HVAC Smart Predictive Maintenance Using Machine Learning and Bayesian Network

Rinta Kridalukmana
Department of Computer Engineering, Universitas Diponegoro, Indonesia
rintakrida@ce.undip.ac.id (corresponding author)

Ayman Elgharabawy
College of Science and Medicine, The Australian National University, Australia
ayman.elgharabawy@anu.edu.au

Fahimeh Ramezani
Australian Artificial Intelligence Institute, University of Technology Sydney, Australia
fahimeh.ramezani@uts.edu.au

Mohsen Naderpour
Australian Artificial Intelligence Institute, University of Technology Sydney, Australia
mohsen.naderpour@uts.edu.au

Wahyul A. Syafei
Department of Electrical Engineering, Universitas Diponegoro, Indonesia
wasyafei@yahoo.com

Received: 15 November 2025 | Revised: 13 January 2026 and 25 January 2026 | Accepted: 2 February 2026

 DOI: https://doi.org/10.48084/etasr.16279

ABSTRACT

Smart Predictive Maintenance (SPM) features for the building's Heating, Ventilation, and Air Conditioning (HVAC) system are crucial for reducing energy consumption, improving scheduling, and detecting potential problems. Popular approaches, such as Machine Learning (ML) and probabilistic methods, are employed for SPM. These methods can be considered forward inference. However, since numerous interdependent HVAC components are involved, SPM requires not only forward but also backward inference (diagnostic capabilities). Given that such abilities have been underexplored, the present study proposes an SPM-based HVAC monitoring system that combines ML and Bayesian Network (BN). While ML is used to predict the status of the HVAC components, BN performs the diagnostic tasks. A case study was conducted at the Sydney Aquarium in Australia to demonstrate the implementation of the proposed approach. The ML model, trained using the Simple Logistic Regression (SLR) method, achieved an accuracy of 0.99, higher than the 0.92 obtained by using the Decision Tree (DT) and Logistic Regression (LR) methods. Furthermore, the BN was used to diagnose and estimate the probability of a component's performance degradation if another component was problematic. Among the key benefits of this proposed system is its potential to enhance operators' understanding of problems with HVAC systems.

Keywords-HVAC; predictive maintenance; machine learning; Bayesian network

I. INTRODUCTION

A building's air quality control system, known as the HVAC system, helps maintain air quality and thermal comfort. HVAC systems rely on their distribution systems, which circulate conditioned air to the desired rooms. These distribution systems depend on the type of coolant and the circulation method, and involve components such as Air Filters (AFs), fan coils, air ducts, water pipes, and water pumps. At facilities requiring a complex HVAC system, such as the Sydney Aquarium, which serves as a case study in this work, maintenance is significant. This is due to the interconnectedness and interdependence of components within the HVAC system. If a component is underperforming, it increases the probability of failure in the others.

There are several approaches to maintenance, including Planned Preventive Maintenance (PPM), Unplanned Reactive Maintenance (URM), and SPM [1]. PPM is usually carried out periodically and may result in increased maintenance costs due to replacing components that are in good condition and still usable. As an alternative, URM can be employed, in which a component is replaced or repaired only when a problem occurs during system operation. URM can reduce costs when deployed for low-priority components, but it leads to higher risk and expenses when used for high-priority components [2]. SPM offers a solution by using continuous sensor network data to monitor and predict the status and condition of system components [3]. Such component observations can provide benefits, including predicting component replacement and optimizing energy use [4]. Additionally, SPM can boost productivity in industries [5] and advance the development of explainable Artificial Intelligence (XAI) to clarify faults in HVAC systems [6]. Especially with the involvement of AI and ML, the HVAC control system can be further enhanced by leveraging digital twin models to simulate and predict its optimization [7]. Although component observations rely on Internet of Things (IoT) technologies, which can be expensive during deployment, their benefits are evident in the long run [8]. The collected data are further processed to infer HVAC-maintenance-related information using three primary approaches: data-driven, hybrid physics–data-driven, and knowledge-driven [9].

The data-driven approach involves ML techniques, which can be supervised, semi-supervised, or unsupervised [10]. The applications of ML in supporting HVAC system maintenance have been investigated, with different objectives including energy management, Fault Detection and Diagnosis (FDD), and indoor environment control [11] across various components of the HVAC system. The indoor environment control includes chillers, Air Handling Units (AHUs), refrigerants, and air quality control [12]. To reduce power consumption, authors in [13] used an Artificial Neural Network (ANN) to predict the energy usage of HVAC systems in a warehouse over 5-h intervals, achieving an error rate of 7%. Regarding FDD, authors in [14] conducted studies on diagnosing malfunctions in electronic components using a statistical ML method that combines a hidden Markov model and a data fusion approach. Their approach also employed clustering and optimization techniques to localize normal and faulty patterns.

Authors in [15] proposed a cloud-based service to monitor refrigerant leakage and compared four ML models, including DTs, K-Nearest Neighbors, Support Vector Machines, and Random Forests, to recognize the leakage problems. The DT and Random Forest achieved the highest accuracy of 95%, with execution times of 0.7 and 3.32 s, respectively. Authors in [16] applied Random Forests to identify six malfunction-related variables with up to nine severity levels, achieving an accuracy of approximately 98%. Authors in [17] deployed a semi-supervised ANN to cluster faults in HVAC components, achieving 86% accuracy.

The data-driven approach is often combined with a physics-based model known as the hybrid physics-data-driven approach. In these techniques, the physics-based model is used to generate synthetic data, followed by a data-driven learning process. An example of such a hybrid mode is presented in [18], which proposed an approach to diagnosing HVAC performance. The synthetic data were generated from two scenarios: healthy and faulty HVAC operation. After that, the former were trained using ML methods such as Support Vector Machine and deep learning. Authors in [19] used a hybrid method to detect faults in fan-coil units. They first developed mathematical models of HVAC subsystem operations to reduce the input space and then utilized polynomial regression to generate steady-state models to predict faults. However, simple rules were also deployed to facilitate fault isolation. The hybrid method is often preferred, especially when the ML model has limited historical data to learn from.

A knowledge-driven approach is a white-box model, usually represented using if-then rules or inference-based methods such as Fuzzy logic and Bayesian inference. Authors in [20-22] employed the knowledge-driven approach. Specifically, authors in [20] proposed a fuzzy classifier to recognize faults in the dehumidifier component of the HVAV system, which was optimized using an adaptive genetic algorithm. Compared with a traditional fuzzy classifier, their optimized classifier achieved better fault-detection performance. In [21, 22], Bayesian inference was used to support fault detection. Authors in [21] employed Bayesian inference to estimate HVAC component observation values from a set of sensors, accounting for the standard deviation of the likelihood function. However, the combination of Markov Chain and Monte Carlo method is used to address the challenge of calculating the Bayesian probability function, particularly to down-sample posterior distributions without complex integration. The results showed that the HVAC system operating state could be recognized to determine the fault position, thereby reducing the deviation rate by up to 98%. In contrast, the Bayesian method in [22] was used to detect HVAC sensor faults and handle incomplete measurements.

