Automated Fault Detection and Diagnosis in Light Commercial Building's HVAC Systems

Milad Babadi Soltanzadeh

A Thesis

In the Department of

Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Master of Applied Science (Building Engineering) at

Concordia University

Montréal, Québec, Canada

August 2024

© Milad Babadi Soltanzadeh

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Milad Babadi Soultanzadeh**

Entitled: **Automated Fault Detection and Diagnosis in Light Commercial Building's HVAC systems.**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Building Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Radu Zmeureanu_____ Chair

Dr. Radu Zmeureanu _____ Examiner

Dr. Fuzhan Nasiri _____ Examiner

Dr. Mohamed Ouf, Dr. Mazdak Nik-Bakht_____Co-Supervisors

Approved by:

Dr. Mohamed Ouf_____Graduate Program Director

Dr. Mourad Debbabi_____Dean of Faculty

2024-08-21

ABSTRACT

Automated Fault Detection and Diagnosis in Light Commercial Building's HVAC systems.

Milad Babadi Soultanzadeh

Fault detection and diagnosis (FDD) in commercial buildings' HVAC systems can significantly reduce energy consumption. Faults in these systems occur due to various issues such as aging and inadequate maintenance. Commercial buildings in Canada covered an area of 709,029,612 m² in 2019, consuming 948,216,746 GJ of energy. HVAC systems are responsible for 25%-50% of this energy consumption. In the United States, faults in HVAC systems contribute to an additional energy consumption of 103 to 500 terawatt-hours (TWh) in the building sector. Detecting and diagnosing faults in HVAC systems can reduce energy consumption by 20% to 30%. Light commercial buildings, defined as commercial buildings with fewer than six stories and less than 2500 square feet, include bank branches, offices, and small industrial facilities. These buildings have similarities in the configuration and size of HVAC components, making it feasible to develop an FDD tailored for this class of buildings, that can be easily scaled up. This goal can be achieved using data-driven methods, which have gained popularity over the past decades by installing various sensors and collecting data integrated with Building Energy Management (BEM) systems. In this thesis, three different FDD methods have been developed and validated on light commercial buildings. The first method is a semi-supervised method that includes various techniques to handle the unlabeled raw data from BEMs, resulting in a final supervised Automatic Fault Detection (AFDD) system. The second method is a fully unsupervised novel AFDD method based on PCA time series fault detection. The third method is primarily based on the inverse model of the Air Handling Unit (AHU) of the HVAC systems. A typical light commercial building in Montreal, Canada, was used for all methods. Additionally, to validate the generalizability of the unsupervised method, another light commercial building, a small industrial facility in Ireland, was used as well.

The first method successfully resulted in an AFDD that can detect and diagnose faults with almost 90% accuracy, performing better in condition-based faults than control faults. The unsupervised method showed very good results in terms of generalizability. It was able to detect faults and report the problematic inputs and locations to the HVAC operators. Although the unsupervised method cannot completely diagnose condition-based faults, it provides very good information based on the system's behavior, enabling operators to diagnose the faults. Finally, the inverse modeling revealed that a physics-based neural network can outperform neural networks and genetic algorithms in modeling the system inversely and detecting anomalies mostly related to energy consumption.

Acknowledgments

I would like to express my deepest appreciation to my supervisors, Dr. Mohamed Ouf and Dr. Mazdak Nik-Bakht, who have enlightened me on my path throughout this journey. Their guidance and support have been invaluable.

I also extend my heartfelt thanks to the team at Strato Automation for providing essential data and consultation during my research on Fault Detection. Their contributions have been crucial to the success of this work.

My warmest thanks go to my wife, Forough, for her unwavering support during my career as an MSc student. Her encouragement and understanding have been my strength.

Lastly, I owe everything to my mother, Touran, whose sacrifices have made this journey possible. Her love and dedication have been my greatest source of inspiration.

Preface

This thesis has been extracted from four different paper manuscripts under the supervision of Dr. Mohamed Ouf and Dr. Mazdak Nik-Bakht for graduation in the Master of Applied Science in Building Engineering. The introduction and literature review are combined and edited from all papers. Chapters 3, 4, and 5 are structured as follows: Chapter 3 is based on papers [1, 2], Chapter 4 is based on paper [3], and Chapter 5 is based on paper [4].

Papers:

1. M. Babadi Soultanzadeh, M. M. Ouf, M. Nik-Bakht, P. Paquette, and S. Lupien, "Fault detection and diagnosis in light commercial buildings' HVAC systems: A comprehensive framework, application, and performance evaluation," Energy Build, vol. 316, p. 114341, Aug. 2024, doi: 10.1016/j.enbuild.2024.114341.

2. M. Babadi Soultanzadeh, M. Ouf, N.-B. Nik-Bakht, P. Paquette, and S. Lupien, "A Framework for Automated Fault Detection in Light Commercial Buildings HVAC System," in ASHRAE Transactions 130, 2024, pp. 590–599.

3. Unsupervised Automated Fault Detection and Diagnosis for Light Commercial Buildings' HVAC systems, has been submitted to the Journal of "Building and Environment".

4. Physics-based inverse model anomaly detection in light commercial buildings' AHU systems, Presented at IBPC 2024, July 26, 2024, Toronto.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1: Introduction

1.1. Background

Due to physical damage, aging, poor quality, and inadequate maintenance of Heating, Ventilation, and Air Conditioning (HVAC) parts, they can experience hard or soft faults (partial or complete failure). While quantifying the energy losses is tough, addressing these issues in commercial buildings could save energy consumption by up to 20-30% [1]. Taking advantage of the commonality of HVAC systems, as well as the similar size and dimensions of light commercial buildings, leads to the development of a fault detection algorithm that is tailored specifically for this type of building. A light commercial building is generally defined as no more than six stories with less than 2,500 square meters of floor area, for example, a small office building, a medical facility, a bank, or a small manufacturing facility. Historical data collected from BEMS can be utilized for automatic fault detection and diagnosis [2]. The occurrence of any fault has the potential to introduce anomalies within historical data records. In other words, when a fault happens, it can disrupt the patterns and trends observed in historical data, making the data deviate from its expected behavior [3]. While fault detection methods might achieve a satisfactory level of precision by identifying evident alterations in the system's patterns, fault diagnosis presents a notably challenging aspect. This difficulty arises from the fact that various faults have the potential to generate identical symptoms. These symptoms can manifest as anomaly shifts in sensor readings, or even as intricate combinations of alterations in multiple sensor values. In essence, the complexity of diagnosing faults lies in distinguishing between distinct underlying issues that might exhibit similar observable effects [4]. Faults in HVAC systems can generally be classified into three categories based on their nature [5]:

- Condition-based faults: These refer to physical conditions that deviate from the norm, such as a stuck valve, fouled coil, or broken fan.

- Behavior-based faults: These involve undesirable system behaviors, like simultaneous heating and cooling, erratic fan speeds, or improper cycling.
- Outcome-based faults: These are identified through performance metrics, such as a reduced coefficient of performance (COP), increased hot water flow, or higher energy consumption than expected.

Each type of fault requires distinct detection and diagnosis approaches to ensure effective system maintenance and optimization.

Another classification of faults in HVAC systems, particularly in Air Handling Units (AHUs), can be as four groups of faults based on their origins: 1) Sensor faults, 2) Hardware (Component) faults, 3) Software (Programming and Controlling) faults, and 4) Communication faults [6,7]. Sensor faults typically involve biases and drifting of measuring values, which can lead to inaccurate readings and misinterpretations of system performance. Hardware faults relate to the failure or malfunction of system components, such as coils fouling, stuck valves, stuck dampers, air duct leakage, valve leakage, and fan failure, all of which can severely impact system functionality and efficiency. Programming faults, categorized as soft faults, can persist in the system for extended periods and diminish system performance, often stemming from errors in programming logic and sequence of operations[8]. Communication faults arise due to protocol incompliance and issues with data collection and storage, posing challenges in maintaining seamless communication between system components and control systems [9]. There are only a few studies specifically focused on soft faults related to programming, controlling, and sequence of operation [10].

1.2. FDD Categorization

Broadly speaking, Fault Detection and Diagnosis (FDD) techniques can be categorized into three main groups: model-based, rules-based, and data-driven methods [11]. In recent decades, numerous researchers have delved into various methodologies within these categories, conducting thorough

investigations and drawing comparisons among their outcomes. These studies have been conducted for both the primary and secondary components of HVAC systems, aiming to enhance the understanding and implementation of effective FDD strategies [12]. The evolution of machine learning techniques, monitoring systems, and the diminishing reliance on explicit models have led to the burgeoning popularity of data-driven methods[13]. A significant advantage of data-driven methods is their independence from human expertise or physical models, as they are based solely on real-world operational data collected from the system. Data-driven methods offer significant advantages in speed and scalability for AFDD systems. Unlike traditional models, they rapidly analyze large datasets, reducing customization time. This enables quick deployment across HVAC systems without specialized model development, facilitating widespread adoption. these methods draw insights directly from the behavior of the system itself. In recent years, data-driven approaches have gained significant traction among researchers due to the increasingly sophisticated technologies and analytical tools used to collect and analyze data [14]. Data-driven methods can be subdivided into two main categories: supervised and unsupervised methods. On the one hand, supervised methods are employed when a dataset with labeled information is accessible, enabling the algorithm to learn from these labels. A dataset that has been labeled refers to a dataset in which all instances are assigned a specific label indicating whether there are faults, fault types, and fault severity. Unsupervised methods, on the other hand, are mostly used for datasets lacking labels and aim to extract patterns and relationships from data without the need for predefined labels [2,11,15]. However, there exist certain semi-supervised methods that attempt to assign labels to unlabeled data by utilizing a small subset of labeled data [14]. An important point to highlight is that data-driven methods frequently adopt a supervised approach [13]. One main problem with data-driven methods, especially the supervised ones, is that when they become trained on the data from a specific system, it is hard to use them on another similar system without adjustments [16]. With the increasing sophistication of data collection technologies and analytical tools, data-driven approaches have gained significant traction among researchers [14]. Their appeal lies in their ability to adapt to the complexities of HVAC systems without the need for prior assumptions or extensive manual intervention, making them particularly valuable in modern buildings' Fault

Detection and Diagnosis (FDD) [2]. A labeled dataset in FDD indicates normal conditions as well as various types of faults with different severity levels. These datasets are scarce in real buildings and are primarily utilized in theoretical fault detection and diagnosis algorithms using experimental or simulated data [14].

1.3. Problem Statement

Many studies in the existing literature rely on experimental or simulated data obtained from laboratory tests, such as ASHRAE RP-1312 [17], or publicly accessible datasets provided by institutions like the Lawrence Berkeley National Laboratory [18]. While these datasets serve the critical purpose of offering benchmark data in a field with limited publicly available resources, they do not accurately reflect the complexities and variations found in real-world systems [19]. Huang et al. discovered that FDD strategies developed using simulated data and applied directly to real building data yield suboptimal results. In practical applications, the data granularity is inadequate, metadata descriptions are lacking, and there is a scarcity of well-labeled faults with varying severity ratings [20]. Therefore, automated fault detection in Building Automation Systems (BAS) presents a significant challenge when dealing with extensive amounts of unlabeled data. The vast size of the datasets, coupled with the presence of numerous missing values and inconsistencies, poses significant obstacles for fault detection algorithms. In such scenarios, the need for expertise from both system professionals and data mining specialists becomes paramount. Considering all factors, it's evident that a generic Automatic Fault Detection and Diagnosis (AFDD) framework, transferable and applicable across different buildings within the same class, is vital, especially for commercial buildings.

