework and trained on a workstation equipped with an NVIDIA RTX 4070 GPU, Intel i9-13900HX CPU, and 32 GB RAM.

To obtain robust and well-justified hyperparameter values, a three-stage strategy was adopted. First, a coarse grid search was conducted to identify plausible ranges for kernel sizes, hidden dimensions, dropout rates, and attention heads. Second, Bayesian optimisation refined the candidate space to balance predictive performance and computational efficiency. Finally, a five-fold cross-validation procedure was employed to validate generalizability and avoid overfitting. The final configuration is outlined in Table 11, where each parameter is reported alongside a rationale, clarifying its functional role in the architecture. Benchmark models were tuned with the same search strategy, and common settings were aligned to ensure comparative fairness.

5.3.2. Design Guidelines

The experiment was conducted in adherence to the following guidelines to enhance the reliability of the results.

- *Basic performance analysis:* The initial effectiveness verification of the proposed method is demonstrated through the timeline figures, confusion matrix, and comparison analysis of training time and statistical index.
- *Comparative Study:* Several advanced DL models, including CNN, LSTM, and BiLSTM, are employed as benchmarks to validate the superior performance of the proposed method, ensuring a comprehensive evaluation and highlighting the strengths and potential limitations of the proposed model.
- *Robust Exploration:* The study employed varied prediction periods, severity levels, and dataset sections to evaluate the model's performance. The prediction period adjustment is achieved by setting the different lengths of the testing set with an unchangeable training dataset (4 months with 172,800 data points). Failure severity is categorised into low and high degrees, providing insights into the model's responsiveness for different situations. Data from Woodland, Redhill, Sengkang, and Lakeside are selected to comprehensively evaluate the proposed method across diverse scenarios.
- *Fairness Setting:* The experiments are conducted 10 times for each scenario, and the average of these performances is defined as the concluded results, which are intended to lessen the impact of various initialisations. Besides, the grid search methodology was employed during parameter tuning to guarantee an equitable and meaningful comparative analysis.

5.3.3. Evaluation Metrics

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are utilised to assess the original overall index predictions in fundamental performance analysis. MAE calculates the average absolute difference between the predicted values and the actual values, providing a straightforward interpretation of the prediction accuracy, while RMSE assigns higher weights to larger errors, making it particularly useful in situations where large errors are especially undesirable.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \widehat{y}_i| \tag{13}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2} \tag{14}$$

Where n is the number of observations, y_i are the actual value, and $\widehat{y}_i$ are the predicted value.

Condition labels are systematically assigned to each data sample based on corresponding values in the Overall Index, as detailed in Table 12. As illustrated in Table 13, 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.

Accuracy, precision, recall, and F1 score are commonly utilised to evaluate the performance of classification models in ML and statistics tasks. Accuracy quantifies the overall correctness of the classification, representing the proportion of accurate predictions (True failure and True normal) among the total number of predictions. Precision measures the proportion of true failure in the total

Table 11
Hyperparameter setting of the proposed model.

Hyperparameter	Value	Justification
Optimisation function	Adam	Offers adaptive learning rates and stable convergence.
Learning rate	0.0001	Ensures stable convergence with Adam optimizer.
Epochs	500	Providing sufficient training horizon, validated by loss monitoring.
Batch size	64	Balances GPU utilization with generalization performance.
CNN kernel sizes	[3, 5]	Captures both short- and mid-range temporal dependencies.
CNN stride	1	Preserves temporal granularity without information loss.
CNN hidden channels	32	Ensures adequate feature extraction capacity while controlling complexity
BiLSTM hidden dimension	64	Provides expressive capacity without overfitting risk.
BiLSTM layers	2	Enhances hierarchical temporal representation over single-layer structures.
Attention heads	4	Enables parallel focus on diverse temporal patterns with limited cost.
Dropout rate	0.3	Mitigates overfitting while maintaining representation power.

Table 12
Label assignment according to overall index.

Overall Index	Condition	Label
[0–100]	Normal	1
[100–200]	Failure (Low Severity)	2
[200–500]	Failure (High Severity)	3

Table 13
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.

failure predictions, highlighting the model's ability to trigger correct failure alarms. Recall, or sensitivity, demonstrates the proportion of true failure among all actual failures, emphasising the model's capacity to find all failure events. The F1-Score is a harmonic mean of precision and recall, providing a balance between these two metrics, particularly useful when the class distribution is uneven. During the practical implementation, appropriate indicators should be selected based on the specific conditions of the fault attributes. For example, prioritising recall over precision when missing an alarm is costlier than a false alarm. Therefore, the F1 score more effectively measures the algorithm's overall performance. The following equations illustrate how the above parameters are derived:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

$$F1 - Score = \frac{2{*}TP}{2{*}TP + FP + FN} \tag{17}$$

$$Recall = \frac{TP}{TP + FN} \tag{18}$$

F1 score also contributes to the uncertainty quantification for predictions, thereby offering valuable reliability insights for maintenance planning. The $F1 - Score$ formula can be converted into the following form (19), where β refers to the balance parameter and plays a crucial role in adjusting the emphasis between precision and recall. During practical deployment, (1-precision) is employed to quantify the rate of false alarms, which correspond to unnecessary maintenance actions, while (1 - recall) quantifies the rate of missed alarms, reflecting maintenance actions that were needed but not performed. This differentiation is critical for optimising maintenance strategies, allowing for the assessment of the trade-offs between the costs associated with unnecessary maintenance actions and the potential losses resulting from unaddressed maintenance needs.

$$F1 - Score = \left(1 + \beta^2\right) \frac{precision{*}recall}{\left(\beta^2{*}precision\right) + recall} \tag{19}$$

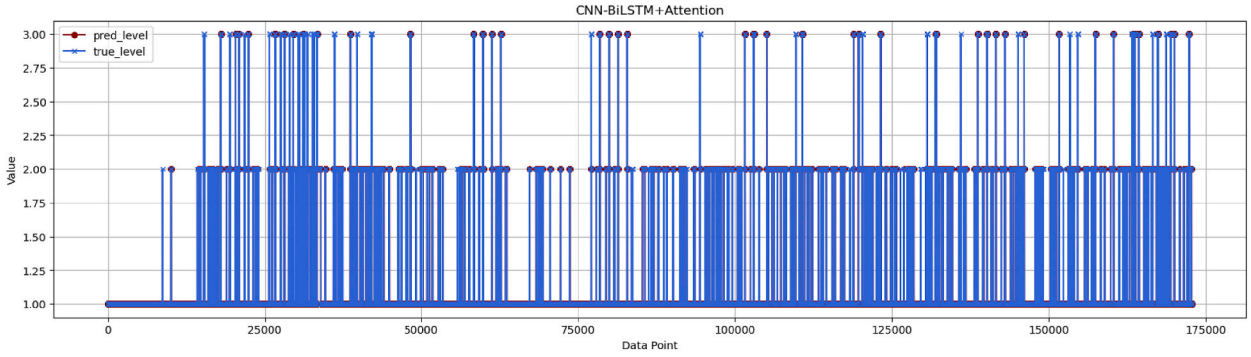

Fig. 12. Failure prediction results of our method (dataset = Woodland, prediction period = 120 days).

6. Result Discussion

This section presents the basic performance evaluation of the proposed method and conducts a comparative analysis with other benchmark models across various scenarios to further explore the interpretability and effectiveness of our method.

6.1. Basic Performance Analysis

This section presents the performance analysis of the proposed model, encompassing both Model Overall Performance and Robustness Evaluation under Sudden Pollutant Spikes.