Data-driven, hybrid, and knowledge-driven methods have been employed to detect faults in specific HVAC components. Therefore, they are suitable for both URM and SPM. In URM, the FDD technique detects current incidents (e.g., leakage, cooler malfunctions), while SPM predicts immediate future incidents or the causal effects on other HVAC components when an incident occurs. Authors in [23] proposed a prediction model for the remaining useful life of HVAC AFs. The physics-based model included inputs from thermal networking based on principles of fluid mechanics, heat transfer, and thermodynamics. These inputs were used to generate synthetic data and were trained using several ML methods. It was indicated that the remaining useful life of HVAC AFs can be predicted with 96% accuracy. Authors in [24] proposed detecting HVAC system-level faults using historical data and knowledge. Interdependencies were identified to support multiple operating modes. However, as the proposed approach was data-driven, the diagnosing feature was considered a forward inference that was limited to predicting fault probability in the immediate future, without identifying causal problems.

Authors in [25, 26] successfully developed ML approaches utilizing a Recurrent Neural Network to detect faults in the SPM environment. In addition to a faulty detector, authors in

[26] proposed an IoT-based architecture to support the SPM solution. However, their approaches need to be verified at a global level of HVAC systems. Authors in [27, 28] introduced an HVAC fault detection system using a data-driven approach, specifically focusing on water chillers and air conditioners. However, the data in [27] were trained using a semi-supervised approach to minimize reliance on prior knowledge of abnormal situations. Besides detecting malfunctions, authors in [29, 30] used ML to optimize energy usage and predict maintenance time. Specifically, they deployed hybrid neural networks. While authors in [29] combined narrow neural networks with Support Vector Machines, other studies combined three types of neural networks, including feed-forward, cascade-forward backpropagation, and Elman backpropagation. Despite their outstanding performance in such optimization and prediction, the case study in [29] was limited to a component, the chilled-beam air conditioning system, while the case study in [30] focused on a residential HVAC system.

Authors in [31] introduced an SPM approach based on AI, namely a causal-based SPM exploiting d-separation and d-connection to model cause-and-effect relationships and generate maintenance explanations. The use of a causal-based approach overcomes the limitations of ML-based approaches that operate in a black-box mode. However, the causal-based method focuses only on forward inference. In HVAC systems, reduced component performance may result from degradation in its constituent parts [32]. Therefore, a diagnostic ability (in terms of backward inference) is required [33].

There is a lack of backward inference in current HVAC SPM research. Therefore, the present study proposes an HVAC system monitoring tool that implements such inference capabilities in SPM using ML and Causal Networks (CN). Three models generated from ML with the SLR, LR, and DT methods were evaluated to classify the condition of the AF as the primary input to simulate the proposed SPM Monitoring Tool (AP-SPM). The outputs from this evaluation were processed using CN, followed by implementation using BN techniques to represent the reliability between components.

BN was chosen for its capabilities to serve forward and backward inference. The former also estimated the probability of the component status and condition as part of predicting maintenance requirements. Multiple sensors—encompassing devices that measure temperature, humidity, and air pressure—were employed to collect significant data. The experimental results showed that the SLR model performed the best (0.99), compared to the DT- and LR-generated models, which obtained an accuracy of 0.92. A sensitivity analysis was used to validate the BN. Based on the magnitude and sensitivity values, the results indicated that the reliability model can identify the probable causal problem when the performance of one of the HVAC components declines. The contributions of this study are:

- A novel way to model HVAC components' relationships using a BN.
- A novel HVAC backward inference diagnostic ability based on ML and the developed BN.

II. THE PROPOSED SMART PREDICTIVE MAINTENANCE

The architecture of the proposed AP-SPM is illustrated in Figure 1. It consists of blocks representing HVAC components, a sensor network, an SPM module, and a visualization. Sensors were installed on the desired HVAC system components to monitor their status. Data from the sensors were received by the AP-SPM engine, which comprised two main parts: the ML detection model and the inter-component reliability model. The results of the AP-SPM engine were visualized in a dashboard tool for technical operators.

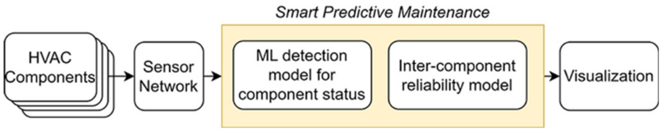

Fig. 1. Architecture of the proposed AP-SPM.

A. HVAC System Components

Figure 2 depicts the components of the HVAC system at the Sydney Aquarium. To assist the HVAC system's cooling and heating functions, seawater and potable water are used. A water pump is utilized to draw water from the sea, before being distributed to the heat exchanger component and then mixed with the potable water. The seawater is filtered to remove sand and other debris. This mixed water is subsequently transported to another heat exchanger to absorb and remove heat from the cooling machine, producing condenser water, which is then pumped to the AHU to assist cooling and coil heating operations. These coils regulate the condition of outdoor air and distribute it to the rooms via air ducts. The latter are equipped with AFs, additional coolers (inter-chillers), and air flow regulators (supply air diffusers). Simultaneously, a ventilation system draws air out of the room and returns it to the AHU, maintaining indoor air circulation.

B. Handling Input Data from Sensors

The input data for the AP-SPM are provided by a network of sensors installed in the HVAC system. Let γ_t denote the set of values from the sensors at a specific time (t), then:

$$\gamma_t = \{\omega^c_{i,t}\}^n_{i=1} \quad (1)$$

where ω is the value of a particular sensor located in cluster c, and i is the sequence number of the sensor.

The primary issue with integrating sensor input into the AP-SPM system is the potential for data loss at any given time. This data loss can be temporary (missing for a moment and then returning to normal), for example, due to interference with the transmission signal or sensor anomalies. Meanwhile, permanent data loss is usually caused by sensor damage and the loss of the sensor's power source. Temporary data loss at time t is calculated using:

$$\omega^c_{i,t} = \frac{\omega^c_{i,t-1} + \omega^c_{i,t-2}}{2} \quad (2)$$

where $\omega^c_{i,t}$ is the average of the last 2 data points from t-1 and t-2.