1.4. Goal and Objectives

The main goal of this research is to develop generalizable AFDD methods tailored for light commercial buildings. These methods aim to be broadly applicable across various buildings within this class, capable of utilizing different types of historical data from BEMS systems, even if the

data is unlabeled, inconsistent, or fragmented. Three different methods have been developed using various machine-learning techniques. The main objectives of this thesis are as follows:

1) Develop a semi-supervised AFDD method for light commercial buildings' HVAC systems.
2) Develop an unsupervised AFDD method for light commercial buildings' HVAC systems.
3) Develop a physics-based neural network inverse model to detect anomalies in the AHUs of light commercial buildings.

The first objective involves starting with raw, unlabeled data from BEMS systems to identify potential faults using a combination of machine learning techniques. This method aims to find patterns leading to anomalies and isolate patterns indicating faults with the confirmation of HVAC operators. The labeled data based on these patterns is then used to finalize the AFDD model through supervised classification. The second objective is designed to work with minimal initial information about the system and to find the most generalizable AFDD method applicable to a wide range of light commercial building data without extensive acquisition. An unsupervised method has been selected, with innovative modifications to enhance its generalizability. The third objective focuses specifically on the AHUs (central systems) of light commercial buildings. This method combines data with physical models to detect anomalies in the energy consumption of AHUs.

Each objective of this thesis has been introduced and investigated in separate chapters. Chapter 3 addresses research objective #1, Chapter 4 covers research objective #2, and Chapter 5 delves into research objective #3.

This thesis is composed of seven chapters. The first chapter introduces the study and establishes its main goals. The second chapter reviews the relevant literature and identifies gaps in previous works. Three different techniques (Objectives) for AFDD are then introduced, validated, and discussed in three separate chapters. Finally, a comprehensive conclusion section summarizes the findings and implications of all three methods.

Chapter 2: Literature Review

To conduct a comprehensive investigation of important previous works, the SCOPUS API was utilized. A query was performed using specific keywords: "AFDD," "HVAC," and "Automatic Fault Detection", for the general data-driven methods. The most relevant papers identified from the search results were manually selected for review. All the papers chosen for this review were published after the year 2000. It is followed by a focus on different datasets used for FDD and, different metrics that have been employed in previous studies, and then concludes with the literature gaps.

2.1. Data-Driven techniques in HVAC AFDD

Since PCA is the main technique used in the Unsupervised method and plays a crucial role in the Comprehensive AFDD framework, a separate subsection has been dedicated to reviewing the PCA literature and another subsection for other data-driven methods.

2.1.1. Data-Driven and ML techniques in HAVC FDD

Clustering is an additional machine-learning tool that holds the potential to achieve high accuracy in the realm of fault detection and diagnosis [21]. Miller et al [22] utilized clustering techniques to establish a multi-variable fault detection system in a residential building. Li et al [23]. employed K-Means clustering following Linear Discriminant Analysis to identify and isolate faults in a chiller system. Li et al [24]. employed density-based clustering in conjunction with PCA for Automated Fault Detection and Diagnosis (AFDD) related to sensor faults in a chiller system. Furthermore, various other researchers have explored different clustering methods for Automated Fault Detection and Diagnosis (AFDD), including Fan et al [25]., Dey et al [26,27]., Gunay et al [28]., Aguilar et al. [29]., and Xu et al [30].

The utilization of classification methods specific decision trees stands out as a valuable method that has found favor among numerous researchers [31]. Yan et al [32]. unveiled that decision tree

classifiers and regressors hold the capability to identify faults while deriving rules that simplify the diagnosis process. Moreover, their findings highlight that the accuracy of actual fault diagnosis using data-driven methods is significantly contingent on the quality of the training data. Issues such as the absence of distinct data patterns and errors in the training data can lead to unreliable classification outcomes. Li et al [33]. employed a virtual sensor-based decision tree to diagnose faults in VRF (Variable Refrigerant Flow) systems. They reported that this approach achieved an accurate detection and diagnosis rate of 69.18% for online data. Furthermore, their study indicated that enhancing the diagnostic methods necessitates the accumulation of more faulty data and training the model using an extensive repository of fault instances. Cappozoli et al [34]. employed predictive classifier and regressor decision trees to identify patterns in the abnormal energy consumption behavior of a smart building. Their study demonstrated that decision trees can encompass various aspects of the system to predict energy consumption and effectively detect anomalies in the process. Piscitelli et al [35]. showcased that tree regression can generate IF-THEN rules, a feature that many other data-driven methods such as artificial neural networks (ANN) lack. In recent years, other researchers have substantiated decision trees' utility as classifiers and regressors. The notable examples of this can be found in the works of Liu et al. [36,37], and Chiosa et al. [38].

Table 1. Summary of previous related research

Main Method	References	Dataset	HVAC System	Techniques	Limitations
Principal Component Analysis (PCA)	Wang S et al. (2005) [39]	Simulated data	AHU	PCA, SPE, T^2	Only sensor faults have been studied
	Du Z et al. (2007) [40]	Simulated data	AHU	PCA, SPE, Contribution plot, Joint plot	The diagnosis potential in HVAC system is limited because of the control loop propagation

	Du Z et al. (2008) [41]	Simulated data	AHU	PCA, Fisher discriminant analysis, SPE	Limited faults have been studied for specific system configurations.
	Hu Y et al. (2012) [42]	Operational data	Screw chiller	PCA, SPE	Only sensor faults have been studied
	Li S et al. (2014) [43]	ASHRAE RP-1312	AHU	PCA, Wavelet	Soft faults have not been studied; Limited fault conditions have been studied
	Li S et al. (2014) [44]	ASHRAE RP-1020 ASHRAE RP- 1312	AHU	PCA, SPE	Fault detection part has been done; the diagnosis part has not been studied
	Cotrufo N et al. (2016) [45]	University Campus data	Chiller	PCA, Ellipsoidal threshold	The diagnosis part can just isolate the variables and could not specify the fault type
	Hu Y et al. (2016) [46]	Operational data	Screw chiller	Preprocessing, PCA	Only sensor faults have been studied
	Guo Y et al. (2017) [47]	Operational data (3 days)	VRF	Preprocessing, PCA, SPE, Contribution plot	Only sensor faults have been studied, the limited period of data
	Li G et al. (2018) [24]	Operational data	Screw chiller	PCA, DBSCAN, Contribution plot	Only sensor faults have been studied
	Shi S et al. (2018) [48]	Experimental data	VRF	PCA, ANN	Only refrigeration amount has been studied
	Burgas L et al. (2021) [49]	Operational data from a university building	AHU	Unfolded-PCA, SPE, T^2	Only sensor faults have been studied
	Zhou Z et al. (2021) [50]	Operational data	VRF	PCA, SVM, CART, BPNN, SMOTE	Limited faults have been studied
Clustering					
	Du Z et al. (2014) [21]	Simulated data	AHU	ANN, Subtracting Clustering	Limited Faults have been studied, high level of FAR and MAR in some fault classes
	Fan C et al. (2015) [25]	Operational data from BAS	Building Energy system	Fuzzy-C-means, Hierarchical	Focused on Energy consumption and sensor

				clustering, k-means, PAM, EWKM, QARM, DT	of power consumption faults
	Dey M et al. (2018) [51]	Operational data from BEM	Fan-coil Unit (FCU)	k-means, SVM	Lack of accuracy in the diagnosis step, Limited faults of FCU have been studied
	Dey M et al. (2020) [26]	Operational data from BEM	Fan-coil Unit (FCU)	k-means, Linkage, Gaussian mixture	Limited faults of FCU have been studied
	Gunay B et al. (2020) [28]	Operational data from a large office	AHU, VAV terminal unit	k-means, Hierarchical clustering, Gaussian mixture	Limited to VAV faults, limited to detection step.
	Aguilar J et al. (2020) [29]	Operational data from a theater salon	Heat pump, Chiller, Cooling tower	k-means, DT, ANN	Limited to specific buildings,
	Xu Y et al. (2021) [30]	Operational data from BMS	Screw chiller	Hierarchical clustering, Partitioning clustering, ARM, FP-Growth	Limited to energy consumption fault.
Classification					
	Yan R et al. (2016) [32]	ASHRAE RP-1312	AHU	Classification and regression tree (CART)	The extracted rules for some faults are not valid
	Li G et al. (2018) [33]	Experimental Data	VRF	DT, Virtual sensor	Limited to specific faults
	Capozzoli A et al. (2018) [34]	Operational data building Energy Consumption	Building Energy system	SAX, CART	Limited to energy consumption faults
	Piscitelli M et al. (2021) [35]	Operational data from a university building	Building Energy system	Follow the leader clustering, ANN, DTR	Limited to energy consumption faults
	Liu X et al. (2021) [36]	Operational data from an office building	Building Energy system	DBSCAN, k-means, CART	Limited to energy consumption faults

Recent Works					
	Alghanmi A et al. (2023) [52]	Simulated data	RTU	PCA, Random Forest, ANN, SVM	Limited to the simulated faults.
	Liang A et al. (2023) [53]	ASHRAE RP-1043	Chiller	PCA, KPCA, k-means	Limited to the sensor faults.
	Zhao T et al. (2023) [54]	Experimental + Simulated data	AHU, Chiller, Cooling tower	PCA, LSTM, KNN	Limited to the specific sensor faults, low FDD efficiency
	Yang X et al. (2023) [55]	ASHRAE RP-1312	AHU	PCA, SPE	Limited to steady-state conditions, inaccuracy in weak faults.
	Fan C et al. (2024) [56]	ASHRAE RP-1312 ASHRAE RP-1043	AHU, Chiller	ANN	Limited faults have been studied. Simple architecture for ANN

Furthermore, Previous research in the field of HVAC systems has extensively utilized inverse models for various tasks, including energy modeling and fault detection. Torabi et al. took a different approach by using a physics-based inverse model to identify hard faults in AHU systems. They utilized supply temperature estimation as a virtual sensor, comparing it with recordings from actual sensors to detect faults effectively [57]. Similarly, Darwazeh et al. employed a model-based inverse model for virtual sensing of AHU systems to detect anomalies in the mixing box. Their method successfully identified return damper leakage, an issue that had gone unnoticed by building HVAC operators [58]. Furthermore, Gunay et al. employed an inverse model-based technique to detect programming logic faults in VAV AHU systems. Through their work, they successfully detected damper position and perimeter heater valve problems, contributing to enhanced system performance [59]. These studies collectively showcase the versatility and effectiveness of inverse modeling techniques in addressing various challenges within HVAC systems, from energy efficiency optimization to fault detection and diagnosis.

2.1.2. PCA in HVAC FDD

PCA is a multivariable statistical tool for analyzing process measurements, which can reveal how different variables change concerning each other and individually [60]. It is essentially a linear transformation that transforms correlated variables (dimensions) into orthogonal spaces, resulting in independent dimensions. These orthogonal dimensions are the eigenvectors of the covariance matrix of the original data, forming the columns of the loading matrix. The projection of sample points onto these eigenvectors can produce the reconstructed data samples when represented in the original dimensions. The number of eigenvectors used for this purpose represents the number of principal components (k) in the PCA method. A specific optimized value for k does not exist, and various methods can be used to determine the best option for the number of principal components, as detailed in the literature [61]. However, this section of the chapter will not delve deeply into the specifics of PCA and just focus on the necessary information related to the recent study.