6.1.1. Model Overall Performance

This section investigates the basic performance of the proposed method to provide an initial understanding of its efficacy. The result of the proposed method is demonstrated through a timeline figure and confusion matrix. Additionally, the comparison of training time and statistical index with other benchmark models is also discussed.

Fig. 12 displays the prediction outcomes of the approach for the 120-day test set (172800 samples) collected from Woodland, while Fig. 13 presents the corresponding confusion matrix. Values 1, 2, and 3 denote normal conditions, low-severity failures, and high-severity failures, respectively. The analysis reveals that the algorithm achieves commendable accuracy across all three conditions. However, a notable limitation is the misclassification of many high-severity failures as low-severity, as shown in Fig. 13. This result primarily arises from two factors. Firstly, insufficient examples of high-severity failures hinder the algorithm's ability to identify such conditions accurately. Secondly, some high-severity conditions are not pronounced enough to distinguish them from low-severity conditions definitively. Addressing these challenges presents a critical avenue for future research to enhance the algorithm's precision in classifying failure severities.

6.1.2. Robustness Evaluation under Sudden Pollutant Spikes

(1) **Identification of Sudden Pollutant Spikes**

To assess the model's ability to handle abrupt environmental variations, the high-severity pollutant condition was defined as an extreme event corresponding to the red threshold line in Fig. 14 When the Air Quality Index (AQI) exceeded 200 US units, the system was considered to experience a sudden pollutant spike. These instances occur intermittently throughout the dataset, forming a sparse distribution of high-severity episodes amidst predominantly low-level conditions. Such peaks represent abrupt deteriorations of air quality caused by transient pollution bursts, ventilation inefficiency, or sensor fluctuations.

(2) **Model Prediction Behavior under Extreme Conditions**

As illustrated in Fig. 14, the proposed framework effectively tracks abrupt pollutant spikes, with the predicted outputs closely aligning with the ground truth during high-severity episodes. The integration of temporal convolution and bidirectional memory

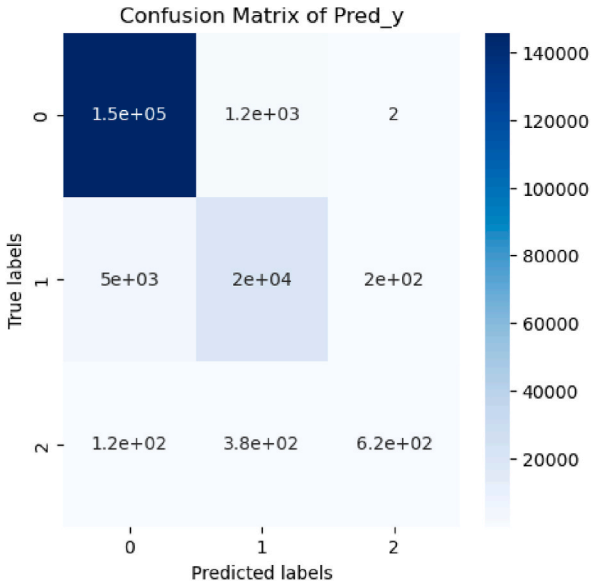

Fig. 13. Confusion matrix (dataset = Woodland, prediction period = 120 days).

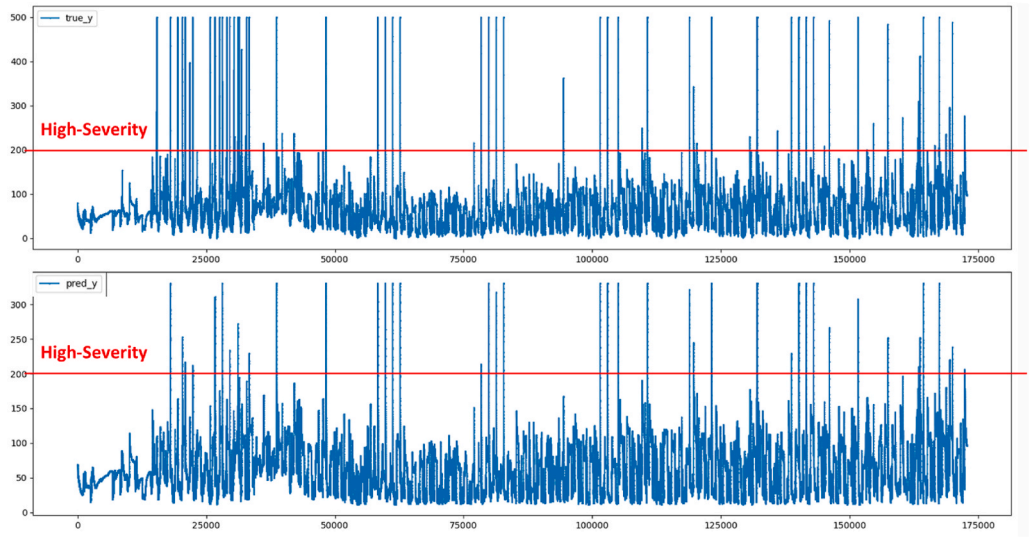

Fig. 14. The distribution of extreme events (observation-up, prediction-Down).

enhances sensitivity to rapid fluctuations, enabling the model to respond promptly to sudden variations. Quantitatively, the confusion matrix in Fig. 13 substantiates the model's robustness under extreme conditions. Despite a pronounced class imbalance, the framework accurately recognizes high-severity failures, maintaining a strong diagonal pattern across all categories. Misclassifications between moderate and high-severity states remain limited, underscoring the network's resilience to noise and its capacity to generalize from infrequent extreme samples. Together, these findings verify that the proposed network exhibits robust performance under sudden pollutant spikes.

6.1.3. Basic Comparison Analysis

To evaluate the model comprehensively, Table 14 compares the training time and computational efficiency between the proposed model and other benchmarks. The "Wall Time" denotes the total elapsed duration, while the "CPU Times Total" reflects the active processing period. In addition, the "Average CPU Load" and "Peak GPU Memory" provide insights into hardware utilization. It is observed that ARIMA achieves the shortest training time due to its simple statistical formulation. SVR and Prophet show longer durations and higher CPU load, indicating greater computational demand. Among deep models, CNN records the best efficiency with moderate CPU and GPU usage, benefiting from its lightweight structure. LSTM and BiLSTM require longer training and higher resource consumption because of their sequential dependencies. Despite having the most complex structure among all models, the proposed method benefits from the feature extraction process, reducing the overall training time and improving the performance. The proposed model achieves a balanced trade-off between training time and resource cost, demonstrating efficient integration of spatial and temporal learning.

As mentioned in Section 5.3.3, condition labels are assigned based on the predicted Overall Index during the practical implementation. Then, the performance is evaluated using accuracy, recall, precision, and the F1-score due to their appropriateness for the classification problem. To further examine the model's performance, Table 15 shows that the proposed method recorded the lowest MAE and the second lowest RMSE for original Overall Index predictions, proving its superior accuracy.

6.2. Comparison Study with Benchmark Models under Various Scenario Settings

This section systematically analyses the proposed method's performance in different failure severity and prediction period settings. The experiments categorise indoor air conditions into three types: normal, low-severity failure, and high-severity failure. Moreover,

Table 14
Training time comparison (dataset = Woodland, prediction period = 120 days).

Model	Wall Time	CPU Times Total	Avg. CPU Load (%)	Peak GPU Memory (MB)
ARIMA	22.4 s	1 min 57 s	42 %	–
SVR	41.8 s	3 min 36 s	61 %	–
Prophet	55.3 s	5 min 12 s	69 %	–
CNN	30.5 s	3min 13s	38 %	1380 MB
LSTM	37 s	4min 4s	44 %	1610 MB
BiLSTM	48.1 s	6min 37s	51 %	1930 MB
CNN-BiLSTM-Attention	40.6 s	4min 27s	49 %	2320 MB

Table 15
MAE & RMSE performance comparison (dataset = Woodland, prediction period = 120 days).