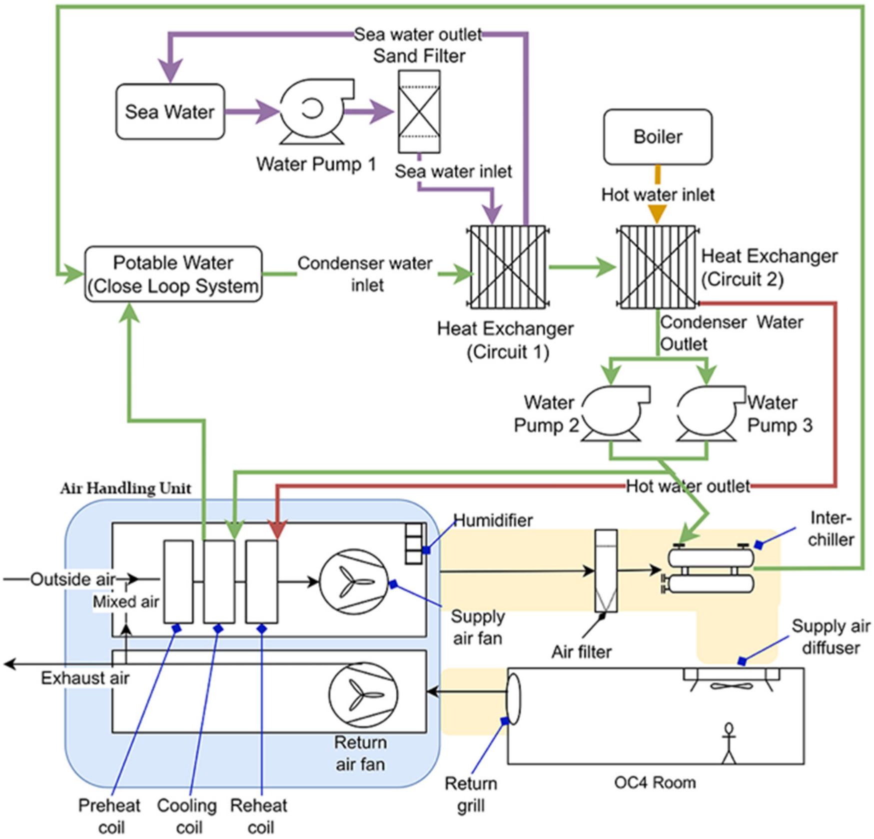

Fig. 2. Illustration of the HVAC system at the Sydney Aquarium.

If, within a given time interval (e.g., 20 h, assuming data are collected every 1 h), 50% of the data are lost either sequentially or randomly, a notification is sent indicating that a sensor anomaly has occurred. During that time unit, if all sensor data are generated by (2), a notification is sent to explain the possibility that the sensor is inactive. In sensor networks, sensors are often located far from the primary power source. Therefore, using batteries is an alternative. Also, the sensor's data-reading cycle needs to be monitored at an appropriate interval (e.g., every 1 or 2 h) to conserve battery life. Some sensors feature a battery capacity check, which significantly simplifies monitoring sensor status.

C. *Component Status Detection*

In HVAC systems, some sensors can be directly interpreted by the AP-SPM system to represent component status, such as pump water pressure. However, there are also components whose status must be observed through data from multiple sensors, such as an AF. For these cases, AP-SPM uses ML methods. The ML method produces a model to perform a classification task φ, which predicts the condition of the components of a particular cluster *c* (e.g., *c* = *x*) at a particular time *t* based on:

$$\varphi = \delta(f(\lambda_t)) \mid c = x \qquad (3)$$

where δ is a function that transforms a set of sensor values from λ_t based on a classification function $f(\lambda_t)$, which produces two types of conditions/states (referred to as binary outcomes), such as 0/1 or true/false.

If the same component is located in different locations or sourced from different vendors, the model needs to be calibrated by retraining it with relevant data. For example, AFs in distribution channels sometimes come from different vendors, types, and sizes. These differences are defined by the cluster code *c*. Various methods can be used to generate ML models, such as SLR, LR, or DT. SLR uses the sigmoid function to model the probability of a predicted outcome *y* having a value of 1 (denoted as *P(y = 1)*) from a combination of input variables *z*. This function can be further defined as:

$$P(y = 1) = \frac{1}{(1+e^{(-\theta_0+\theta_1 z)})} \qquad (4)$$

where e is the Euler's constant ($\pm$ 2.781), θ_0 is the bias value, and θ_1 is the weight coefficient of the input variable z. The expression $(-\theta_0 + \theta_1 z)$ is a linear combination that produces a positive or negative value, so that the fraction value, which is the sigmoid function, ranges between 0 and 1.

D. *Inter-Component Reliability Model*

The inter-component reliability model shows the influence of one component on another in an HVAC system. In the AP-SPM system proposed in this article, a CN model applying the BN technique is used. This technique was chosen since it can perform both forward and backward diagnosis. A BN can be defined as a graph *G*:

$$G = ((N, A), \wp) \quad (5)$$

where N is a set of nodes, A is a set of paths connecting nodes, and $\wp$ is a set of probability distributions. For a node other than the root node, the probability distribution of that node is a conditional probability distribution that indicates the dependency between that node and its contributing nodes (parent nodes). Each node $n \in N$ has a finite number of states and is mutually exclusive, meaning that two events cannot occur simultaneously. The state of each node (n) can be defined as:

$$K_n = \{K_1^n, K_2^n, \dots, K_j^n\} \quad (6)$$

In developing reliability relationships between components, knowledge from relevant experts is essential. All nodes of the BN, representing the reliability relationships between components, receive input or evidence from the HVAC sensors or the ML model. The BN inference is activated by two triggers. The first trigger is when a sensor observing HVAC component conditions reaches a threshold value. The second trigger occurs when the ML detects real-time problems with observed HVAC components. Depending on the BN design, conflicting predictions are possible; these can be avoided by using mutually exclusive enforcement during the conditional probability table.

E. Visualization

A web-based application is used to visualize the results of the AP-SPM system. This system monitors the condition of sensors, HVAC system components, and room conditions. The application also provides notifications to inform the technical operators responsible for HVAC system operations. Additionally, notifications can be sent to operators via other channels, such as email or text messages.

III. RESULTS AND DISCUSSION

A. Case Description, Scope of Observation, and Data Collection

To demonstrate the operation of the proposed AP-SPM system, the case study focuses on input initiation from the AF component in the HVAC system, with predictive maintenance being the primary focus. Additionally, the condition of the AF component is determined using observations from multiple types of sensors, namely those measuring air velocity, temperature, humidity, and the air pressure difference across the AF.

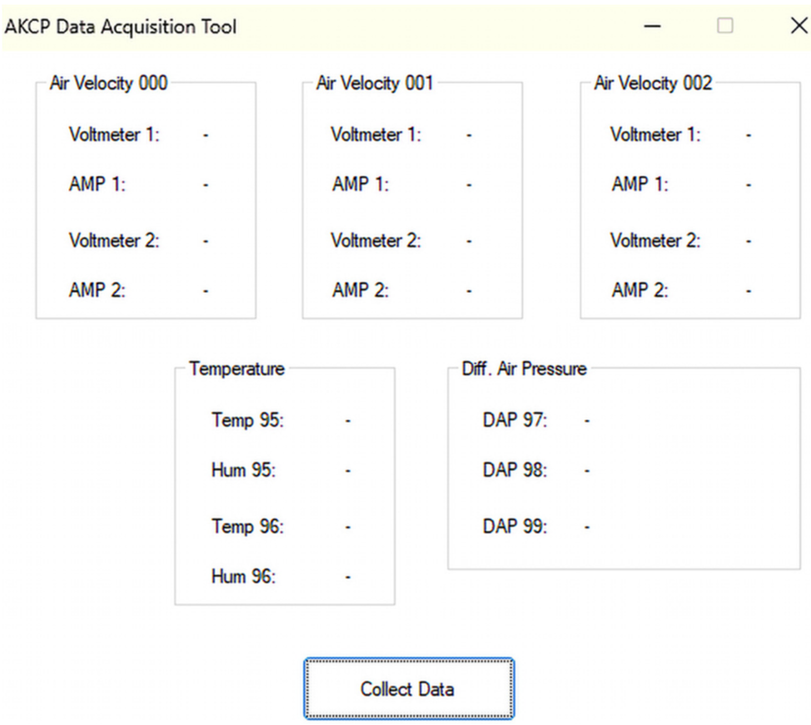

Fig. 3. Software interface during data acquisition.