The difference between the original data and the reconstructed data (projected on principal component directions) can be represented by Equation (1), where $\vec{e}$ is the vector of reconstruction error, $\vec{X}$ is the original sample vector, and $\hat{\vec{X}}$ is the reconstructed sample vector. Figure 1illustrates the process of calculating the projected error.

$$\vec{e} = \vec{X} - \hat{\vec{X}} \tag{1}$$

The reconstructed vector can be found using Equation (2), where P refers to the loading matrix with k number of eigenvectors.

$$\hat{\vec{X}} = \vec{X}PP^T \tag{2}$$

SPE will be calculated as the squared sum of the elements of the reconstructed errors, as shown in Equation (3):

$$SPE = \|\vec{e}\|^2 \tag{3}$$

This means that a sample vector with n dimensions, and hence an n-dimensional Reconstructed Error (RE) vector leads to a scalar SPE. Traditionally, in fault detection, a threshold is used for

SPE, as shown by equation (4) [62]. Where C_α is the normal deviation corresponding to the (1- α) percentile of SPEs higher than threshold and λ is the eigenvalue corresponding to the j^{th} eigenvector. Also, θ represents the first to third moments of eigenvalues from k+1 th to nth components of eigenvectors, and h_0, is a term that adjust the SPE distribution.

$$Threshold_{SPE} = \theta_1 \left[\frac{C_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_2} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}}$$

$$\theta_i = \sum_{j=k+1}^{n} \lambda_j^i, \quad i = 1,2,3 \tag{4}$$

$$h_0 = \frac{1 - 2\theta_1 \theta_3}{3\theta_2^2}$$

The following section will review previous literature on the FDD of HVAC systems. Since the foundational method of chapter 4 is PCA, the literature review will delve into a more detailed study of this method.

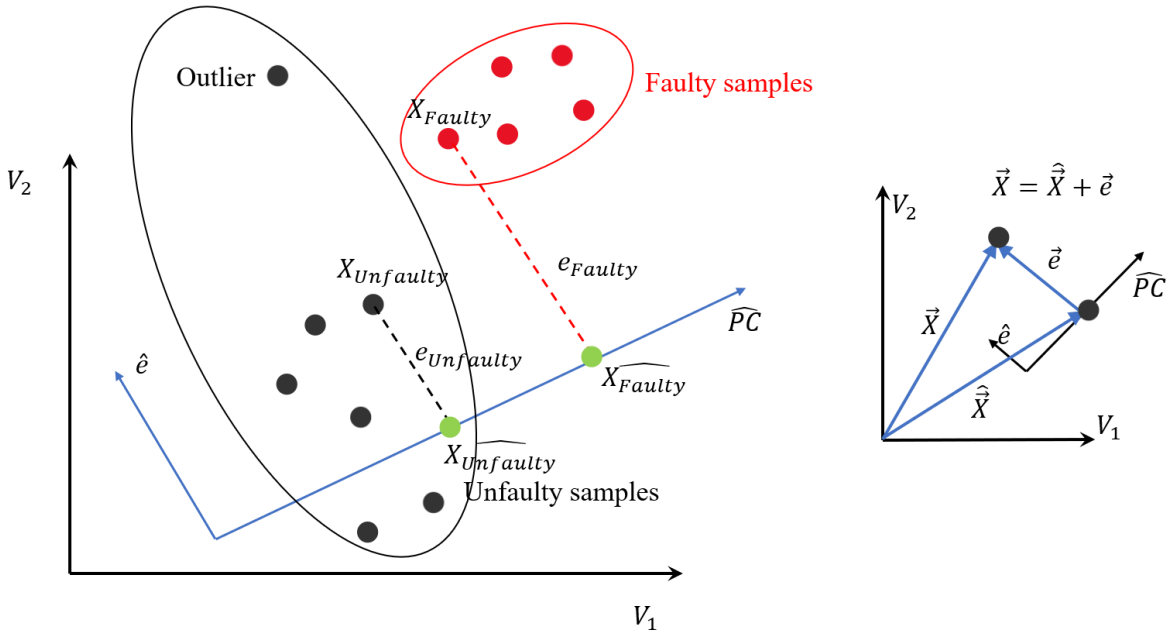


Figure 1. Concept of PCA-Based AFDD.

Among data-driven techniques, Principal Component Analysis (PCA) stands out as a highly favored method, particularly in the context of fault detection. PCA is a statistical technique and can be classified as a pattern recognition method [14]. Some of the earliest applications of PCA in HVAC FDD were reported by Wang S. and Xiao F. et al. [63–68]. These initial studies focused primarily on fault detection. They later extended their methods to the diagnosis of sensor faults in various HVAC systems, including Air Handling Units (AHU), Variable Air Volume (VAV) terminals, and Vapor Compression Systems (VCS). They utilized a variety of data sources for their studies, including simulated data, field data, and experimental data published by ASHRAE. Du et al [69]. combined various rules with PCA to detect and diagnose sensor faults in a Chiller + AHU system, experimenting with different numbers of principal components (PCs). Using a labeled dataset, they revealed that a robust FDD process for each sensor fault requires a specific number of PCs (cumulative variance). They expanded their research using simulated labeled data and a combination of multi-level PCA and Fisher Discriminant Analysis to enhance the detection and

diagnosis of sensor faults. They discovered that drifting sensor faults can persist in the system for a long time, so they simulated exaggerated drifting for the sensors to make these faults more detectable [70]. Xiao F. and Wang S. et al [71]. expanded their research on AFDD by using PCA with an expert-based multivariate decoupling method for sensor fault detection and isolation. They revealed that although PCA is powerful for detecting sensor faults in various engineering processes, it has significant weaknesses in the diagnosis phase. They found that the Q-contribution plot, commonly used for detection, is not effective even for sensor faults diagnosis in HVAC systems and they used expert-system-based diagnostics after detecting using PCA. Wang et al [72]. utilized normal operational data from an HVAC system, including AHUs, chillers, and cooling towers, and simulated fault scenarios to develop a system-level sensor fault detection and diagnosis method. They employed performance indicators for each level to enhance the detection and diagnosis process. The research on PCA-based AFDD of HVAC systems continued with the studies summarized in Table 2.

Table 2. PCA-based HAV system AFDD studies.

Study	Methodology	Dataset Used	Type of Faults Detected	Key Findings
Hu et al [62]. (2012)	Adaptive PCA	Simulated data	Sensor faults	Errors in normal data used for PCA training decrease FDD efficiency; PCA is less effective for negative faults.
Li et al [60]. (2014)	PCA-Wavelet model	ASHRAE RP-1312	Outdoor damper stuck, heating coil leaking	Data pretreatment is essential; improved detection of outdoor damper stuck vs. heating coil leaking faults.
Li et al [73]. (2014)	PCA with similarity analysis	ASHRAE RP-1312	Various faults	Enhanced PCA-based FDD with better performance using ASHRAE RP-1312 data.
Padilla et al [74]. (2015)	Passive-active sensor fault detection	Simulated and experimental data	Sensor faults	SPE outperformed T2 for experimental tests; both performed equally well with simulated data.
Yan et al. [75]. (2016)	PCA with k-distance and OPTICS	Simulated data	Sensor faults	Appropriate thresholds for specific faults and systems are crucial for efficient FDD.
Hu et al [76]. (2016)	PCA with preprocessing	Operational data	Sensor faults	Steady-state condition consideration and outlier removal (z-score) significantly improved FDD efficiency.

Guo et al [77,78]. (2017)	PCA with expert-based multivariable decoupling and SG method	Simulated data	System and sensor faults	FDD performance varied (59%-84%) for different faults; diagnosis performance >90% for all faults.
Li et al [79]. (2018)	PCA with DBSCAN	Operational data	Sensor faults	Clustering operational conditions and applying PCA separately enhanced FDD performance.
Montazeri et al [80]. (2020)	PCA, KPCA, and ANN	Simulated data, ASHRAE RP-1312	Sensor faults	PCA FDD performance was 60%, KPCA was 62%, ANN reached 98.7%.
Gu et al [81]. (2020)	Gaussian Mixture Model with PCA	Operational data	System faults	PCA for data dimensionality reduction significantly increased FDD performance.
Burgas et al [82]. (2021)	Unfolded-PCA	Operational data	Sensor faults, leakages	SPE was more effective for detecting sensor faults; Hotelling's T^2 was more efficient for identifying leakages.
Yang et al [83]. (2022)	PCA with thermal load matching	ASHRAE RP-1312	Sensor faults	FDD performance highly depends on fault severity; method for identifying best normal condition training set.
Liang et al [84]. (2023)	Hybrid clustering-isolation forest-PCA	ASHRAE RP-1043	Sensor faults	Well-chosen feature combination significantly improved sensor fault detection performance.
Wen et al [85]. (2023)	DBSCAN, SG smoothing, and PCA	Operational data	Temperature sensor faults	Enhanced sensor FDD for temperature sensors; applied to sensors not part of feedback regulation systems.
Yang et al [86]. (2023)	PCA with data window analysis	ASHRAE RP-1312	Sensor faults	Larger data window sample size enhances fault-free condition ratio by capturing more system dynamics.
Ma et al [87]. (2024)	BPNN-PCA	Simulated and operational data	Sensor faults	Improved fault diagnosis accuracy when using BPNN-PCA compared to BPNN alone.
Li et al [88]. (2024)	PCA and Bayesian inference	Field data, ASHRAE RP-1312	Sensor faults	Detectable faults in AHU system; combined PCA with Bayesian inference for HVAC sensor FDD and automatic calibration.

Regarding Table 2. PCA has established itself as a valuable tool in HVAC FDD, both as a standalone method and in combination with other techniques, or for dimensionality reduction. The reviewed literature indicates that PCA-based algorithms predominantly focus on sensor faults, particularly those not involved in feedback loops. A limited number of studies address component faults, often relying on labeled data for diagnosis or focusing solely on fault detection. PCA-based

algorithms are highly sensitive to the quality of training data, with outliers and inconsistencies significantly impacting performance. The SPE has proven to be more effective than Hotelling's T^2 for detecting and diagnosing sensor faults and malfunctions. For specific fault detection in particular systems, appropriate thresholds and cumulative covariance values must be applied. Furthermore, PCA is sensitive to dynamic and transient conditions, leading many researchers to use steady-state data for training or to implement smoothing techniques. PCA-based FDD algorithms are also more effective for faults of a certain severity, meaning that faults with lower severity might go undetected. Overall, while PCA-based methods are powerful, their effectiveness depends on careful data preparation and the application of tailored thresholds and techniques.

2.2. Operational, Experimental, and Simulated Data in HVAC AFDD

As shown in Table 1. the types of data utilized in previous studies can generally be categorized into three groups - simulated data, experimental data, and operational data. Typically, HVAC AFDD methods rely on operational data to develop techniques for vapor compression systems or overall energy consumption in buildings.