Model	Overall index	
	MAE	RMSE
ARIMA	8.63	27.05
SVR	7.84	25.42
Prophet	7.26	23.96
CNN	6.65	23.14
LSTM	7.14	21.25
BiLSTM	6.17	18.14
CNN-BiLSTM-Attention	5.99	19.21

with the determined training set, various prediction periods ranging from 30 (43,200 samples) to 180 days (259,200 samples) are considered to examine the multi-period prediction ability of the proposed model. The Woodland dataset is utilised in this experiment. The experiment results are depicted in Figs. 15–17, providing a detailed analysis of the performance of the proposed and benchmark models. This comparison aims to elucidate the model's strengths and weaknesses, thereby contributing to a deeper understanding of its efficacy within existing frameworks. Accuracy, Precision, Recall, and F1-Score are utilised as evaluation metrics to provide a multi-faceted view of the model's effectiveness, addressing its precision in prediction, ability to identify relevant instances correctly, and overall reliability.

According to the results, the proposed model consistently achieves the highest performance across most evaluation metrics and scenarios. Although ARIMA exhibits rapid convergence in short-term forecasts, its performance degrades sharply as the prediction horizon extends, indicating limited adaptability to nonlinear and multivariate patterns. Both SVR and Prophet demonstrate moderate accuracy in shorter horizons but fail to sustain stability beyond 90 days, mainly due to their dependence on smooth trend assumptions. Among DL models, CNN maintains strong short-term precision, while LSTM and BiLSTM show sensitivity to extended temporal dependencies, leading to gradual declines in recall and F1-score as the prediction period increases. In terms of recall, the performance of our model varies across scenarios, exhibiting a clear advantage only in low-severity failure situations. As a comprehensive measure of model effectiveness, the F1-score validates the efficacy of the proposed method, which excels in almost all testing scenarios.

The proposed algorithm demonstrates considerable advantages across all prediction periods and event types, except for individual high-severity scenarios where the CNN and BiLSTM model exhibits marginally better performance. Addressing the data scarcity issue in high-severity cases is expected to further enhance the benefits of the proposed method on the model's ability to handle high-severity scenarios and its inherent advantages.

Furthermore, our observations indicate that LSTM and BiLSTM methods experience a decline in performance as the prediction period extends, a trend not observed in either the CNN or proposed model. This observation suggests that the CNN component more effectively integrates temporal features, enabling the model to maintain stability over extended prediction periods. The exceptional resilience and accuracy of the proposed model can be attributed to its ensemble approach, which leverages the LSTM's accuracy in time-series prediction and enhances algorithmic stability through the integration of CNN with the MHA mechanism.

6.3. Comparison Study with Benchmark Models under Different Datasets Selection

This section investigates the applicability and efficacy of the proposed method across different air quality scenarios, and the F1 score is employed for the model evaluation index because of its holistic measure of precision and recall. Three different datasets gathered from the eastern, western, and southern regions of Singapore are utilised to provide a comprehensive representation of environmental diversity and verify the model's robustness. Additionally, such research extends its implications to cost-effective regional energy optimisation because the adaptability of strategies is critical for the different datasets with unique characteristics [32].

Fig. 18 presents a detailed comparison of the performance of various DL models and traditional models across different datasets, underscoring the adaptability and superior performance of the proposed method across different datasets. The results demonstrate that the proposed model consistently achieves the highest F1-scores across most conditions, maintaining stability and superior performance throughout all datasets. While traditional approaches perform reasonably well in short-term predictions, their performance declines rapidly as the prediction horizon extends, indicating limitations in capturing nonlinear and multivariate temporal dependencies. For normal and low-severity failure conditions, the proposed model exhibits a clear and consistent advantage over all comparison models, reflecting its ability to generalize effectively across datasets. Under high-severity failure scenarios, the proposed method continues to outperform other algorithms and even expands its performance margin with increasing prediction periods, demonstrating improved adaptability to complex and volatile environments. In contrast, other methods show a decreasing F1-trend as the prediction period lengthens.

Overall, the findings confirm the robustness, adaptability, and long-term stability of the proposed model across different datasets and failure severities. These results highlight its strong potential for scalable deployment in real-world indoor environmental prediction and energy optimisation tasks.

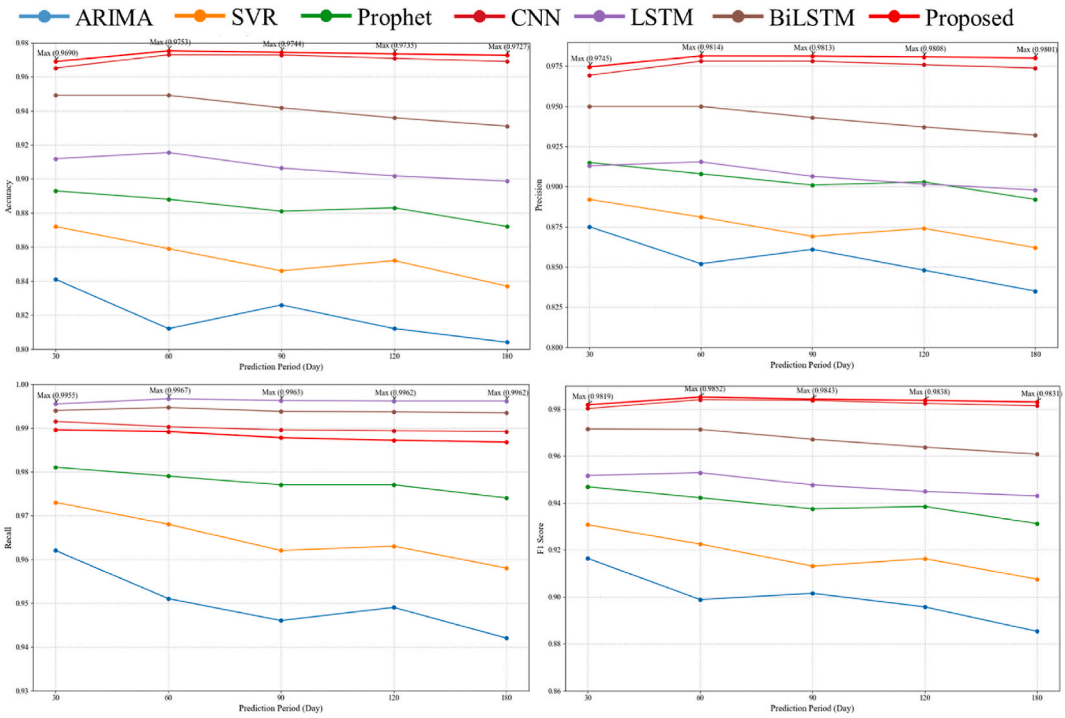

Fig. 15. Comparative results for normal events under various scenario settings.