All sensors used in this experiment were manufactured by AKCP. Software for data acquisition was developed, and its interface is shown in Figure 3. An example of sensor installation is provided in Figure 4.

Data were collected over a three-month period, during which the sensors recorded observations every 15 min. In addition to observation data on AFs that were fit for use (class 0), observation data were collected on those that were no longer fit for use due to dirt and labeled as class 1. There were 22.190 instances, with 90% being class 0; the remaining were class 1. Feature engineering was applied to the collected data, particularly to differentiate the operational hours from the date/time. Moreover, missing values were also handled during data collection by using (2).

Fig. 4. Sensor installation in the air duct.

The observation was divided into two different conditions. The first condition represented normal operation, with all components functioning properly and clean AFs installed. In the second condition, the water pump experienced performance degradation, affecting the AF, while dirty AFs were also deployed as a benchmark. These two conditions were used to label the dataset; the normal condition was assigned as class 0, and the other one as class 1. They were selected as they were crucial for maintaining the air quality distributed to the OC room. A plot of the sensor data acquisition results is presented in Figure 5. In areas with limited access to power, data are transmitted via radio waves to a computer that serves as a database. The sensors are battery-powered and have a lifespan of up to 10 years when the data collection frequency is set to every 1 h. Data communication from the sensor to the database is displayed in Figure 6.

B. ML Model

Based on the observations and data collection, an ML model was developed using the SLR method. This model was created employing a supervised learning approach, where the collected data are labeled according to the specified conditions for a component. Regarding the AF component, Figure 7 shows the time series of pressure fluctuations for a dirty filter (red curve) and a clean filter (blue curve). The training process used 10-fold cross-validation. Hence, the training dataset percentage follows the rule of (k-1)/k of the total data, which is 90%. The remaining 10% is used for testing.

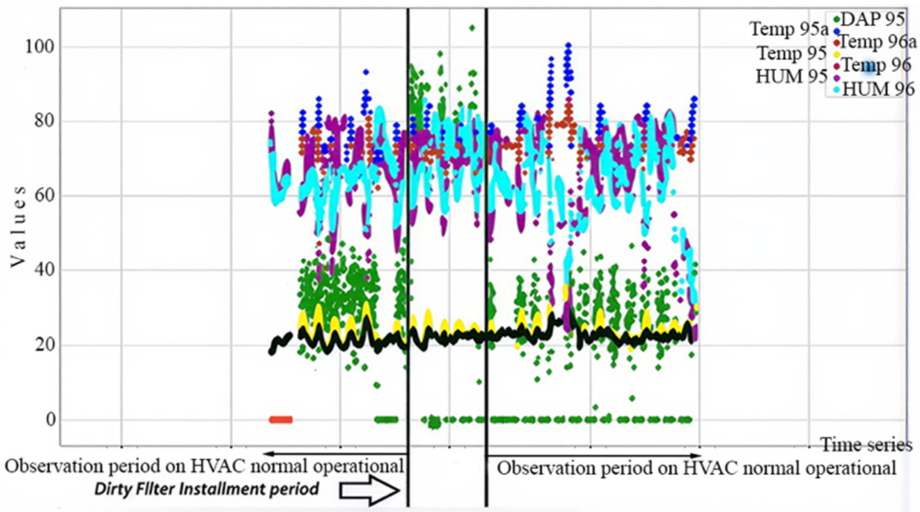

Fig. 5. Plot of data acquisition results from the sensor on an AF.

Fig. 6. Sensor network architecture used in the case study.

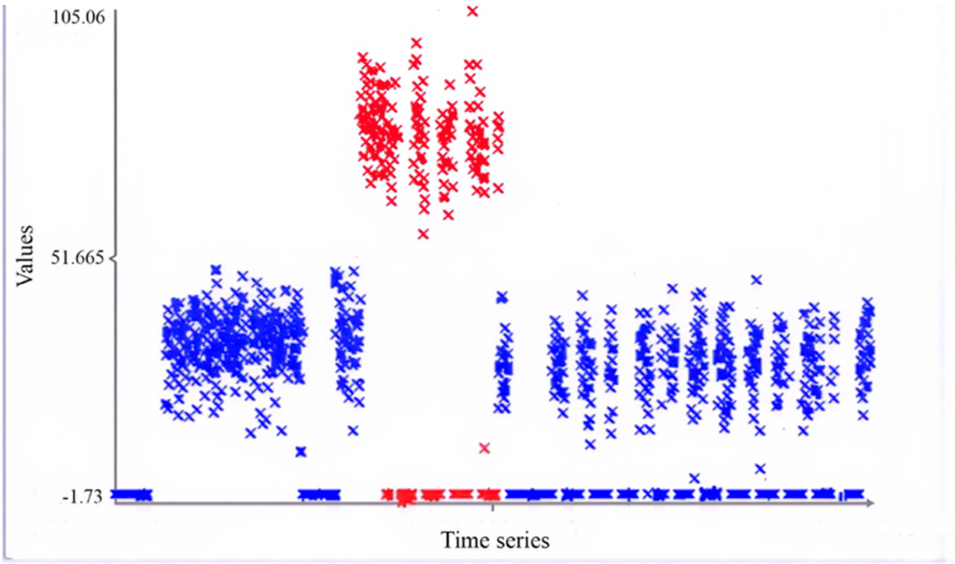

Fig. 7. Data labeling plot (red: filter in unusable condition; blue: filter in usable condition).

To train the SLR method, some significant hyperparameters were configured. The ridge value was set to 1.0^{-8}, and the conjugate gradient descent was set to false. Finally, the seed value was set to 1. Because the training dataset was imbalanced, the ridge estimator was applied to stabilize coefficient estimates, particularly for small samples in the minority class. The ML model testing results are outlined in Tables I and II. Table I provides cross-validation performance, with the average accuracy of 0.99 across all folds and a small Root Mean Square Error (RMSE) of 0.03, indicating no overfitting and a close match to the actual values. The results were also supported by additional evaluation parameters, such as precision, recall, and F1-score, as presented in Table II. The effect of an imbalanced training dataset can also be minimized, as indicated by the Precision-Recall Curve (PRC) areas of 1 and 0.997 for classes 0 and 1, respectively. Figure 8 displays the confusion matrix, suggesting that operational hours and air pressure differences were the most significant factors in recognizing AF conditions and performance.

TABLE I. RESULTS OF ML MODEL TESTING WITH SLR APPROACH TO DETECT AF CONDITION

Parameters	Values
Number of correct predictions	0.99
Number of incorrect predictions	0.01
Kappa statistic	0.99
Mean absolute error	0.0024
RMSE	0.03
Relative absolute error	0.75%
Root relative squared error	7.49%

Besides SLR, other approaches were also utilized to compare the performance of ML models. Using DT, the correct prediction rate only reached 0.92. Similarly, deploying the LR approach, the prediction accuracy also reached 0.92. The accuracy gaps between SLR and the other two methods can be attributed to several factors, including random shifts in the dataset during cross-validation splitting. Moreover, SLR tended to focus on the most important features as it utilized a single predictor.