Typically, operational data is primarily utilized for fault detection, whereas experimental and simulated data with detailed labels are employed for both fault detection and diagnosis. Experimental data are predominantly employed for vapor compression systems and Air Handling Units (AHUs). Two widely recognized AFDD experimental datasets published by ASHRAE are titled "ASHRAE RP-1312"[17]and "ASHRAE RP-1043" [89] concerning AHUs and vapor compression systems, respectively. Additionally, other experimental tests for AFDD on vapor compression systems are more prevalent in the literature compared to AHUs, which are relatively rare occurrences.

Various fault scenarios have been emulated in the experimental datasets. These emulated scenarios can offer insights into different fault types, but they may not accurately reflect real-world conditions. For instance, consider the study conducted by Li et al.[43] on heating coil valve

leakage. In this study, the leakage fault was simulated by adjusting (overriding) the control valve signal to reduce the flow to a specific amount when the leakage was at 10%. However, it is highly unlikely for a system to experience such a precise severity of fault. In the event that this scenario was to occur, the control valve signal would likely be adjusted to a higher position to compensate for the leakage, driven by the control loop's aim to maintain the supply air setpoint, which completely deviates from the experiment condition. As a result, in experiments, the decrease in supply air temperature leads to a reduction in mixed air temperature. This relationship has been assumed as an expert rule in classification techniques [32]. However, relying solely on these assumptions can introduce inaccuracies when AFDD encounters real-world conditions.

Another challenge with experimental data is the information they offer. These experiments are often designed for specific purposes, resulting in differences in the number and location of measurements compared to real-world scenarios. For instance, consider the mixed air temperature sensor. Regarding the size and maintenance aspects, there are AHU systems in real-world settings that do not include a mixed air sensor[90], yet experimental setups may include this sensor and it is widely used in AFDD methods [91]. Lastly, experimental datasets often represent a specific HVAC configuration in terms of size, components, and control logic. For example, replacing water heating coils with electric heating coils, or water cooling coils with direct expansion coils, could significantly alter the features and inputs required for AFDD. While these datasets can be valuable for academic purposes, their limited applicability to generic AFDD purposes is evident due to their specificity to particular configurations.

2.3. FDD Metrics

Generally speaking, fault detection involves binary classification, distinguishing between two classes: Normal and Fault. Alternatively, fault diagnosis involves multiclass classification, differentiating between types of faults and their severity. However, different studies used different metrics for FDD based on the techniques they applied, which is always regarded as a challenge in the implementation of AFDD[6]. In previous studies, various metrics have been used to evaluate HVAC AFDD, including True Positive Rate (TPR), True Negative Rate (TNR), False Positive

Ratio (FPR), False Negative Rate (FNR), No Detection Rate (NDR), Correct Diagnosis Rate (CDR), and Misdiagnosis Rate (MDR) [14]. In addition to the metrics mentioned, other evaluation methods widely used for AFDD classification techniques include Confusion Matrix, Accuracy, Precision, F-Score, Recall, and ROC Curve. While each of these metrics gives some insight into AFDD performance, Shi. Z and O`Brien. W [92] suggests that FPR and FNR, along with Detection Time, are the most important. Ideally, an HVAC AFDD method should minimize these three metrics as much as possible, aiming for zero. However, achieving a Detection Time of zero is not feasible since a fault must occur before it can be detected. These metrics provide further insights into the performance and effectiveness of AFDD methods. Therefore, for a generalization purpose, in this study, with input from building operators and control system designers, FPR and FNR, also known as False Alarm Ratio (FAR) and Missing Alarm Ratio (MAR), respectively, along with Detection Time, have been selected for evaluation.

2.4. Literature Gaps

However, there appears to be a significant gap between academic research and practical applications as a result of differing methods of collecting and labeling data when dealing with AHU and VAV systems. This discrepancy can affect the ability to accurately detect and diagnose faults in HVAC systems[19]. Laboratory experimental data, such as ASHRAE RP-1312 [17], or simulated data provided by institutions such as Lawrence Berkeley National Laboratory [18] are the basis for the majority of the above-mentioned studies in the FDD literature. Despite helping to address the scarcity of publicly available benchmark datasets, these datasets may not be true representatives of real-world conditions [93]. Huang et al [20]. discovered that FDD strategies developed with simulated data and then applied directly to real building data did not produce satisfactory results. It is often the case that data granularity, metadata descriptions, and well-labeled faults with different severity levels are lacking in real-world applications. To better reflect real-world conditions, existing building datasets should be collected and studied to supplement the applicable FDD method [94]. Another crucial aspect to consider is that a significant portion of developed FFD algorithms focus on steady-state conditions for their application. This emphasis on

steady-state conditions is often due to the inherent challenge of accurately modeling and predicting transient conditions. As a result, transient behavior complexities are typically ignored in favor of more manageable steady-state scenarios [95].

Although many researchers have employed different data-driven techniques to develop tailored AFDD solutions for specific HVAC systems, these methods often lack transparency during testing and implementation, providing little information on how faults are detected and isolated [32]. However, it is a significant challenge for building operators to understand the mechanisms and inferences of data-driven methods, especially black-box supervised ones, which rely on complex relationships between data, and to trust the results [96]. Consequently, real building operators need an AFDD system that can provide at least minimal information about the operational conditions leading to faults, enabling them to analyze and trust the results.

Commercial buildings' HVAC systems are designed and installed based on the specific conditions of each building. Factors such as weather, internal loads, occupant behavior, and schedules vary from building to building, leading to different historical data behaviors. Consequently, a data-driven AFDD tailored for one building may not apply to others [14]. So, a generalizable AFDD that can be applied to different buildings within the same class with acceptable performance is a significant need in the commercial buildings HVAC AFDD market.

The availability of data for the training stage of AFDD is another significant challenge. Obtaining comprehensive faulty data that includes all types and severities of faults is nearly impossible. Additionally, the BEMS data from existing buildings often suffers from low quality, short and discontinued time ranges, numerous missing values, and inconsistencies [97]. Furthermore, many previous studies on AFDD have developed methods using simulated data, experimental data, or real operational labeled data containing specific faults with specific severities for particular systems, limiting their generalizability [98,99]. Therefore, there is a need for a data-driven AFDD system that can utilize raw, unlabeled data from BEMs, representing another crucial aspect for the advancement of commercial building AFDD systems.

Another significant gap in previous works is the lack of consideration for the historical behavior of systems when using PCA-based FDD. Most researchers have used only cleaned data without sequence or continuity, or just a segment of cleaned steady-state data for training PCA. Although PCA has been applied for time series analysis in various fields (like quality control) [61], PCA-based FDDs have yet to leverage time series analysis fully.

Previous literature has paid limited attention to the detection and diagnosis of programming logic faults in HVAC systems. However, inverse models, combined with rules outlined in ASHRAE Guideline 36, offer a promising approach to addressing this issue [59].

Chapter 3: Semi-supervised AFDD Method

3.1. Objectives

The main goal of this method is to establish a framework for fault detection and diagnosis in light commercial buildings using unlabeled raw data from BEMS for both steady-state and transient conditions. The utilization of a simple case study involving one AHU and four VAV boxes offers a valuable opportunity to develop a flexible framework that can be extrapolated to diverse real-world scenarios. This overarching goal can be further divided into the following sub-objectives:

- Developing a method to detect anomalies (potential faulty patterns) within the operational data, after the removal of sensor faults;
- Investigating anomalies to distinguish unseen or hidden faulty patterns from outliers and extracting the corresponding faulty rules. Subsequently, labeling the dataset based on this method; and
- Applying an AFDD classification to the normal and faulty classes generated through the labeling process and evaluating its performance for application in real conditions.

3.2. Proposed Framework

The first step in the framework involves gathering the essential sensor values using Input/Output (I/O) reports or other available information of the BAS. Because of the diversity of data collection methods, sensor data may vary in terms of time intervals. The sensor data should be synchronized with a specific time step to produce a consistent time series. Several factors may influence the suggested time steps, including the data collection protocols, the volume of data, and the time sensitivity of the data. However, 5-15 minutes is recommended. Data frames derived from BEMS systems are raw and unlabeled, making it imperative to resolve three crucial challenges including missing values, incompleteness, and incompliance. There are several reasons why these challenges

may arise, including sensor faults or inconsistencies in communication protocols [19,95]. To overcome these challenges, an initial strategy involves embarking on data visualizations coupled with statistical analysis. Given that the process of identifying outliers and differentiating them from faults is a more advanced stage, the current focus centers on addressing incompliant, inconsistent, and missing values.

Dimensional reduction techniques have become indispensable when dealing with unlabeled raw data frames without labels. PCA is one of the most widely applied FDD methods among these various techniques. By using PCA, a set of high-dimensional variables with possible correlations can be reduced to a set of low-dimensional, linearly uncorrelated variables [50]. A primary objective of PCA in this study is to reduce dimensionality while extracting essential information from the dataset. In summary, Principal Component (PC) analysis is a multi-step process that begins with data preparation. First, missing values are removed, ensuring a complete dataset for analysis. Subsequently, the data is normalized and centered. This is a crucial step to ensure that variables are on a consistent scale and that analysis is not skewed by differences in units or magnitudes. Next, a covariance matrix is constructed based on the processed data. This matrix captures the relationships and variances among the variables. Eigenvalue decomposition is then applied to this covariance matrix, resulting in a set of eigenvalues and corresponding eigenvectors. The eigenvalues represent the variance explained by each principal component, offering insights into the significance of these components in the data. Meanwhile, the eigenvectors define the directions in the original feature space along which the data exhibits the most variation. To build the PC matrix, the data is transformed using these eigenvector directions, effectively projecting it onto an orthogonal set of axes. Each PC captures a different level of variance in the data. The detailed procedure is described in [100].

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a powerful algorithm used in data analysis and pattern recognition, particularly in large spatial databases. Its primary purpose is to uncover meaningful clusters or groups of data points within a dataset based on their similarity[101]. What sets DBSCAN apart is its ability to identify clusters of arbitrary shapes.

Instead of assuming that clusters are spherical or have specific geometric forms, it looks for areas in the dataset where data points are densely packed together. These dense regions are separated by areas where data points are more sparsely distributed[24]. DBSCAN's approach is particularly useful in scenarios where traditional clustering algorithms might struggle. For example, it can effectively identify irregularly shaped clusters or clusters of varying sizes. It's also quite robust when dealing with datasets that contain noise or outliers—data points that don't fit neatly into any cluster and are located in regions of lower data density. In addition to clustering, DBSCAN has the added benefit of outlier detection [102]. By design, it recognizes low-density areas as potential outliers. These are data points that exist in regions where there isn't a clear cluster. Identifying outliers can be crucial in various applications, as well as fault detection [103].