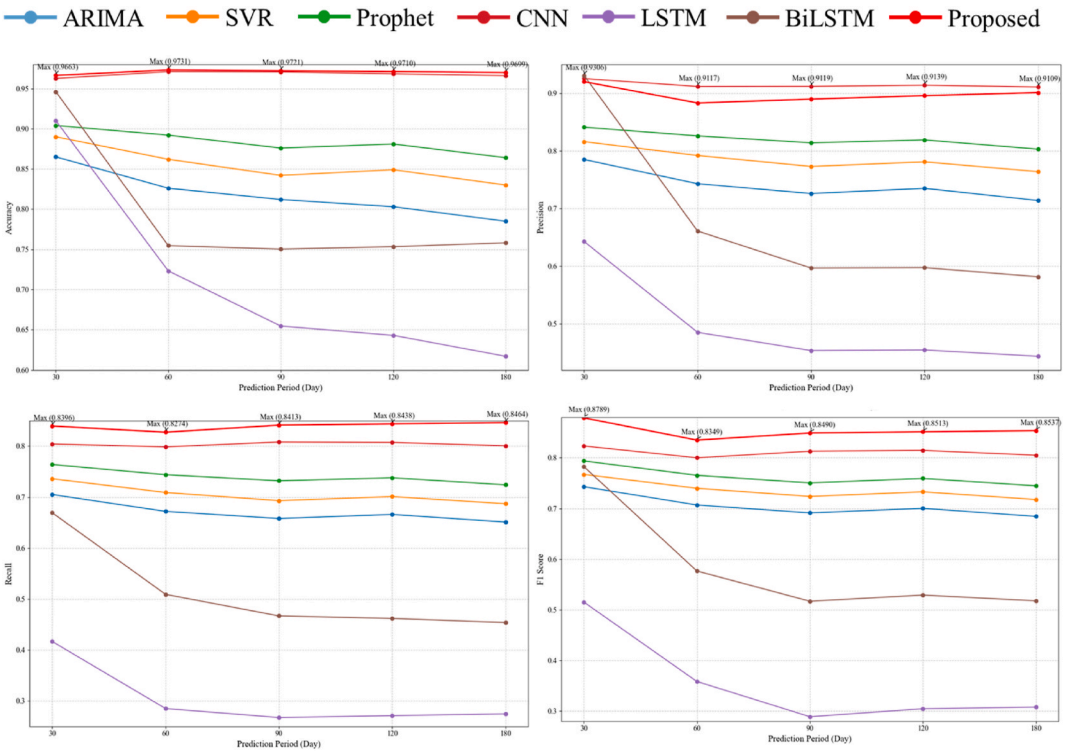

Fig. 16. Comparative results for low-severity events under various scenario settings.

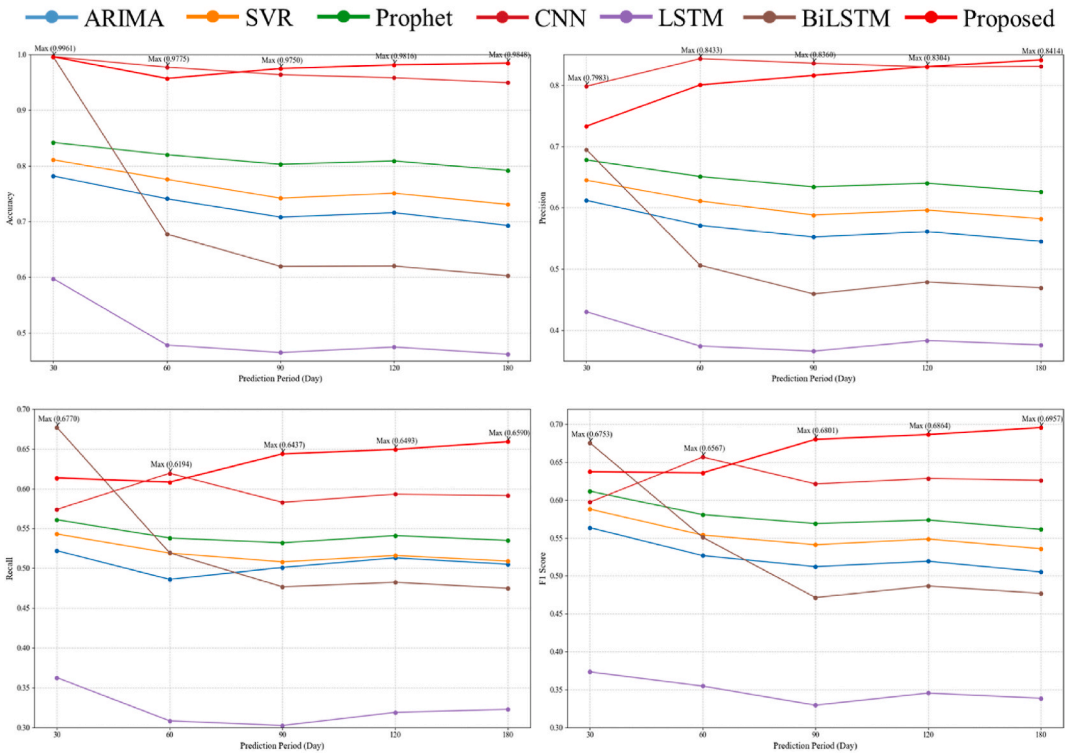

Fig. 17. Comparative results for high-severity events under various scenario settings.

6.4. *Platform Deployment*

This section details the creation of the prototype system designed for the verification and implementation of the DT concept. The platform encompasses two integral subsystems: (i) an as-built BIM visualization platform and (ii) an IAQ monitoring system. Together, these provide both spatial context and predictive intelligence.

(1) As-built BIM Visualization Platform

BIM has become a recognised standard in the construction industry for enhancing integration, efficiency, and collaboration. However, most BIM-enabled solutions depend on as-designed models, limiting their applicability to pre-BIM era buildings, whereas as-built BIM technologies enable advanced facility management for structures lacking original BIM data. In this study, IScan2BIM [74] were utilised to establish the as-built BIM models for target buildings to achieve visualization and offer detailed layout information, as shown in Fig. 19 (Left). The methodology flow is delineated in Fig. 20, and the practical demonstration can be accessed at https://www.youtube.com/watch?v=RE-vNYs_i_w. Firstly, an autonomous scanning robotic integrating advanced Light Detection and Ranging (LiDAR) technology named "BIMBot" is developed to produce a 3D point clouds dataset of the target building [75], which is followed by the AI engine tailored for BIM reconstruction. The semantic information of civil structures is extracted from point clouds using the BEACon algorithm [60] and then converted to the IFC format for the BIM reconstruction, whereas the MEP systems are constructed via Pipenet [59]. Finally, the Amazon Web Services cloud platform is utilised for the model integration and demonstration. IScan2BIM serves as the central platform for the integration of IAQ information (see Fig. 21).

(2) IAQ Monitoring System

To complement the visualization platform, a prototype application was developed using the Guietta and Tkinter Python libraries to demonstrate real-time integration among sensor data, predictive modelling, and user interfaces. The Kaiterra Sensedge Mini sensors continuously acquire $PM_{2.5}$, CO_2, TVOC, O_3, temperature, and humidity data at 1-min intervals. After data cleaning and normalization, the time-series are fed into the trained PyTorch model, which predicts IAQ indices, identifies dominant pollutants, and issues early-warning alerts upon threshold exceedance. As illustrated in Fig. 19 (right), a desktop interface visualizes real-time sensor and model outputs, including IAQ indices, alarm levels, and temporal dynamics. The current prototype operates using replayed data from an online Excel dataset for verification; however, its architecture supports direct sensor-to-model communication via MQTT or REST protocols, enabling real-time inference. Future developments will integrate a cloud-based time-series database and deploy the inference engine as a Flask or Django web service to facilitate online access, cross-platform monitoring, and interoperability with