TABLE II. DETAILED ACCURACY BY CLASS OF ML MODEL GENERATED USING SLR

Parameters	Class 0	Class 1	Weighted average
True positive rate	0.999	0.998	0.999
False positive rate	0.002	0.001	0.002
Precision	0.999	0.998	0.999
Recall	0.999	0.998	0.999
F1-score	0.999	0.998	0.999
MCC	0.997	0.997	0.997
Receiver Operating Characteristic (ROC) area	0.999	0.999	0.999
PRC area	1.000	0.997	0.999

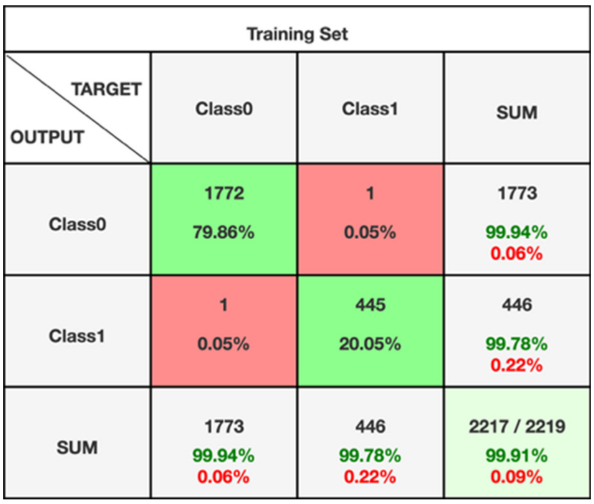

Fig. 8. Confusion matrix of the SLR model detecting AF conditions.

C. Inter-Component Reliability Model

An inter-component reliability model is used to calculate the probability of another component's condition, given that one component meets the standard requirements. As shown in Figure 2, there are two primary cycles in the HVAC system at the case study site: the water cycle and the air cycle. The components of the water and air cycles are presented in Table III. The former has at least 7 components, while the latter has 9, and the potential malfunctions of each component can be identified. Component failures in the water supply cycle stem from water pump performance, as presented in Table IV. Water pump performance can be monitored using the water pressure indicator installed on each pump. In the heat exchanger component, reduced performance is often caused by insufficient cold-water supply, leading to reduced heat transfer. Therefore, it is necessary to continuously monitor the water temperature and adjust the flow rates of hot and cold water to the heat exchanger.

Another problem in the HVAC system lies in the air supply cycle. The main component in this cycle is the AHU. AHU's performance is greatly influenced by three coils—Preheating Coil (PC), Reheating Coil (RC), and Cooling Coil (CC)—which are also highly dependent on the performance of the Heat exchanger circuit 1 (H1) and Heat exchanger circuit 2 (H2) components (heat exchangers). In the air cycle, no mechanism was found to monitor the air supply to the room through the distribution duct. Therefore, maintenance involves periodically replacing the duct AFs (every 3 months). At the case study site, there are over 90 AFs of various types and sizes. Sometimes, AFs that are still in good condition are still replaced during routine maintenance. This will undoubtedly increase the maintenance costs. The problems related to the air supply cycle are summarized in Table V.

The reliability relationships between components can be defined using the identification of potential problems in the components of the HVAC system. However, to reduce complexity, the scope for demonstrating component reliability modeling of the AP-SPM system is narrowed to the air supply cycle, specifically problem ASF-PRB2, as presented in Table V. This relationship is modeled using a BN, as shown in Figure 9. Each component is assumed to be a node in the BN. Each node has two states: true indicates the presence of a problem, while false indicates otherwise.

TABLE III. COMPONENTS OF THE HVAC SYSTEM

Cycle	Components	Symbol	Function
Water cycle	Sea water pump 1	P1	Seawater intake
	Sand filter	SAF	Cleaning sand from seawater taken by P1
	Heat exchanger circuit 1	H1	Turning hot water into cold water
	Heat exchanger circuit 2	H2	Turning hot water into cold water
	Pump 2	P2	Distributing cold water
	Pump 3	P3	Distributing cold water
	Boiler	BL	Hot water supply
Air cycle	Preheat Coil	PC	Changing the air temperature from below the cold temperature to above the cold temperature
	Cooling coil	CC	Removing heat from the air to be distributed
	Reheat Coil	RC	By utilizing hot water, this coil increases the temperature of the air from outside to lower the humidity of the air that will be distributed to the room.
	Air supply fan	ASF	Pushing air into the air duct system
	Humidifier	HM	Maintain air humidity
	Air filter	AF	Filtering dust before the air is distributed to the room
	Inter-chiller	IC	Maintaining the temperature of the air distributed to the room
	Supply air diffuser	SAD	Components for air distribution
	Return air fan	RAF	Removing air from the room for circulation

TABLE IV. PROBLEMS IN THE WATER SUPPLY CYCLE

Main problem	Component code-problem symbol	Indicator
Reduced water suction power due to a blockage in the water inlet pipe	P1-PRB1	Indicated by the water pressure indicator
Reduced water suction power due to the need for pipe maintenance	P1-PRB2	Indicated by the water pressure indicator
The sand filter needs to be cleaned	SF-PRB1	Indicated by the water pressure indicator
Overheated potable water	H1-PRB1	Indicated by the temperature indicator
Insufficient seawater supply	H1-PRB2	Indicated by the water pressure indicator
Insufficient portable water supply	H1-PRB3	Indicated by the water pressure indicator
Decreasing heat transfer performance	H1-PRB4	-
Insufficient cold-water supply	H2-PRB1	-
Not cold enough water, temperature coming out	H2-PRB2	-
Decreasing heat exchanger performance	H2-PRB3	-
Decreasing the water pump's pressure	P2-PRB1	Indicated by the water pressure indicator
Decreasing the water pump's pressure	P3-PRB1	Indicated by the water pressure indicator

Figure 10 illustrates the state of the BN if PAF = true. The model indicates that a problem with the AF had a 60% probability of impacting the Inter-Chiller (IC). Through backtracking, the model also indicates a 59% probability that the AF problem was caused by the coils (PC, RC, and CC, represented by the node PC-RC-CC), having reduced temperature regulation at the AHU air inlet.

Sensitivity analysis was used to validate the BN by measuring the effects of the changes in the model's probability parameters on the output probabilities. Two key parameters were considered: sensitivity and magnitude. Sensitivity indicates the extent to which an output variable depends on its inputs, while the magnitude quantifies that dependency. To measure these values, the observation node and target output should be determined. The observation node represents the component to be monitored, and its specific outcome is referred to as the target output.