While DBSCAN is effective at grouping similar data points and pinpointing anomalies within a large dataset, it falls short when it comes to describing the specific properties or features that define each cluster (or the anomaly). To understand the underlying rules or behaviors that make data points within a cluster similar to each other, we must turn to additional techniques and methods. In other words, DBSCAN provides the "what" by identifying clusters and anomalies, but to uncover the "why" or the reasons behind these patterns, we need to employ complementary approaches. It is also necessary when we deal with an unlabeled dataset in FDD to recognize the outliers from faults and during the isolation process. Nevertheless, in previous studies, researchers have often turned to methods such as Association Rule Mining (ARM) and Decision Trees to unearth the hidden behaviors and characteristics of clusters[36]. In this study, despite dealing with a numerical dataset, the Decision Tree method has been chosen for the specific purpose of uncovering insights and patterns within the data. Following the identification of the underlying rules governing each cluster, the study focused on the rules associated with anomaly clusters. To determine whether anomalies were caused by faults or simply outliers, a comprehensive analysis of the rules governing anomalies was carried out. The Decision Tree (DT) rules were used to generate labels and consultations with system experts were conducted to streamline this process. As a result, these labels were assigned to the original dataset and outliers were removed. Consequently, the dataset

has been structured and classified, enabling a more informative and structured presentation for the FDD process. In a broader sense, labeling was instrumental to the study's goal of identifying and understanding fault conditions within the dataset, ultimately facilitating subsequent analyses, and enhancing interpretability.

To address the issue of dataset imbalance, the SMOTE oversampling technique was applied in the final phase of the proposed framework. This involved increasing the number of fault samples across all fault classes to achieve a more balanced dataset. This oversampled dataset was then used to train a DT classifier. The trained classifier was then rigorously tested and evaluated using the original dataset without any modifications. The objective of this step was to evaluate the classifier's performance and ability to detect faults accurately in the field while accounting for the complexities of recording HVAC data on light commercial buildings. Figure 2. represents the comprehensive methodology adopted by this study.

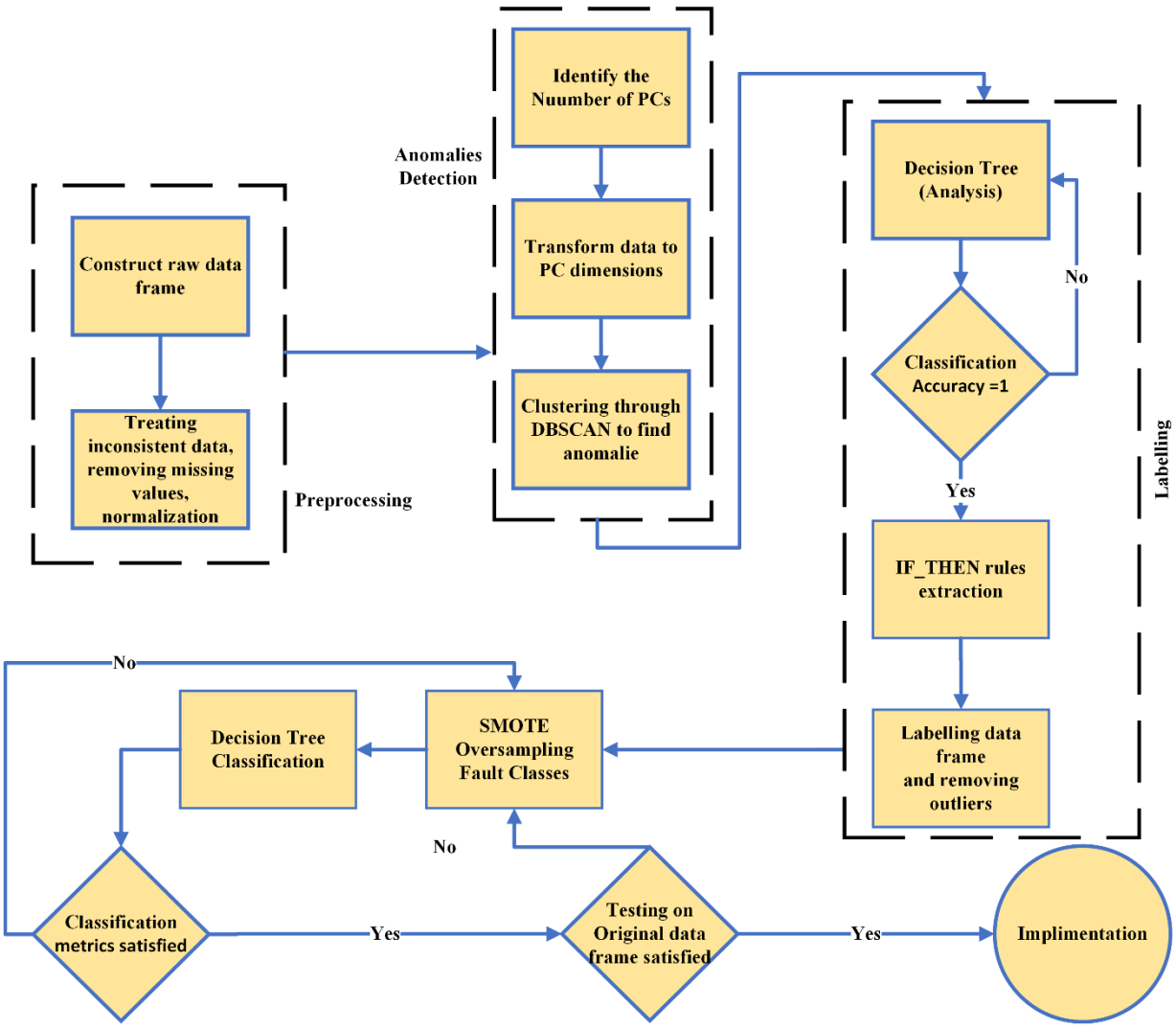
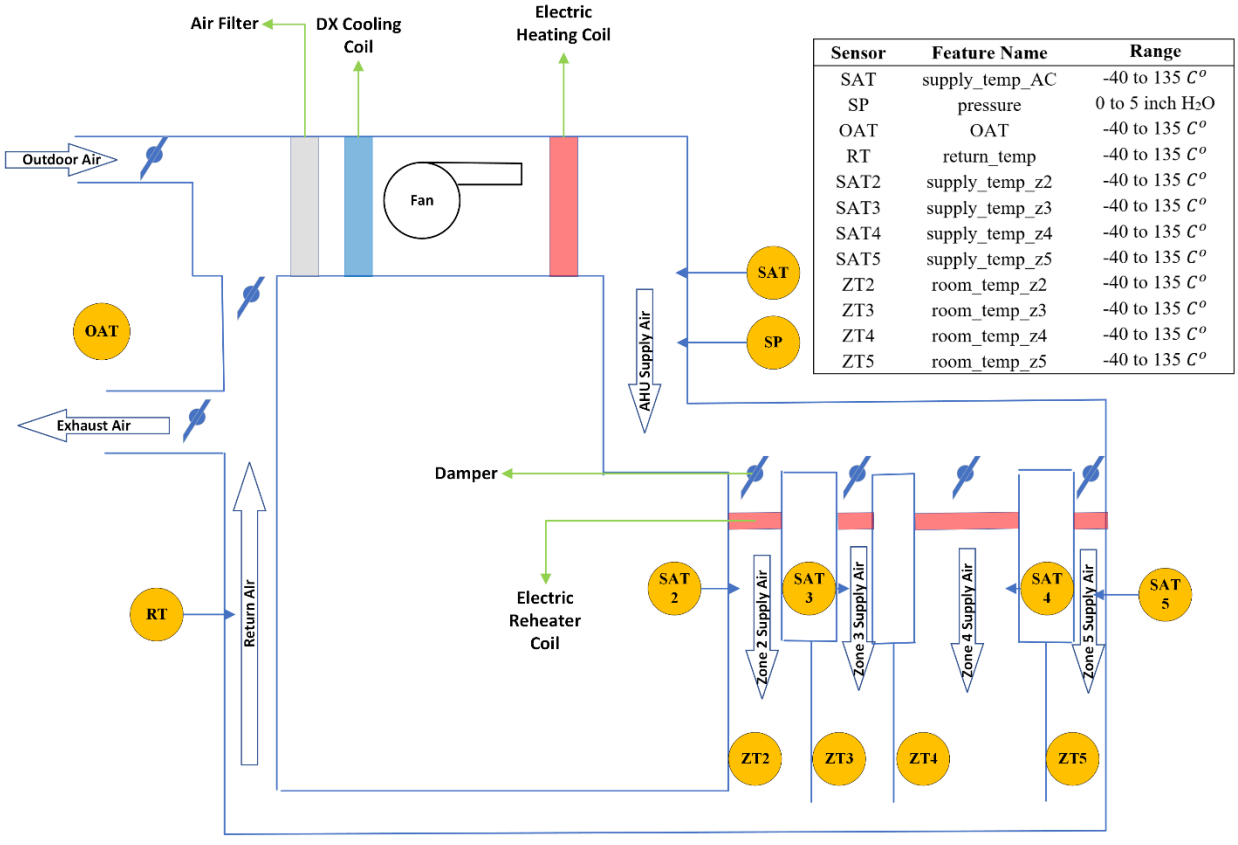


Figure 2. Flow diagram of the methodology

3.3. Case study and data collection

To evaluate the proposed comprehensive framework for the effective detection and diagnosis of HVAC faults in light commercial buildings, a comprehensive investigation was carried out on a real-world example of this building category. This specific building, located in the city of Montréal, Canada, was chosen as the focal point of the study. It features an HVAC configuration comprising

a single-duct AHU augmented with four reheating VAV boxes. The maximum heating capacity for the AHU electrical heating coil is 21.2 kW. The maximum flow rate of the supply fan, rated at 1.5 HP, is 2500 CFM. The cooling capacity is equivalent to 6 tons of refrigeration. The steady-state efficiency of the heating coil ranges between 80.5% to 81.1%, while the steady-state efficiency of the direct expansion cooling coil is 80%. The Energy Efficiency Ratio (EER) in cooling conditions is 12, and the Seasonal Energy Efficiency Ratio (SEER) is 14. The data nature and long data period make the building a valuable reference point for the development and assessment of FDD strategies in light commercial buildings. Figure 3. provides a schematic representation of the HAVC system configuration in the studied building. The architectural layout of the building is presented in Figure 4. Zone 2 to Zone 5 refer to the administration room, conference room, R&D room, and customer room, respectively.