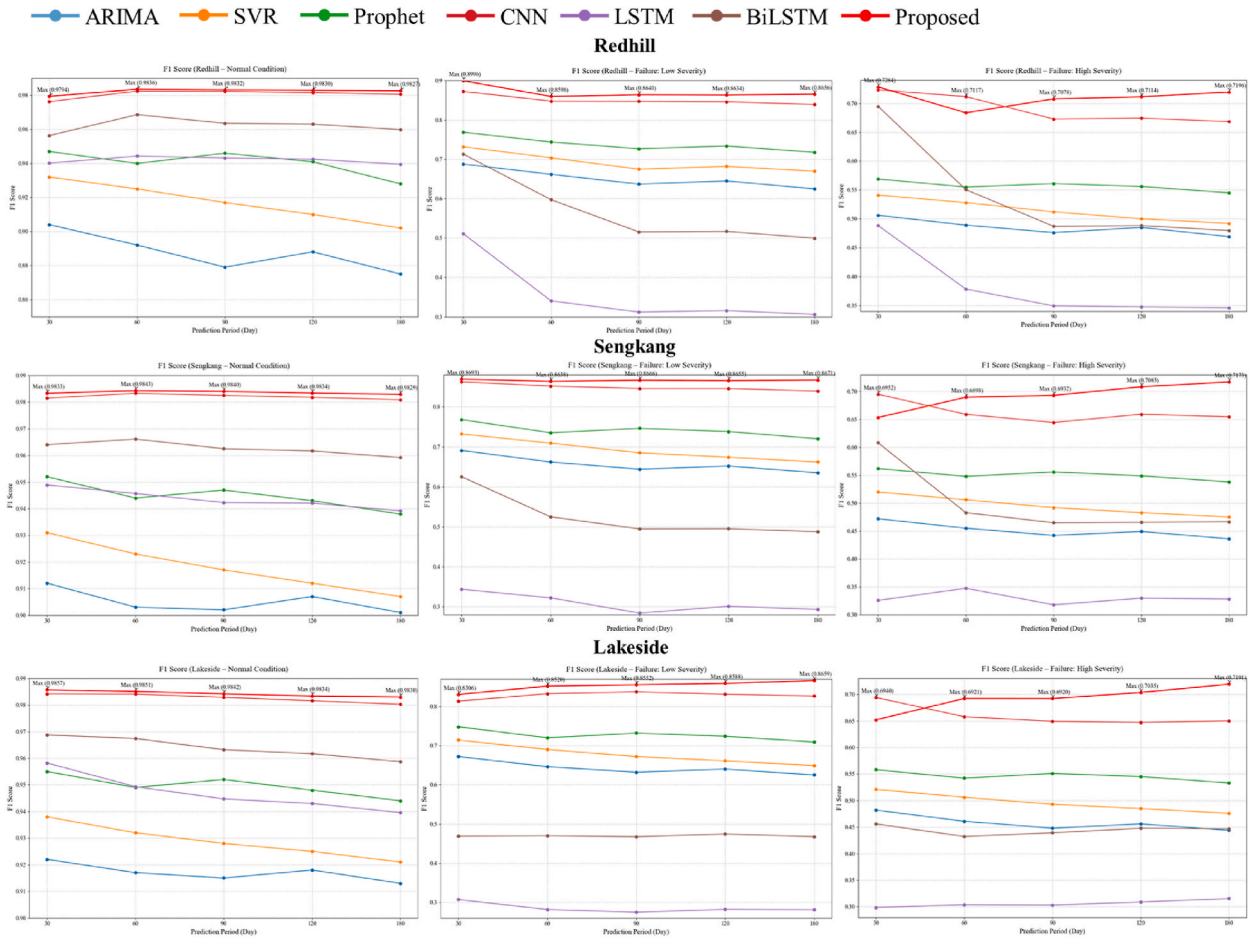

Fig. 18. Comparative results of different datasets.

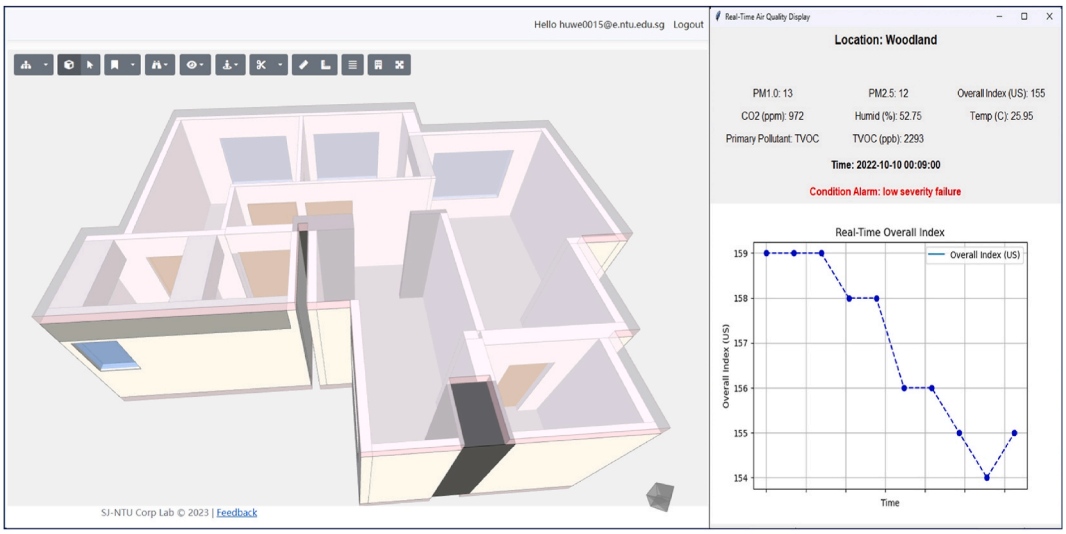

Fig. 19. Prototype DT system.

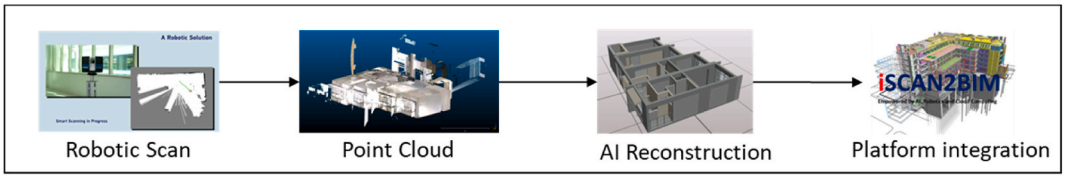

Fig. 20. IScan2BIM methodology.

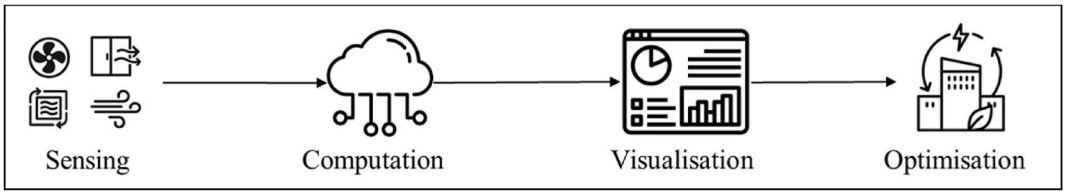

Fig. 21. Deployment flow.

facility management systems.

(3) Future Implementation

Future development will extend the prototype into a fully functional DT system integrating real-time sensing, predictive analytics, and as-built BIM visualization within a unified cloud framework. The Kaiterra Sensedge Mini network will continuously transmit $PM_{2.5}$, CO_2, TVOC, O_3, temperature, and humidity data via MQTT or REST APIs to a cloud-based time-series database. A real-time preprocessing module will filter anomalies, interpolate short gaps, and synchronize multi-sensor inputs before feeding the cleaned data into the developed model deployed as a Flask microservice. The model will generate IAQ forecasts, dominant pollutant identification, and early warnings, which are delivered through API endpoints to the operator interface and the IScan2BIM platform. The cloud-based BIM environment will map sensor and model outputs onto the 3D spatial layout, highlighting high-risk areas and temporal variations, while the dashboard will visualize real-time trends and alerts. Collectively, these modules establish a bidirectionally linked DT ecosystem that synchronizes physical and virtual spaces to enable real-time monitoring, predictive maintenance, and intelligent facility management.