TABLE V. PROBLEMS IN THE AIR SUPPLY CYCLE

Main problem	Component code-problem symbol	Indicator
The AHU's power to absorb air from outside is insufficient	ASF-PRB1	Air flow rate and temperature sensors
Air supply performance and air quality decline	ASF-PRB2	Air temperature sensor, air pressure
Temperature controller problem	PC-PRB1	Air temperature sensor
Air conditioner problem	CC-PRB1	Air temperature sensor
Failed to raise the air temperature	RC-PRB1	Air temperature sensor
Dirty air filter	AF-PRB1	-
Low air quality and air pressure, less cool	IC-PRB1	Air temperature sensor, air flow rate
The fan's pulling power to remove air from the room is reduced	RAF-PRB1	Air flow rate sensor

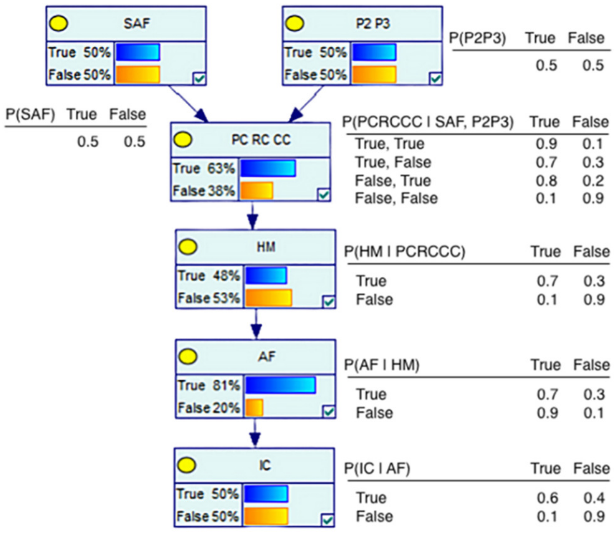

Fig. 9. Reliability model between components in the air supply cycle and their conditional probability tables.

The observation node is HM, with the target output being $P_{HM} = true$. In this regard, it is examined which nodes in the BN are the most sensitive to and have the greatest impact on the HM's performance degradation. The results of the sensitivity analysis are illustrated in Figure 11. It was revealed that the target output is the most sensitive to the poor condition of the three coils (PC + RC + CC, represented by the node PC-RC-CC), as indicated by the value of $P_{PCCCRC} = true$ in the AHU, with sensitivity and magnitude values of 0.625 and 0.0875, respectively. However, although the opposite condition of these coils ($P_{PCCCRC} = false$) has the second-highest sensitivity value, its magnitude is low (0.0075). This means that these coils have a negligible influence on the HM's performance decline. Furthermore, the sensitivity analysis on the HM node also shows that the other components that reduce HM performance are P2 and P3, with each having a magnitude of 0.27.

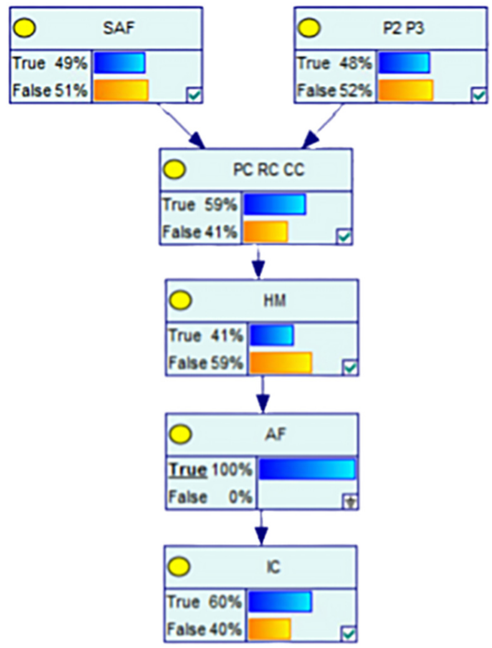

Fig. 10. State of the BN if PAF = true.

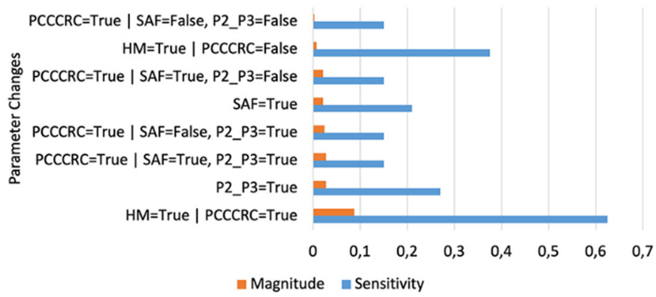

Fig. 11. Influential parameters generated by the sensitivity analysis test with the probability of interest $P_{HM} = true$.

Furthermore, the sensitivity analysis was conducted by changing the probability of interest (observation node) to the three coils having poor condition ($P_{PCCCRC} = true$). As depicted in Figure 12, the failure of two pumps (node P2P3) obtained sensitivity and magnitude values of 0.45 and 0.045, respectively. This means that node P2P3 was considered the most contributing component to the failure of the three coils, in line with the expert comments and data from the existing HVAC monitoring system. The pump's performance has a greater impact than SAF, as it plays a significant role in supplying the hot and cool water to coil components during the preheat, cooling, and reheat processes.

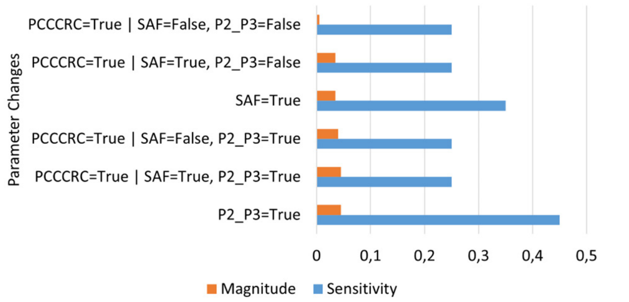

Fig. 12. Influential parameters generated by the sensitivity analysis test with the probability of interest $P_{PC\ CC\ RC} = true$.

D. Visualization

Figure 13 illustrates the AP-SPM system. Figure 13(a) shows the monitoring dashboard of the theater room condition at the study site. If the room icon is blue, the room's air quality meets the standard requirements, while red indicates otherwise. Figure 13(b) shows the sensor status in Cluster 1, which contains information on the measured variables, location, measurement results, and battery status.

E. Discussion

Based on the experimental results, the BN-based reliability model can be used to predict the factors contributing to the decline in component performance in the HVAC system. Although this reliability model is generally highly dependent on the root node's value as input, if there is additional evidence in the middle node (intermediate node) (for example, the case in Figure 9), the diagnosis function can point to both the parent node and the child node.

(a)

(b)