Sensor	Feature Name	Range
SAT	supply_temp_AC	-40 to 135 C^o
SP	pressure	0 to 5 inch H_2O
OAT	OAT	-40 to 135 C^o
RT	return_temp	-40 to 135 C^o
SAT2	supply_temp_z2	-40 to 135 C^o
SAT3	supply_temp_z3	-40 to 135 C^o
SAT4	supply_temp_z4	-40 to 135 C^o
SAT5	supply_temp_z5	-40 to 135 C^o
ZT2	room_temp_z2	-40 to 135 C^o
ZT3	room_temp_z3	-40 to 135 C^o
ZT4	room_temp_z4	-40 to 135 C^o
ZT5	room_temp_z5	-40 to 135 C^o

Figure 3. The studied building`s HVAC system schematic and sensor locations

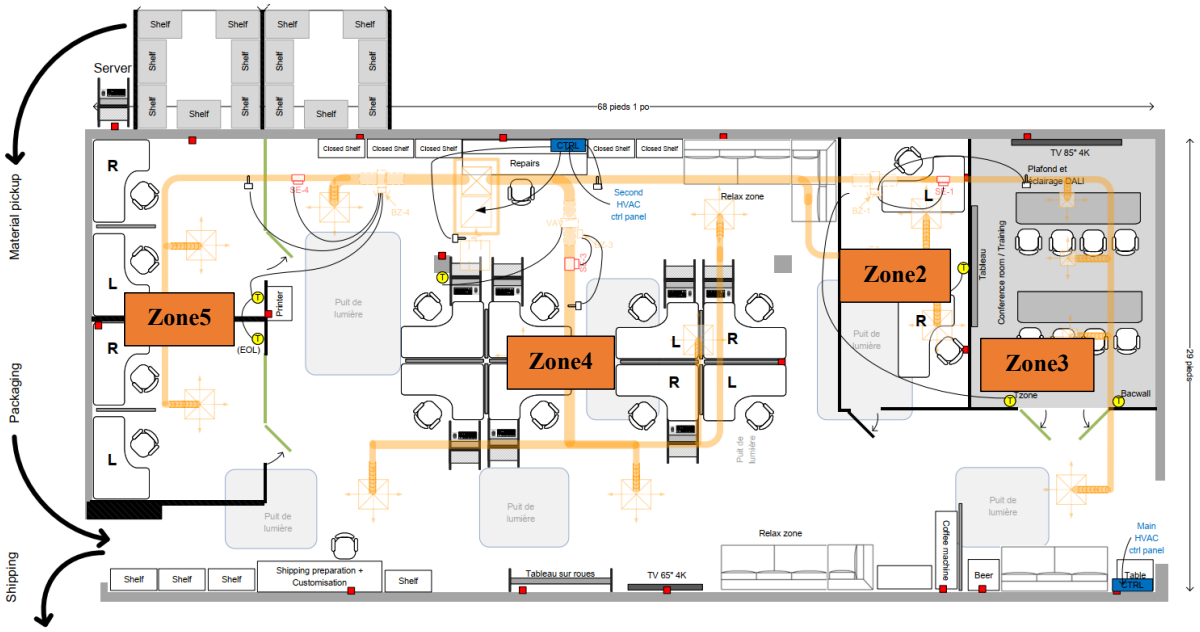


Figure 4. Floor layout of the studied building

In this study, data from the BEMS were obtained, covering the period from January 18, 2022, to March 8, 2023. The data collection method involved recording sensor values only when they changed, which ensured efficient storage and reduced redundancy. This type of sensor value is recorded in the system with a timestamp. In all other situations, the sensor values remained constant, so data was collected at regular intervals of 15 minutes, a common BEMS time step. The size of this dataset presented a significant challenge in ensuring efficient access to and processing of specific tags and instances. Furthermore, the varying time steps introduced an added complexity to the synchronization of information for a thorough analysis due to the existence of varying time steps. As a result of the size and irregularity of the dataset, a methodical approach was adopted to overcome the challenges presented by the dataset. It was first necessary to use a systematic extraction process to identify the various instances (representing specific zones or devices) and tags within the dataset (comprising of the sensor type and instance) before conducting the next step.

Table 3. Features and Descriptions from BEMS I/O Report

Feature name	Description	Type	Collection Method	Feature name	Description	Type	Collection Method
supply_temp_AC	Supply (discharge) air temperature of AHU (C)	Analog Value	Sensor	'OAT'	Outdoor air temperature (C)	Analog Value	Sensor
return_temp	Return air temperature to AHU (C)	Analog Value	Sensor	supply_temp_z2	Supply temperature after VAV reheating for Zone 2 (Administration room) (C)	Analog Value	Sensor
room_sp_z2	Room temperature setpoint for Zone 2 (Administration room) (C)	Analog Value	Control Value	supply_temp_z3	Supply temperature after VAV reheating for Zone 3 (Conference room) (C)	Analog Value	Sensor

room_sp_z3	Room temperature setpoint for Zone 3 (Conference room) (C)	Analog Value	Control Value	supply_temp_z4	Supply temperature after VAV reheating for Zone 4 (R&D room) (C)	Analog Value	Sensor
room_sp_z4	Room temperature setpoint for Zone 4 (R&D room) (C)	Analog Value	Control Value	supply_temp_z5	Supply temperature after VAV reheating for Zone 5 (Costumer room) (C)	Analog Value	Sensor
room_sp_z5	Room temperature setpoint for Zone 5 (Costumer room) (C)	Analog Value	Control Value	pressure	Supply pressure at AHU discharge (Inch H_2O)	Analog Value	Sensor
room_temp_z2	Room Temperature for	Analog Value	Sensor	fan_ac	AHU fan state	Binary Value	Control Value

	Zone 2 (Administration room) (C)						
room_temp_z3	Room Temperature for Zone 3 (Conference room) (C)	Analog Value	Sensor	cooling_ac	Cooling mode state	Binary Value	Control Value
room_temp_z4	Room Temperature for Zone 4 (R&D room) (C)	Analog Value	Sensor	heating_ac	Heating mode state	Binary Value	Control Value
room_temp_z5	Room Temperature for Zone 5 (Costumer room) (C)	Analog Value	Sensor	cooling_fan_demand	Cooling demand on fan	Binary Value	Control Value
room_temp_avg	Averaged room air temperatures of	Analog Value	Calculated	heating_fan_demand	Heating demand on fan	Binary Value	Control Value

	Zones (C)			
supply_temp_avg	Averaged supply air temperature of Zones (C)	Analog Value	Calculated	

3.4. Preprocessing and data cleansing

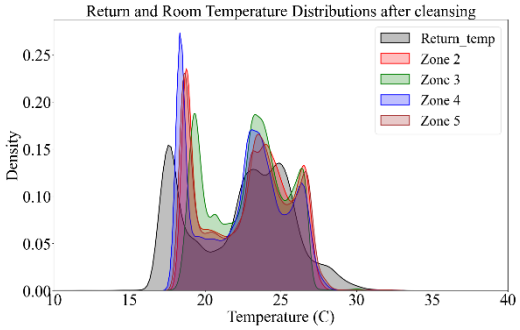
During this phase, the primary effort was directed toward investigating and rectifying issues related to data inconsistencies and gaps caused by missing values. As part of this phase, a deep understanding of the system domain is essential to distinguish unexpected values based on insights into the system's inherent characteristics and the context of the data being analyzed [104]. The majority of temperature data inconsistencies were attributed to examples when the sensors were not properly connected, resulting in sensor records reaching their maximum allowable values as per the configuration within the BAS. Figure 5(a), and Figure 5(b) show the temperature data distributions after removing inconsistent temperature data.

Besides the temperature data, the distribution of pressure sensor values, following the data cleansing process, is visualized in Figure 5(c). Initially, upon a cursory examination, it was noted that the minimum pressure when the AHU fan was turned off appeared to be 0.2 H_2O instead of the expected 0. However, upon conducting a thorough investigation within the BAS, it was uncovered that the sensor signal type had been incorrectly configured as 2-10 V for 1 Inch of H_2O, instead of the correct 0-10 V for 1 inch of H2O. Subsequently, this issue was rectified within the system, and the pressure data was successfully transformed from the [0.2, 1] interval to the intended [0, 1] interval.

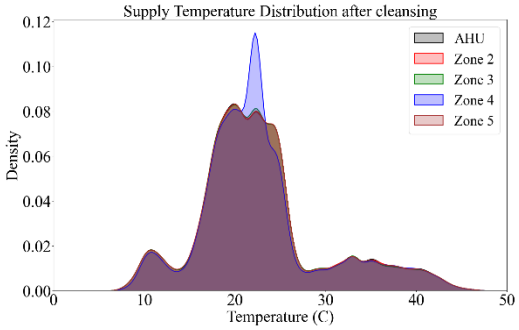
As outlined in Table 3, the BEMS I/O report defines various binary values in addition to analog values. In most cases, a binary value refers to the specific control state of the system. An illustration

of the distribution of each binary value is presented in Figure 5(d). As evident from Figure 2(d), the system predominantly operates in the heating mode in comparison to cooling. Additionally, the probability of the AHU fan being in the "on" state is notably higher than being in the "off" state.

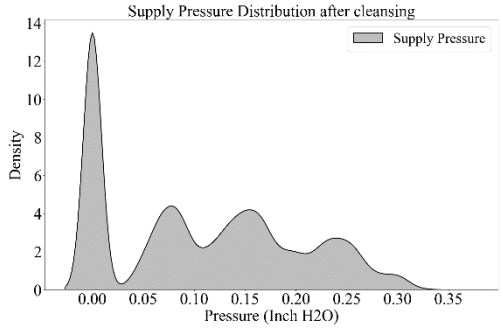
In addition to the temperature sensors depicted in the preceding figures, an additional sensor, the Outdoor Air Temperature (OAT) sensor, was examined. As a result of its location near intake or exhaust ducts and susceptibility to diverse ambient interferences like wind speed and solar radiation, the OAT sensor presents one of the most challenging aspects of HVAC systems. It is necessary to compare these data with those from a nearby weather station[95], in this case, Pierre Elliott Trudeau Airport weather data[105], to determine the accuracy of the OAT sensor. Referring to Figure 6(a) and consulting with domain experts specializing in the HVAC system of the building under study, it was observed that OAT sensor values closely align with the corresponding weather data. According to Figure 6(b), the average deviation was calculated as 0.8 Celsius, with a standard deviation of 1.83. Upon conducting a thorough investigation and consulting with domain experts, it was determined that this phenomenon was attributed to exhaust air leakage. These fluctuations could be considered negligible in terms of impacting the accuracy of the OAT readings. Figure 6(c) provides a specific instance of this behavior, observed on February 1st, 2022, and Figure 5(d) shows the OAT sensor values distribution after cleansing.



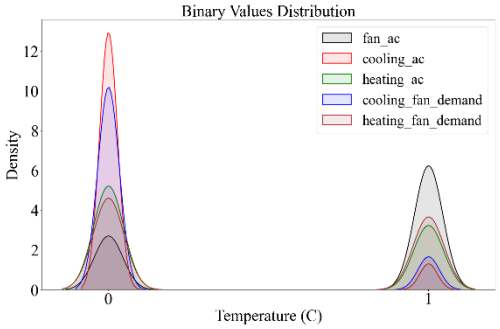
(a) Room temperatures and return temperature



(b) Supply temperature

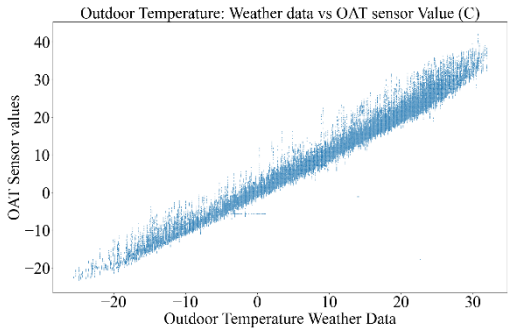


(c) Supply pressure

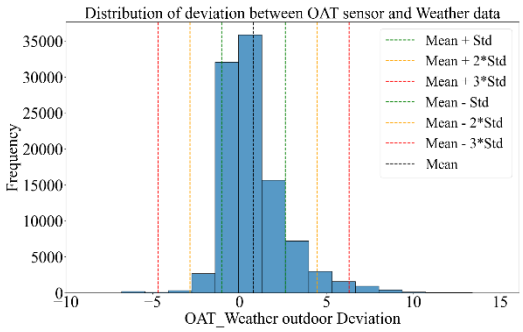


(d) Binary values

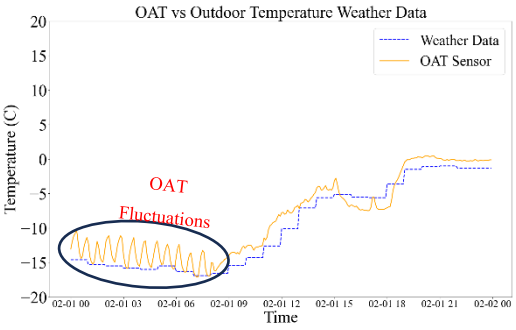
Figure 5. Data distribution after cleansing and vying incompliance data



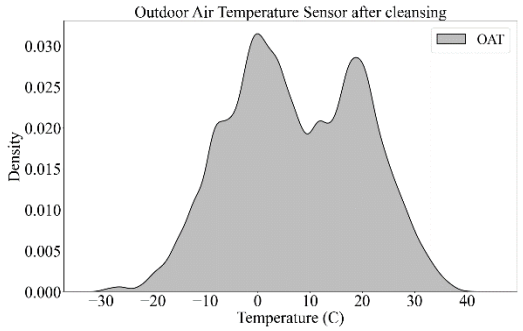
(a) OAT vs. Weather data



(b) Distribution of deviation between OAT values and Weather data



(c) OAT sensor value fluctuation on Feb 1, 2022



(d) Cleaned OAT distribution

Figure 6. OAT data inconsistency analysis

3.5. Result

The purpose of this section is to present results associated with the implementation of the proposed framework in the light commercial building case study. The results sections have been arranged according to the framework steps.