7. Conclusion

Indoor climate maintenance in the building industry has garnered increasing attention from both academic circles and because of its profound influence on the well-being and productivity of occupants [76]. Furthermore, there is a growing imperative to understand how technological advancements in indoor climate systems extend benefits beyond academic interest and contribute meaningfully to societal welfare, thereby potentially reducing healthcare costs and improving the general quality of life. This study delves into the feasibility of incorporating the DL algorithm and DT perspective for failure prediction practices within the indoor climate domain. Practical air quality datasets collected from different buildings are utilised as the foundational resource to support the study and analysis. The contributions are listed below.

1. A DT-enabled PdM framework for the failure prediction of IAQ has been proposed, incorporating building environment virtualisation and indoor climate demonstration. The proposed framework facilitates the future implementation of DT solutions and additional function integration.
2. An ensemble DL model combining CNN, BiLSTM, and MHA mechanisms have been developed for predicting indoor climate failures, leveraging their unique strengths to enhance the model's predictive accuracy and reliability.
3. An evaluation and comparative analysis have been conducted with three other benchmark models using 4 different datasets, demonstrating the superior performance of the proposed method under different scenarios and the adaptability to different datasets. The F1 score of the proposed method consistently exceeds 0.8 for normal and low-severity events, whereas it approximates 0.7 for the high-severity cases.
4. A prototype DT system for as-built BIM reconstruction and indoor climate failure alarm has been established to showcase the DT concept and enhance the visualization during the experiment.
5. A comprehensive real-world dataset on IAQ has been collected from different locations in Singapore, addressing the issue of data scarcity and contributing to future research in this field.

Future efforts could further advance the PdM implementation in indoor climate control areas. Firstly, some pollutants such as SO_2 and NO_2 are not considered due to the limitation of IoT devices, can be addressed by the utilization of enhanced air quality sensor

systems in the future. Secondly, the integration of functionalities such as energy optimisation and smart ventilation control into the DT platform can provide more holistic and preventative solutions. Thirdly, the prototype DT system currently does not support real-time cloud data uploading, and its user interface could be further refined to enhance user-friendliness. Privacy protection, computing resources, and economic benefits also need more exploration during the practical deployment. Lastly, future work will not only look to bridge technical gaps but also examine this research's societal implications.

Besides, future extensions aim to extend the proposed framework toward large-scale building complexes comprising multiple interconnected subsystems, including HVAC, lighting, and occupant-control networks. Achieving such scalability requires multi-agent DT coordination, cross-building data fusion, and distributed or federated learning strategies to enable edge-level PdM. Moreover, the underlying algorithmic architecture could be enhanced through the incorporation of adaptive attention mechanisms and graph-based spatial–temporal representations, which can capture intricate inter-zone dependencies across buildings. These advancements would strengthen model generalisation, improve computational efficiency, and enable real-time, data-driven decision-making for district-level building clusters.

Overall, this study advances the PdM field by transitioning from condition monitoring and anomaly detection to failure prediction through the development of the advanced DL approach. Besides, the proposed DT framework lays a solid foundation for future research into scalable, intelligent, and collaborative PdM systems capable of supporting complex, interconnected building environments. The research findings and insights lay the groundwork for DT-enabled PdM strategies, while the acknowledged limitations serve as a roadmap for directing future research and development efforts.

CRediT authorship contribution statement

Wei Hu: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Cheng Chang:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Han Wu:** Writing – original draft, Methodology, Investigation, Conceptualization. **Kang Lai:** Writing – original draft, Methodology, Conceptualization. **Yiyu Cai:** Writing – review & editing, Project administration, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is partially supported by the Natural Science Foundation of Jiangsu Province under Grant BK20250728. The authors acknowledge Nanyang Technological University for providing the experimental site for BIM reconstruction, as well as Delft University of Technology and Chang'an University for their collaborative contributions to this study. No additional external funding was received from other institutions.

Data availability

Data will be made available on request.

References

[1] J. Bower, et al., Monitoring ambient air quality for health impact assessment, World Heal. Organ. Reg. Publ. - Eur. Ser. (85) (1999).
[2] W. Hu, X. Wang, K. Tan, Y. Cai, Digital twin-enhanced predictive maintenance for indoor climate : a parallel LSTM-autoencoder failure prediction approach, Energy Build. 301 (November) (2023) 113738, https://doi.org/10.1016/j.enbuild.2023.113738.
[3] S.H. Khajavi, N.H. Motlagh, A. Jaribion, L.C. Werner, J. Holmstrom, Digital Twin: Vision, benefits, boundaries, and creation for buildings, IEEE Access 7 (2019) 147406–147419, https://doi.org/10.1109/ACCESS.2019.2946515.
[4] I. Errandonea, S. Beltrán, S. Arrizabalaga, Digital Twin for maintenance: a literature review, Comput. Ind. 123 (2020), https://doi.org/10.1016/j.compind.2020.103316.
[5] M. Jasiulewicz-Kaczmarek, S. Legutko, P. Kluk, Maintenance 4.0 technologies - new opportunities for sustainability driven maintenance, Manag. Prod. Eng. Rev. 11 (2) (2020) 74–87, https://doi.org/10.24425/mper.2020.133730.
[6] T.O. Alamri, J.P.T. Mo, Optimisation of preventive maintenance regime based on failure mode system modelling considering reliability, Arab. J. Sci. Eng. 48 (3) (2023) 3455–3477, https://doi.org/10.1007/s13369-022-07174-w.
[7] Y. Wang, H. Cheng, M.Q.H. Meng, Spatiotemporal Co-Attention hybrid neural network for pedestrian localization based on 6D IMU, IEEE Trans. Autom. Sci. Eng. 20 (1) (2023) 636–648, https://doi.org/10.1109/TASE.2022.3164966.
[8] A. Zhang, Z. Min, Y. Wang, M.Q.H. Meng, Towards an accurate augmented-reality-assisted orthopedic surgical robotic system using bidirectional generalized point set registration, in: IEEE Int. Conf. Intell. Robot. Syst., 2023, pp. 4600–4607, https://doi.org/10.1109/IROS55552.2023.10341401.
[9] H. Cheng, Y. Wang, M.Q.H. Meng, A vision-based robot grasping system, IEEE Sens. J. 22 (10) (2022) 9610–9620, https://doi.org/10.1109/JSEN.2022.3163730.
[10] N.M. Tiglao, M. Alipio, R. Dela Cruz, F. Bokhari, S. Rauf, S.A. Khan, Smartphone-based indoor localization techniques: State-of-the-art and classification, Meas. J. Int. Meas. Confed. 179 (November 2020) (2021) 109349, https://doi.org/10.1016/j.measurement.2021.109349.
[11] S.J. Emmerich, K.Y. Teichman, 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. 23 (7) (2017) 1159–1166, https://doi.org/10.1080/23744731.2016.1274627.
[12] P.E.D. Love, J. Matthews, The 'how' of benefits management for digital technology: from engineering to asset management, Autom. Constr. 107 (July) (2019), https://doi.org/10.1016/j.autcon.2019.102930.