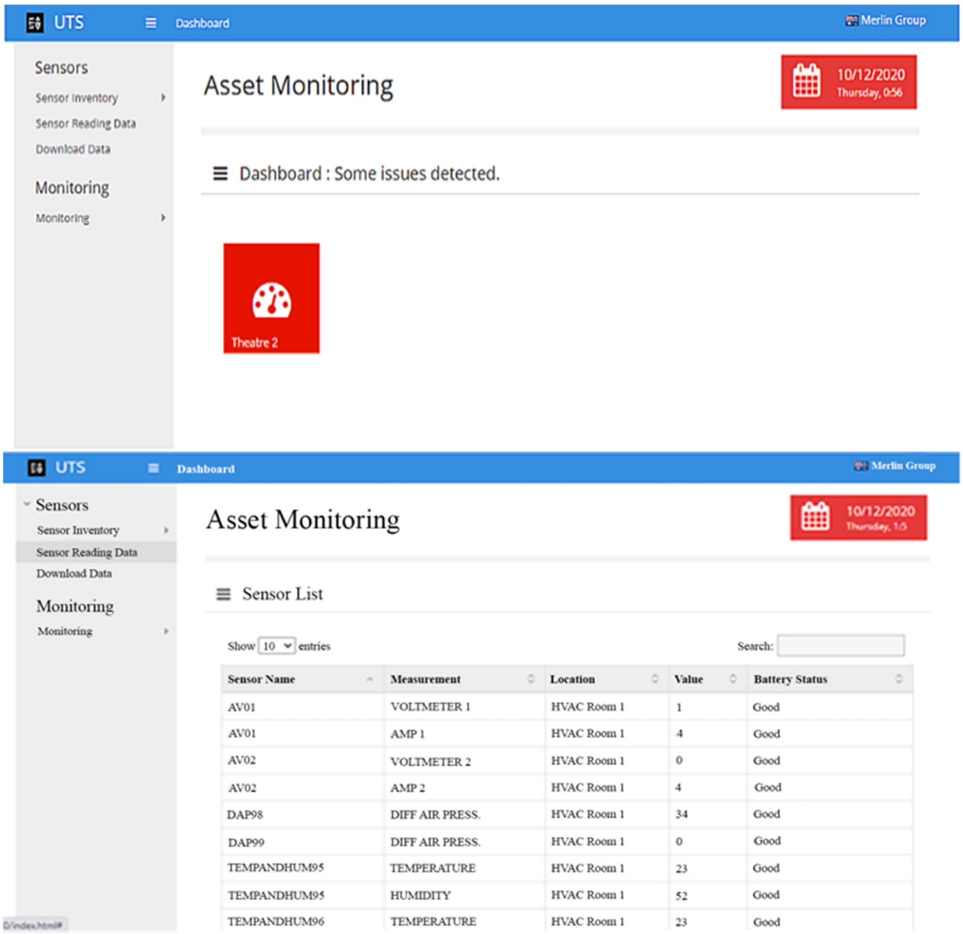

Fig. 13. AP-SPM system visualization: (a) dashboard, (b) sensor status.

This simplifies HVAC system operators' work in investigating component problems and can serve as a guide for junior technicians. Also, the presence of components physically located in areas that make sensor installation challenging due to space constraints should be considered. For example, an AF in a distribution duct is in a narrow space that does not accommodate a sensor, posing a challenge. Therefore, this reliability model must be created separately based on the existing room clusters.

IV. CONCLUSION

This study introduces the SPM Monitoring Tool (AP-SPM) for monitoring the performance of Heating, Ventilation, and

Air Conditioning (HVAC) components. The tool combines the Machine Learning (ML) and Causal Network (CN) methods, employing Simple Logistic Regression (SLR) as the ML model, which achieved a root mean square error (RMSE) of 0.03 and an accuracy of 0.99. For the CN component, a Bayesian network (BN) technique is implemented. The experimental results confirm the ML model's performance in identifying component conditions in the HVAC system. In addition, the developed reliability model demonstrates its ability to predict the probability of component deterioration when another component is identified as problematic. Even though the experiments used data from another HVAC monitoring system to validate the results, it is possible to integrate the AP-SPM system with the current HVAC system, especially at the database level.

However, the proposed AP-SPM system has some limitations, particularly in recognizing a reduction in HVAC component performance, since non-electrical components are involved and play a critical role in the development of the inter-component reliability model. In the long run, the model needs to be redesigned if the HVAC system architecture changes or to be implemented in other HVAC systems. Considering these limitations, future work should focus on developing ML models to detect problems in non-electrical components and conducting broader tests, particularly to relate the inter-component reliability model for energy use optimization.

REFERENCES

[1] A. Koons-Stapf, "Condition Based Maintenance: Theory, Methodology, & Application," in *Conference: Reliability and Maintainability Symposium*, Tarpon Springs, FL, USA, Jan. 2015, pp. 1–23.

[2] Y. Yu, D. Woradechjumroen, and D. Yu, "A Review of Fault Detection and Diagnosis Methodologies on Air-Handling Units," *Energy and Buildings*, vol. 82, pp. 550–562, Oct. 2014, https://doi.org/10.1016/j.enbuild.2014.06.042.

[3] M. M. Tehrani, Y. Beauregard, M. Rioux, J. P. Kenne, and R. Ouellet, "A Predictive Preference Model for Maintenance of a Heating Ventilating and Air Conditioning System," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 130–135, 2015, https://doi.org/10.1016/j.ifacol.2015.06.070.

[4] I. Sittón-Candanedo, E. Hernández-Nieves, S. Rodriguez-González, F. de la Prieta-Pintado, and J. M. C. Rodríguez, "Fault Predictive Model for HVAC Systems in the Context of Industry 4.0," in *XVIII Conferencia De La Asociacion Espa ´ Nola Para La Inteligencia Artificial*, Granada, Spain, 2018, Art. no. 10.

[5] D. Agustina, F. Fitri, Zilrahmi, R. S. Winanda, and D. P. Sari, "Utilization of Machine Learning for Predictive Maintenance in Improving Productivity in Manufacturing Industry," presented at the Proceedings of the 38th International Conference of the Polymer Processing Society, St. Gallen, Switzerland, 2024, Art. no. 060001, https://doi.org/10.1063/5.0224329.

[6] C. Tsallis, P. Papageorgas, D. Piromalis, and R. A. Munteanu, "Application-Wise Review of Machine Learning-Based Predictive Maintenance: Trends, Challenges, and Future Directions," *Applied Sciences*, vol. 15, no. 9, Apr. 2025, Art. no. 4898, https://doi.org/10.3390/app15094898.

[7] S. Lu *et al.*, "Exploring the Comprehensive Integration of Artificial Intelligence in Optimizing HVAC System Operations: A Review and Future Outlook," *Results in Engineering*, vol. 25, Mar. 2025, Art. no. 103765, https://doi.org/10.1016/j.rineng.2024.103765.

[8] I. Hector and R. Panjanathan, "Predictive Maintenance in Industry 4.0: A Survey of Planning Models and Machine Learning Techniques," *PeerJ Computer Science*, vol. 10, May 2024, Art. no. e2016, https://doi.org/10.7717/peerj-cs.2016.

[9] I. Matetić, I. Štajduhar, I. Wolf, and S. Ljubic, "A Review of Data-Driven Approaches and Techniques for Fault Detection and Diagnosis in HVAC Systems," *Sensors*, vol. 23, no. 1, Dec. 2022, Art. no. 1, https://doi.org/10.3390/s23010001.

[10] Z. Chen *et al.*, "A Review of Data-Driven Fault Detection and Diagnostics for Building HVAC Systems," *Applied Energy*, vol. 339, Jun. 2023, Art. no. 121030, https://doi.org/10.1016/j.apenergy.2023.121030.

[11] S. L. Zhou, A. A. Shah, P. K. Leung, X. Zhu, and Q. Liao, "A Comprehensive Review of the Applications of Machine Learning for HVAC," *DeCarbon*, vol. 2, Sept. 2023, Art. no. 100023, https://doi.org/10.1016/j.decarb.2023.100023.

[12] Y. Guo, Y. Liu, Y. Wang, Z. Wang, Z. Zhang, and P. Xue, "Advance and Prospect of Machine Learning Based Fault Detection and Diagnosis in Air Conditioning Systems," *Renewable and Sustainable Energy Reviews*, vol. 205, Nov. 2024, Art. no. 114853, https://doi.org/10.1016/j.rser.2024.114853.