3.5.1. PCA results

Reporting the results will be started by the first phase of the framework which is PCA. As a guiding metric for determining the optimal number of PCs, the cumulative variance percentage has been calculated. The cumulative variance percentage indicates how much variance is accounted for by the selected PCs of the dataset. According to various studies cited in [45], an acceptable level of data variance explanation lies between 75% and 90%. As seen in Figure 7, 89.51 % and 93.59% of the variance in the data can be explained by n=4 and n=5 number of PCs, respectively. Therefore, n=5 was chosen for the number of PCs in the dimensional reduction step. The distribution of the data projected onto reduced dimensions is also shown in Figure 8 and confirms the existence of distinct clusters. These clusters likely reflect diverse system behaviors or patterns within the dataset. A further observation is that certain data points are located far from densely populated areas. These isolated data points may be potential faults or anomalies within the dataset. Furthermore, the arbitrary shapes of clusters can be extracted from projected data plotting as an additional significant piece of information.

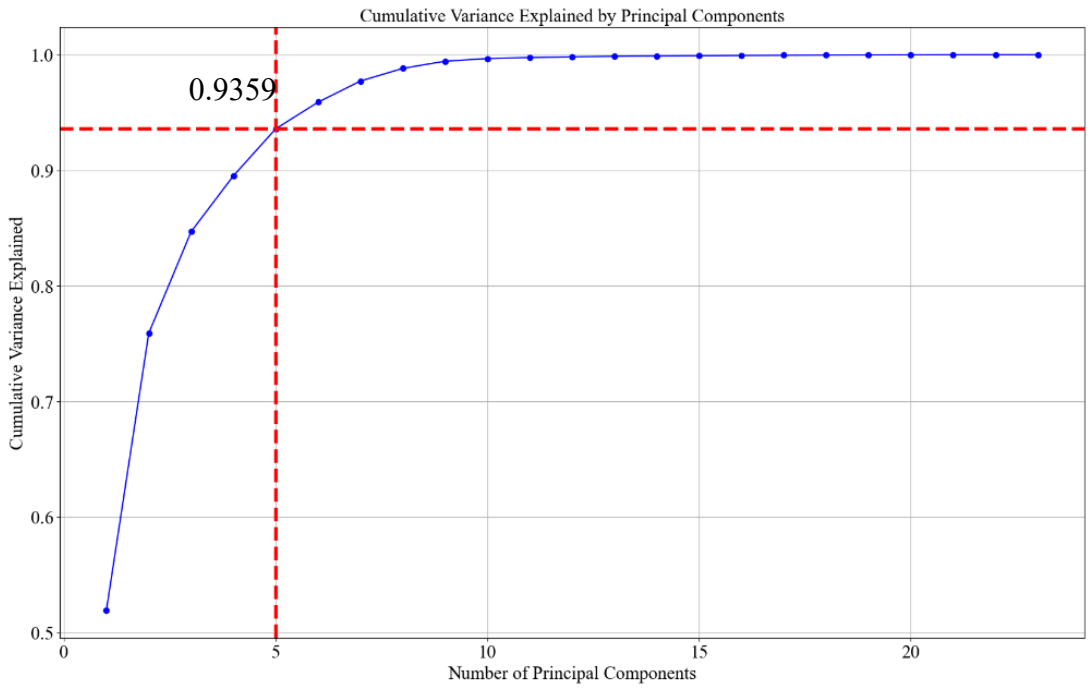


Figure 7. Cumulative Variance Explained by Number of PCs

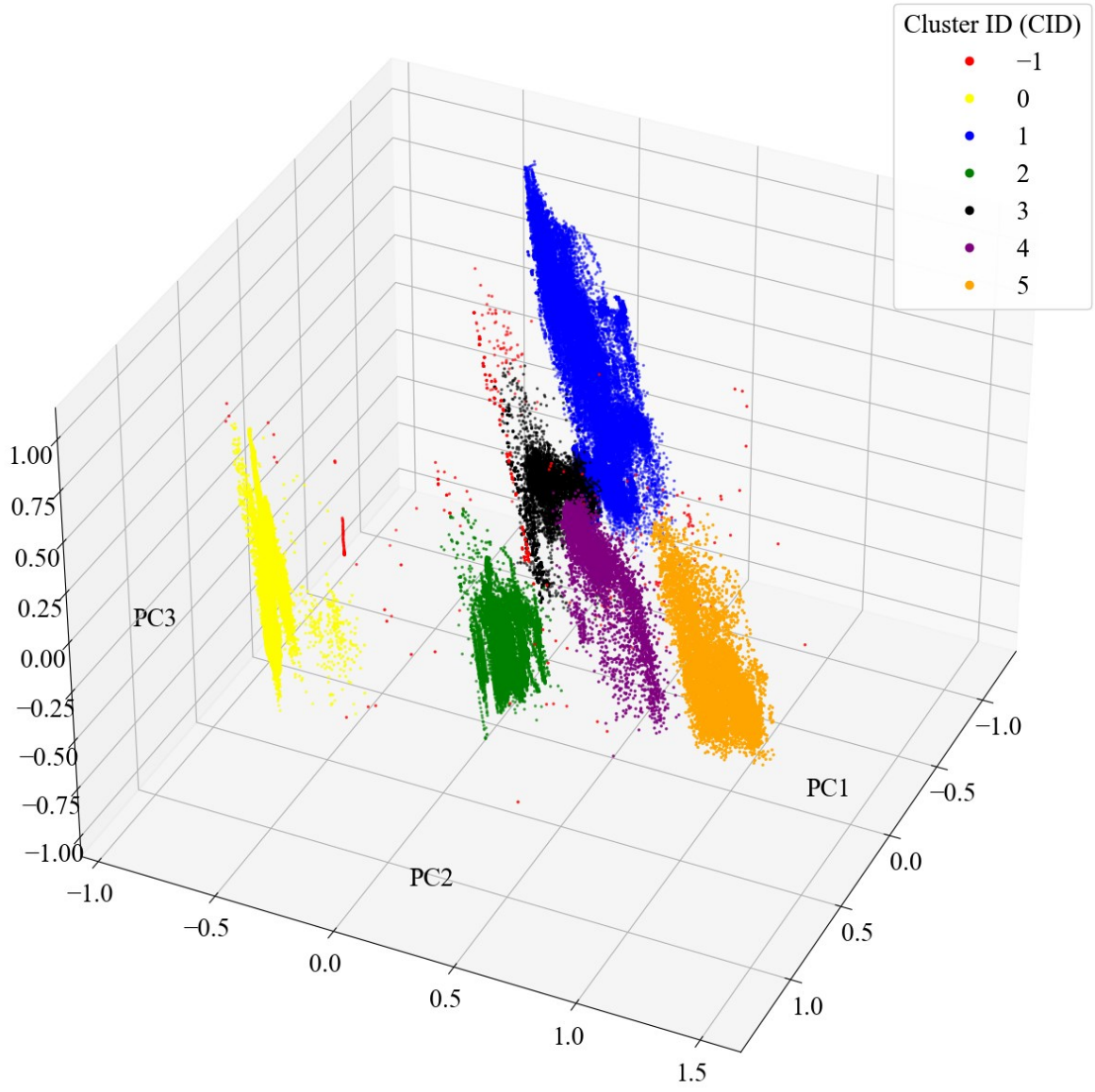
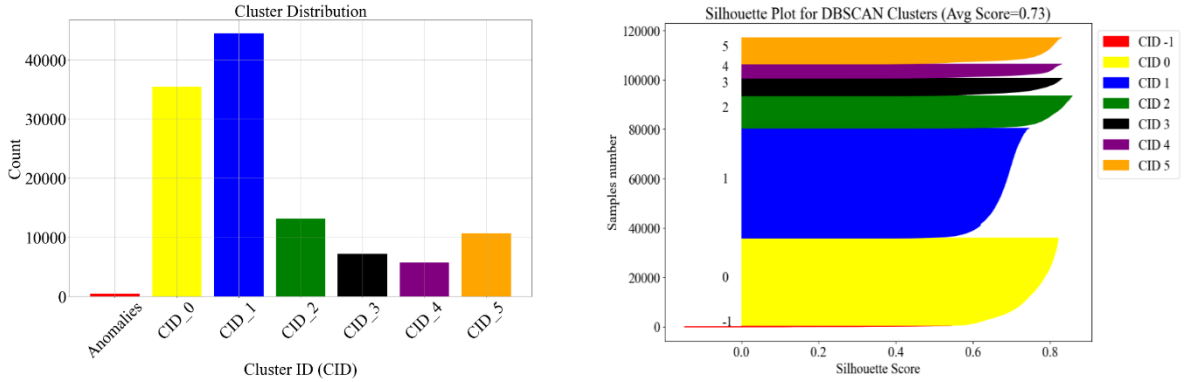


Figure 8. DBSCAN clustering presentation in the space of the first 3 PCs

3.5.2. DBSCAN results

The DBSCAN algorithm relies on two essential parameters: "eps" and "min_samples". That “eps” is the threshold for considering two points as neighbors and “min_samples’ is the number of neighbors that a point should have to be considered as a core point. A detailed description and

calculation of the silhouette score can be found in [106]. Based on the results, eps = 0.25 and min_samples = 200 were selected to guide the next phase of the investigation.Figure 8 also illustrates the results of DBSCAN, highlighting clusters as well as outliers, which may be anomalies or potential faults. Figure 9(a) shows the distribution of sample numbers in each cluster. The Cluster with CID #-1 is considered and named as the anomalies containing 482 samples. In Figure 9(b), silhouette scores are displayed for each cluster, specifically for the selected parameters of DBSCAN. There is a range of -1 to 1 for the silhouette score. An object with a silhouette score close to 1 has an excellent match with their own cluster but a poor match with their neighboring clusters, while an object with a score close to -1 has an adverse match with their cluster. Notably, it highlights that data within Cluster IDs (CID) #2 exhibit a high degree of similarity compared to data in other clusters. A further observation is that some data points within the anomalies class (CID #-1) have silhouette scores less than 0, indicating a lack of similarity among these data points. This may indicate that different fault patterns or anomalies exist within this class.