[13] H.H. Hosamo, H.K. Nielsen, A.N. Alnmr, P.R. Svennevig, K. Svidt, A review of the digital twin technology in the AEC-FM industry, Front. Built Environ. 8 (2022), https://doi.org/10.3389/fbuil.2022.1013196.
[14] F. Hodavand, I.J. Ramaji, N. Sadeghi, Digital twin for fault detection and diagnosis of building operations: a systematic review, Buildings 13 (6) (2023), https://doi.org/10.3390/buildings13061426.
[15] G. Piras, S. Agostinelli, F. Muzi, Digital twin framework for built environment: a review of key enablers, Energies 17 (2) (2024), https://doi.org/10.3390/en17020436.
[16] J. Zhao, H. Feng, Q. Chen, 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. 49 (October 2021) (2022) 1–12, https://doi.org/10.1016/j.jobe.2022.104028.
[17] I. Yitmen, S. Alizadehsalehi, İ. Akıner, M.E. Akıner, An adapted model of cognitive digital twins for building lifecycle management, Appl. Sci. 11 (9) (2021) 4276, https://doi.org/10.3390/app11094276.
[18] D. Quirk, J. Lanni, N. Chauhan, Digital twins: details of implementation, ASHRAE J. 62 (10) (2020) 20–24.
[19] Y. Peng, M. Zhang, F. Yu, J. Xu, S. Gao, Digital twin Hospital buildings: an exemplary case study through continuous lifecycle integration, Adv. Civ. Eng. 2020 (2020), https://doi.org/10.1155/2020/8846667.
[20] L. Zhao, et al., Digital twin evaluation of environment and health of public toilet ventilation design based on building information modeling, Buildings 12 (4) (2022), https://doi.org/10.3390/buildings12040470.
[21] M. Pregnolato, et al., Towards Civil engineering 4.0: concept, workflow and application of digital Twins for existing infrastructure, Autom. Constr. 141 (July) (2022) 104421, https://doi.org/10.1016/j.autcon.2022.104421.
[22] Y. Zu, R. Dai, Distributed path planning for building evacuation guidance, Cyber-Physical Syst 3 (1–4) (2017) 1–21, https://doi.org/10.1080/23335777.2017.1326983.
[23] A. Bonci, A. Carbonari, A. Cucchiarelli, L. Messi, M. Pirani, M. Vaccarini, A cyber-physical system approach for building efficiency monitoring, Autom. Constr. 102 (August 2018) (2019) 68–85, https://doi.org/10.1016/j.autcon.2019.02.010.
[24] B. Montalbán Pozas, B. Muriel Holgado, M. Lucas Bonilla, S. Barroso Ramírez, P. Bustos García de Castro, Iterative optimization of a social inmotics-based method in order to make buildings smart and resilient, Sustain. Cities Soc. 82 (January) (2022), https://doi.org/10.1016/j.scs.2022.103876.
[25] H. Pruvost, F. Forns-Samso, O. Gnepper, O. Enge-Rosenblatt, Integrating energy system monitoring and maintenance services into a BIM-based digital twin, in: IECON Proc. (Industrial Electron. Conf., 2023, pp. 1–6, https://doi.org/10.1109/IECON51785.2023.10311659.
[26] Y. Zhou, Y. Su, Z. Xu, X. Wang, J. Wu, X. Guan, A hybrid physics-based/data-driven model for personalized dynamic thermal comfort in ordinary office environment, Energy Build. 238 (2021) 110790, https://doi.org/10.1016/j.enbuild.2021.110790.
[27] B.L. Thomas, D.J. Cook, Activity-aware energy-efficient automation of smart buildings, Energies 9 (8) (2016), https://doi.org/10.3390/en9080624.
[28] M. Shahinmoghadam, W. Natephra, A. Motamedi, BIM- and IoT-based virtual reality tool for real-time thermal comfort assessment in building enclosures, Build. Environ. 199 (April) (2021) 107905, https://doi.org/10.1016/j.buildenv.2021.107905.
[29] Z. Ni, Y. Liu, M. Karlsson, S. Gong, A sensing system based on public cloud to monitor indoor environment of historic buildings, Sensors 21 (16) (2021), https://doi.org/10.3390/s21165266.
[30] J. Zhang, H.H.L. Kwok, H. Luo, J.C.K. Tong, J.C.P. Cheng, Automatic relative humidity optimization in underground heritage sites through ventilation system based on digital twins, Build. Environ. 216 (January) (2022), https://doi.org/10.1016/j.buildenv.2022.108999.
[31] B. Keskin, B. Salman, 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. 148 (5) (2022) 1–15, https://doi.org/10.1061/(asce)co.1943-7862.0002271.
[32] S. HosseiniHaghighi, P.M.Á. de Uribarri, R. Padsala, U. Eicker, Characterizing and structuring urban GIS data for housing stock energy modelling and retrofitting, Energy Build. 256 (2022) 111706, https://doi.org/10.1016/j.enbuild.2021.111706.
[33] B. Bass, J. New, W. Copeland, Potential Engergy, Demand, Emissions, and Cost Savings Distributions for Buildings in a Utility ' S Service Area, 2021.
[34] W. Huang, Y. Zhang, W. Zeng, Development and application of digital twin technology for integrated regional energy systems in smart cities, Sustain. Comput. Informatics Syst. 36 (August) (2022) 100781, https://doi.org/10.1016/j.suscom.2022.100781.
[35] S. Agostinelli, F. Cumo, Machine learning approach for predictive maintenance in an advanced building management system, WIT Trans. Ecol. Environ. 255 (2022) 131–138, https://doi.org/10.2495/EPM220111.
[36] Z. Wu, Y. Chang, Q. Li, R. Cai, A novel method for tunnel digital twin construction and virtual-real fusion application, Electron 11 (9) (2022), https://doi.org/10.3390/electronics11091413.
[37] S. Kaewunruen, J. Sresakoolchai, W. Ma, O. Phil-Ebosie, Digital twin aided vulnerability assessment and risk-based maintenance planning of bridge infrastructures exposed to extreme conditions, Sustain. Times 13 (4) (2021) 1–19, https://doi.org/10.3390/su13042051.
[38] V. Villa, B. Naticchia, G. Bruno, K. Aliev, P. Piantanida, D. Antonelli, Iot open-source architecture for the maintenance of building facilities, Appl. Sci. 11 (12) (2021), https://doi.org/10.3390/app11125374.
[39] A.B. Nassif, M.A. Talib, Q. Nasir, F.M. Dakalbab, Machine learning for Anomaly detection: a systematic review, IEEE Access 9 (2021) 78658–78700, https://doi.org/10.1109/ACCESS.2021.3083060.
[40] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, A. Amira, Artificial intelligence based anomaly detection of energy consumption in buildings: a review, current trends and new perspectives, Appl. Energy 287 (November 2020) (2021) 116601, https://doi.org/10.1016/j.apenergy.2021.116601.
[41] H.H. Hosamo, H.K. Nielsen, D. Kraniotis, P.R. Svennevig, K. Svidt, Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings, Energy Build. 281 (2023) 112732, https://doi.org/10.1016/j.enbuild.2022.112732.
[42] E. de Rautlin de la Roy, T. Recht, A. Zemmari, P. Bourreau, L. Mora, Data analytics for smart buildings: a classification method for anomaly detection for measured data, J. Phys. Conf. Ser. 2042 (1) (2021), https://doi.org/10.1088/1742-6596/2042/1/012015.
[43] R. Chandra, et al., A Survey of failure mechanisms and statistics for critical electrical equipment in buildings, in: IECON Proc. (Industrial Electron. Conf., 2020, pp. 1955–1961, https://doi.org/10.1109/IECON43393.2020.9254225, vol. 2020-Octob.
[44] H. Kayan, Y. Majib, W. Alsafery, M. Barhamgi, C. Perera, AnoML-IoT: an end to end re-configurable multi-protocol anomaly detection pipeline for internet of things, Internet of Things (Netherlands) 16 (July) (2021) 100437, https://doi.org/10.1016/j.iot.2021.100437.
[45] H.M. Chen, C.C. Hou, Y.H. Wang, A 3D visualized expert system for maintenance and management of existing building facilities using reliability-based method, Expert Syst. Appl. 40 (1) (2013) 287–299, https://doi.org/10.1016/j.eswa.2012.07.045.
[46] Q. Lu, X. Xie, A.K. Parlikad, J.M. Schooling, Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance, Autom. Constr. 118 (March) (2020) 103277, https://doi.org/10.1016/j.autcon.2020.103277.
[47] A. Malki, E.S. Atlam, I. Gad, Machine learning approach of detecting anomalies and forecasting time-series of IoT devices, Alexandria Eng. J. 61 (11) (2022) 8973–8986, https://doi.org/10.1016/j.aej.2022.02.038.
[48] A. Gálvez, A. Diez-Olivan, D. Seneviratne, D. Galar, Fault detection and RUL estimation for railway HVAC systems using a hybrid model-based approach, Sustain. Times 13 (12) (2021), https://doi.org/10.3390/su13126828.
[49] J. Xia, Y. Feng, C. Lu, C. Fei, X. Xue, LSTM-based multi-layer self-attention method for remaining useful life estimation of mechanical systems, Eng. Fail. Anal. 125 (December 2020) (2021) 105385, https://doi.org/10.1016/j.engfailanal.2021.105385.
[50] B. Li, W. Cai, A novel CO2-based demand-controlled ventilation strategy to limit the spread of COVID-19 in the indoor environment, Build. Environ. 219 (February) (2022) 109232, https://doi.org/10.1016/j.buildenv.2022.109232.
[51] Y. Wei, J. Jang-Jaccard, W. Xu, F. Sabrina, S. Camtepe, M. Boulic, LSTM-Autoencoder-Based anomaly detection for indoor air quality time-series data, IEEE Sens. J. 23 (4) (2023) 3787–3800, https://doi.org/10.1109/JSEN.2022.3230361.
[52] Y. Liu, Z. Pang, M. Karlsson, S. Gong, Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control, Build. Environ. 183 (July) (2020) 107212, https://doi.org/10.1016/j.buildenv.2020.107212.
[53] Y.S. Jung, T.W. Kang, C. Chun, Anomaly analysis on indoor office spaces for facility management using deep learning methods, J. Build. Eng. 43 (2021) 103139, https://doi.org/10.1016/j.jobe.2021.103139.