[13] F. Dadras Javan, I. A. Campodonico Avendano, B. Najafi, A. Moazami, and F. Rinaldi, "Machine-Learning-Based Prediction of HVAC-Driven Load Flexibility in Warehouses," *Energies*, vol. 16, no. 14, Jul. 2023, Art. no. 5407, https://doi.org/10.3390/en16145407.

[14] Y. Guo, J. Wall, J. Li, and S. West, "Intelligent Model-Based Fault Detection and Diagnosis for HVAC System Using Statistical Machine Learning Methods," in *ASHRAE 2013 Winter Conference*, Dallas, TX, USA, Jan. 2013, vol. 119, no. 1.

[15] Z. Li *et al.*, "Machine Learning Based Diagnosis Strategy for Refrigerant Charge Amount Malfunction of Variable Refrigerant Flow System," *International Journal of Refrigeration*, vol. 110, pp. 95–105, Feb. 2020, https://doi.org/10.1016/j.ijrefrig.2019.10.026.

[16] N. Es-sakali, Z. Zoubir, S. Idrissi Kaitouni, M. O. Mghazli, M. Cherkaoui, and J. Pfafferott, "Advanced Predictive Maintenance and Fault Diagnosis Strategy for Enhanced HVAC Efficiency in Buildings," *Applied Thermal Engineering*, vol. 254, Oct. 2024, Art. no. 123910, https://doi.org/10.1016/j.applthermaleng.2024.123910.

[17] D. Borda, M. Bergagio, M. Amerio, M. C. Masoero, R. Borchiellini, and D. Papurello, "Development of Anomaly Detectors for HVAC Systems Using Machine Learning," *Processes*, vol. 11, no. 2, Feb. 2023, Art. no. 535, https://doi.org/10.3390/pr11020535.

[18] S. Gharsellaoui, M. Mansouri, M. Trabelsi, M.-F. Harkat, S. S. Refaat, and H. Messaoud, "Interval-Valued Features Based Machine Learning Technique for Fault Detection and Diagnosis of Uncertain HVAC Systems," *IEEE Access*, vol. 8, pp. 171892–171902, 2020, https://doi.org/10.1109/ACCESS.2020.3019365.

[19] A. Ranade, G. Provan, A. El-Din Mady, and D. O'Sullivan, "A Computationally Efficient Method for Fault Diagnosis of Fan-Coil Unit Terminals in Building Heating Ventilation and Air Conditioning Systems," *Journal of Building Engineering*, vol. 27, Jan. 2020, Art. no. 100955, https://doi.org/10.1016/j.jobe.2019.100955.

[20] Y. Gao, C. Ma, and T. Wang, "Fault Diagnosis for Cooling Dehumidifier Based on Fuzzy Classifier Optimized by Adaptive Genetic Algorithm," *Heliyon*, vol. 8, no. 12, Dec. 2022, Art. no. e12057, https://doi.org/10.1016/j.heliyon.2022.e12057.

[21] Z. Liu, Z. Huang, J. Wang, C. Yue, and S. Yoon, "A Novel Fault Diagnosis and Self-Calibration Method for Air-Handling Units Using Bayesian Inference and Virtual Sensing," *Energy and Buildings*, vol. 250, Nov. 2021, Art. no. 111293, https://doi.org/10.1016/j.enbuild.2021.111293.

[22] K. H. Ng, F. W. H. Yik, P. Lee, K. K. Y. Lee, and D. C. H. Chan, "Bayesian Method for HVAC Plant Sensor Fault Detection and Diagnosis," *Energy and Buildings*, vol. 228, Dec. 2020, Art. no. 110476, https://doi.org/10.1016/j.enbuild.2020.110476.

[23] A. Gálvez, A. Diez-Olivan, D. Seneviratne, and D. Galar, "Fault Detection and RUL Estimation for Railway HVAC Systems Using a Hybrid Model-Based Approach," *Sustainability*, vol. 13, no. 12, Jun. 2021, Art. no. 6828, https://doi.org/10.3390/su13126828.

[24] K. Verbert, R. Babuška, and B. De Schutter, "Combining Knowledge and Historical Data for System-level Fault Diagnosis of HVAC Systems," *Engineering Applications of Artificial Intelligence*, vol. 59, pp. 260–273, Mar. 2017, https://doi.org/10.1016/j.engappai.2016.12.021.

[25] H. Shahnazari, P. Mhaskar, J. M. House, and T. I. Salsbury, "Modeling and Fault Diagnosis Design for HVAC Systems Using Recurrent Neural Networks," *Computers & Chemical Engineering*, vol. 126, pp. 189–203, Jul. 2019, https://doi.org/10.1016/j.compchemeng.2019.04.011.

[26] S. Fernandes, M. Antunes, A. R. Santiago, J. P. Barraca, D. Gomes, and R. L. Aguiar, "Forecasting Appliances Failures: A Machine-Learning Approach to Predictive Maintenance," *Information*, vol. 11, no. 4, Apr. 2020, Art. no. 208, https://doi.org/10.3390/info11040208.

[27] A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, M. Rampazzo, and F. Simmini, "Data-Driven Fault Detection and Diagnosis for HVAC Water Chillers," *Control Engineering Practice*, vol. 53, pp. 79–91, Aug. 2016, https://doi.org/10.1016/j.conengprac.2016.04.018.

[28] S. A. Aghili, M. Khanzadi, and Z. M. Vaghefi, "A Data-driven Based Framework for Resilience Assessment of Hospital HVAC Systems," *Journal of Building Engineering*, vol. 112, Oct. 2025, Art. no. 113656, https://doi.org/10.1016/j.jobe.2025.113656.

[29] N. Hajimirza Amin, A. Etemad, and A. Abdalisousan, "Data-Driven Performance Analysis of an Active Chilled Beam Air Conditioning System: A Machine Learning Approach for Energy Efficiency and Predictive Maintenance," *Results in Engineering*, vol. 23, Sept. 2024, Art. no. 102747, https://doi.org/10.1016/j.rineng.2024.102747.

[30] M. Ray, P. Samal, and C. K. Panigrahi, "Implementation of a Hybrid Technique for the Predictive Control of the Residential Heating Ventilation and Air Conditioning Systems," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8772–8776, Jun. 2022, https://doi.org/10.48084/etasr.5027.

[31] M. Kliangkhlao, K. Haruehansapong, K. Yeranee, and B. Sahoh, "Causal Artificial Intelligence–Driven Approach for HVAC Preventive Maintenance Explanation," *IEEE Access*, vol. 12, pp. 121064–121076, 2024, https://doi.org/10.1109/ACCESS.2024.3452416.

[32] C. Abay and H. Apaydın Özkan, "Artificial Neural Network Based Fault Detection and Classification Method for Air Conditioners," *Eskişehir Technical University Journal of Science and Technology A - Applied Sciences and Engineering*, vol. 25, no. 2, pp. 240–249, Jun. 2024, https://doi.org/10.18038/estubtda.1379759.

[33] D. Das, N. Bhattacharjee, and D. Boro, "Real-Time IoT Data Analysis for HVAC System Maintenance," *Engineering Research Express*, vol. 7, no. 3, Sept. 2025, Art. no. 0352b4, https://doi.org/10.1088/2631-8695/ae0107.