[54] A.R. Santiago, M. Antunes, J.P. Barraca, D. Gomes, R.L. Aguiar, Predictive maintenance system for efficiency improvement of heating equipment, in: 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, 2019, pp. 93–98, https://doi.org/10.1109/BigDataService.2019.00019. April.
[55] W. Hu, K.Y.H. Lim, Y. Cai, Digital Twin and Industry 4.0 Enablers in building and construction: a survey, Buildings (2022) 1–27.
[56] British Standards Institution (BSI), BSI Standards Publication Maintenance — Maintenance Terminology, 2010, p. 36.
[57] T. Czerniawski, F. Leite, Automated digital modeling of existing buildings: a review of visual object recognition methods, Autom. Constr. 113 (July 2019) (2020) 103131, https://doi.org/10.1016/j.autcon.2020.103131.
[58] C. Rausch, C. Haas, Automated shape and pose updating of building information model elements from 3D point clouds, Autom. Constr. 124 (August 2020) (2021) 103561, https://doi.org/10.1016/j.autcon.2021.103561.
[59] Y. Xie, S. Li, T. Liu, Y. Cai, As-built BIM reconstruction of piping systems using PipeNet, Autom. Constr. 147 (December 2022) (2023) 104735, https://doi.org/10.1016/j.autcon.2022.104735.
[60] T. Liu, Y. Cai, J. Zheng, N.M. Thalmann, BEACon: a boundary embedded attentional convolution network for point cloud instance segmentation, Vis. Comput. (2021), https://doi.org/10.1007/s00371-021-02112-7.
[61] Y. LeCun, K. Kavukcuoglu, C. Farabet, Convolutional networks and applications in vision, in: ISCAS 2010 - 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems, 2010, pp. 253–256, https://doi.org/10.1109/ISCAS.2010.5537907.
[62] J. Zhang, L. Zi, Y. Hou, M. Wang, W. Jiang, D. Deng, A deep learning-based approach to enable action recognition for construction equipment, Adv. Civ. Eng. 2020 (2020), https://doi.org/10.1155/2020/8812928.
[63] B. Zhao, H. Lu, S. Chen, J. Liu, D. Wu, Convolutional neural networks for time series classification, J. Syst. Eng. Electron. 28 (1) (2017) 162–169, https://doi.org/10.21629/JSEE.2017.01.18.
[64] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, P.A. Muller, Deep learning for time series classification: a review, Data Min. Knowl. Discov. 33 (4) (2019) 917–963, https://doi.org/10.1007/s10618-019-00619-1.
[65] Z. Huang, W. Xu, K. Yu, Bidirectional LSTM-CRF models for sequence tagging [Online]. Available: http://arxiv.org/abs/1508.01991, 2015.
[66] W. Li, F. Qi, M. Tang, Z. Yu, Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification, Neurocomputing 387 (2020) 63–77, https://doi.org/10.1016/j.neucom.2020.01.006.
[67] S. Siami-Namini, N. Tavakoli, A.S. Namin, The performance of LSTM and BiLSTM in forecasting time series, in: Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019, 2019, pp. 3285–3292, https://doi.org/10.1109/BigData47090.2019.9005997.
[68] K. Xia, J. Huang, H. Wang, LSTM-CNN architecture for human activity recognition, IEEE Access 8 (2020) 56855–56866, https://doi.org/10.1109/ACCESS.2020.2982225.
[69] Z. Xu, S. Li, W. Deng, Learning temporal features using LSTM-CNN architecture for face anti-spoofing, in: Proc. - 3rd IAPR Asian Conf. Pattern Recognition, ACPR 2015, 2016, pp. 141–145, https://doi.org/10.1109/ACPR.2015.7486482.
[70] A. Veltman, D.W.J. Pulle, R.W. De Doncker, Attention is all you need, in: 31st Conference on Neural Information Processing Systems (NIPS 2017), 2016, pp. 47–82, https://doi.org/10.1007/978-3-319-29409-4_3.
[71] B. Li, F. Cheng, X. Zhang, C. Cui, W. Cai, A novel semi-supervised data-driven method for chiller fault diagnosis with unlabeled data, Appl. Energy 285 (January) (2021) 116459, https://doi.org/10.1016/j.apenergy.2021.116459.
[72] 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.
[73] A. Kanchan, K. Gorai, P. Goyal, A review on air quality indexing system, Asian J. Atmos. Environ. 9 (2) (2015) 101–113, https://doi.org/10.5572/ajae.2015.9.2.101.
[74] Intelligent scan to BIM [Online]. Available: https://www.youtube.com/watch?v=RE-vNYs_i_w.
[75] N. Liang, Y.P. Ang, K. Yeo, X. Wu, Y. Xie, Y. Cai, BIMBot for autonomous laser scanning in built environments, Robotics 13 (2) (2024) 1–18, https://doi.org/10.3390/robotics13020022.
[76] Y. Liu, Z. Pang, M. Karlsson, S. Gong, Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control, Build. Environ. 183 (May) (2020) 107212, https://doi.org/10.1016/j.buildenv.2020.107212.