(a) Number of samples in each cluster

(b) DBSCAN silhouette score of each cluster (eps=0.25 and min_samples=200)

Figure 9. Clustering Evaluation and Sample Distribution in DBSCAN Clustering

3.5.3. Rules extraction, labeling, and faulty pattern results

To understand the relationships among data points within the anomalies cluster, the DT classifier attempted to be overfitted on the entire dataset to unravel every detail and relation between data in each cluster, especially in instances of anomalies. The process of overfitting may increase the depth

and complexity of the tree, as well as the length of the extracted rules, but to distinguish noise from anomalies, overfitting is an essential step. Regarding DT results, at least 8 DT depths needed to be overfitted using the whole data set. The DT classification provides a set of rules (antecedent) associated with anomaly occurrences, and these rules are documented in Table 4. Rules extracted from the DT classifier. This table provides a comprehensive overview of the rules for anomalies and encompasses all instances within the anomalies class. Figure 10 shows the top 5 attributes that have the greatest contribution to the classification, and consequently, to the identification of faulty patterns.

Table 4. Rules extracted from the DT classifier

	Antecedents	Sample numbers	Label
1	(heating_fan_demand = 0) and (fan_ac = 1) and (cooling_fan_demand = 0) and (heating_ac = 1) and (pressure <= 0.337)	283	Communication_issue
2	(heating_fan_demand = 1) and (heating_ac = 0) and (fan_ac = 0)	75	Fan_Stuck
3	(heating_fan_demand = 1) and (heating_ac = 0) and (fan_ac = 1) and (cooling_ac = 1)	47	Control_Fault1
4	(heating_fan_demand = 1) and (heating_ac = 0) and (fan_ac = 1) and (cooling_ac = 0) and (supply_temp_avg > 39.0) and (OAT > -8.444)	22	Heating_Coil_temp_high
5	(heating_fan_demand = 0) and (fan_ac = 1) and (cooling_fan_demand = 1) and (cooling_ac = 0) and (heating_ac = 1)	21	Control_Fault2
6	(heating_fan_demand = 0) and (fan_ac = 1) and (cooling_fan_demand = 1) and (cooling_ac = 1) and (heating_ac = 1)	10	Control_Fault2
7	(heating_fan_demand = 1) and (heating_ac = 0) and (fan_ac = 1) and (cooling_ac = 0) and (supply_temp_avg <= 39.0) and (supply_temp_z4 <= 11.806) and (supply_temp_z2 <= 10.556)	7	Outlier
8	(heating_fan_demand = 0) and (fan_ac = 1) and (cooling_fan_demand = 0) and (heating_ac = 0) and (cooling_ac = 0) and (supply_temp_z2 > 33.056) and (room_temp_z5 > 19.889)	7	Outlier
9	(heating_fan_demand = 0) and (fan_ac = 0) and (cooling_fan_demand = 1)	6	Outlier
10	(heating_fan_demand = 0) and (fan_ac = 0) and (cooling_fan_demand = 0) and (supply_temp_z4 > 38.833)	2	Outlier

11	(heating_fan_demand = 1) and (heating_ac = 0) and (fan_ac = 1) and (cooling_ac = 0) and (supply_temp_avg <= 39.0) and (supply_temp_z4 <= 11.806) and (supply_temp_z2 > 10.556) and (pressure > 0.353)	1	Outlier
12	(heating_fan_demand = 0) and (fan_ac = 0) and (cooling_fan_demand = 0) and (supply_temp_z4 <= 38.833) and (supply_temp_z3 <= 8.694)	1	Outlier
	Total Number	482	Anomalies

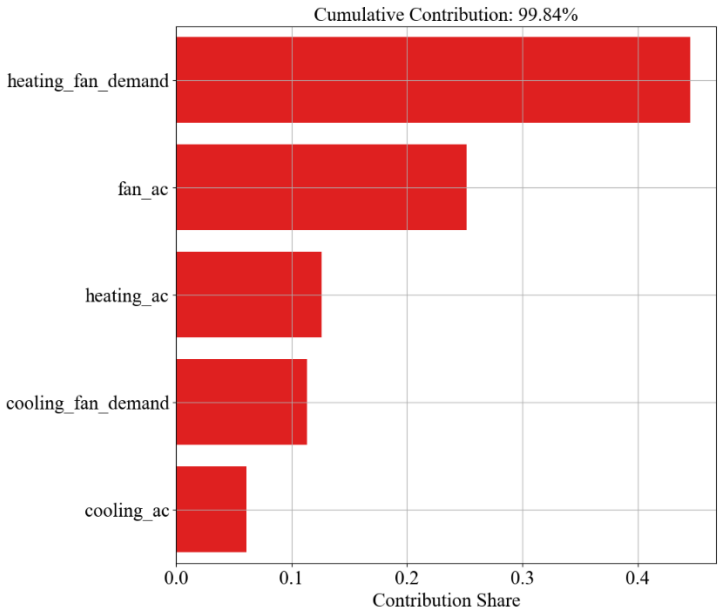


Figure 10. Contribution share of important features in DT classification for anomaly rules extraction.

In the next step, rules with more than 10 samples were considered as faulty patterns of the system. Afterward, these patterns were discussed with the building operators to comprehensively label the faults. Several of these patterns have repeating behavior (i.e., repeated in consecutive time steps) over a specific period, whereas others represent isolated data points with similar patterns. With 283 samples, there is a faulty pattern within ~58% of the anomaly clusters data, that point to a communication issue fault. A communication problem existed between the BEMS and the sensors. A critical symptom of the pattern is that *heating_fan_demand* and *cooling_fan_demand*, were 0,

while *fan_ac* and *heating_ac* were 1, which indicates there was no demand for heating, but heating still took place. Another symptom of this faulty pattern can be observed when the *supply_temp_AC* is increasing, yet the temperature in room zone 4 (*room_temp_z4*) remains constant. It appears that this faulty pattern occurred between the 16th and 18th of February 2023. Figure 11. shows the *room_temp_z4* and *supply_temp_AC* for three normal days (13th to 15th February) and faulty days. Furthermore, the right y-axis represents the binary values *heating_ac* and *heating_fan_demand*.

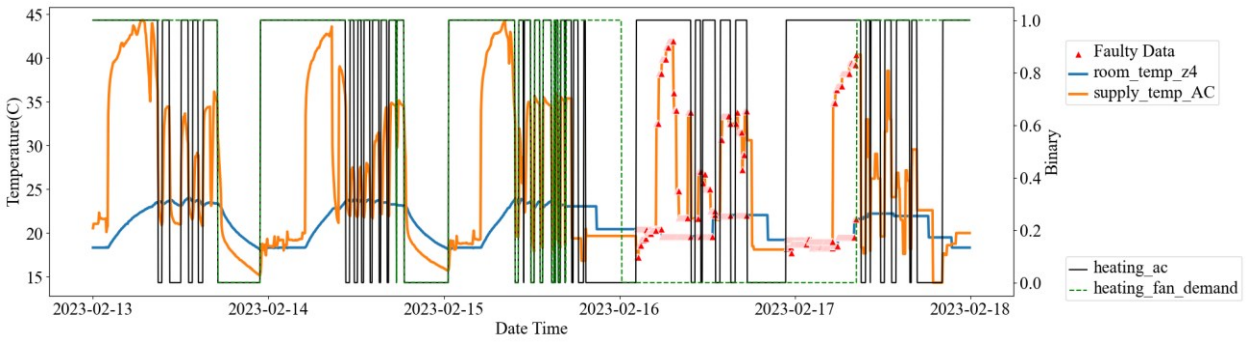


Figure 11. Communication faulty patterns between the 16th to 18th of February 2023.

Regarding Figure 12, the occurrence of Fan_Stuck fault represents a repeating fault pattern. a notable event transpired when the system mode shifted from unoccupied to occupied. During this transition, an anomaly emerged as the *heating_fan_demand* was recorded as 1, indicating a request for heating, while both *fan_ac* and *heating_ac* remained at 0 which means there are no fan working and no heating provided. This discrepancy implies that the HVAC system was not functioning properly. Consequently, both *supply_temp_AC* and room_temp_z4 experienced simultaneous and consistent decreases, attributable to the system's inactivity. This fault persisted from 6:00 AM and extended until 11:35 AM, at which point the fan resumed operation. As a result, *supply_temp_AC* began to rise, gradually warming the various zones within the building.

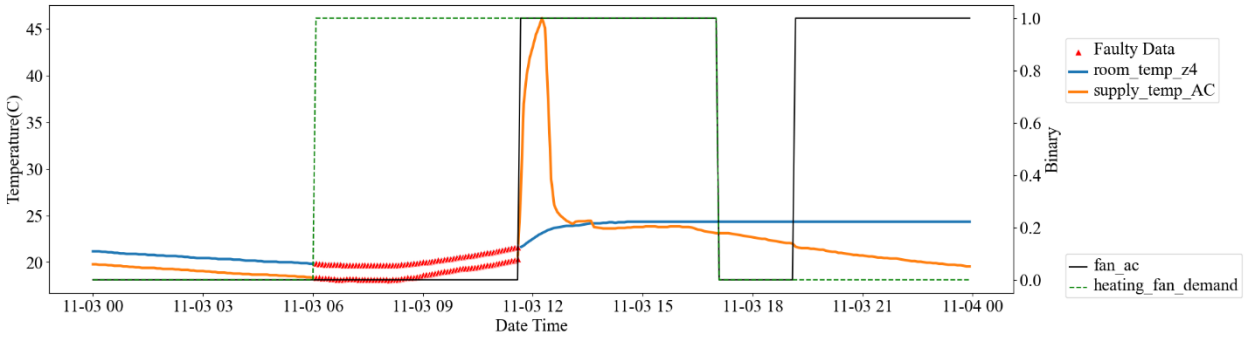
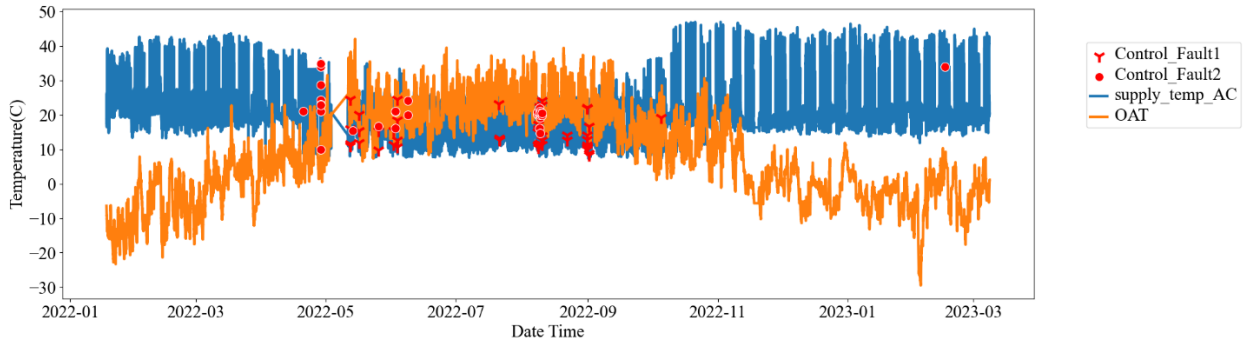
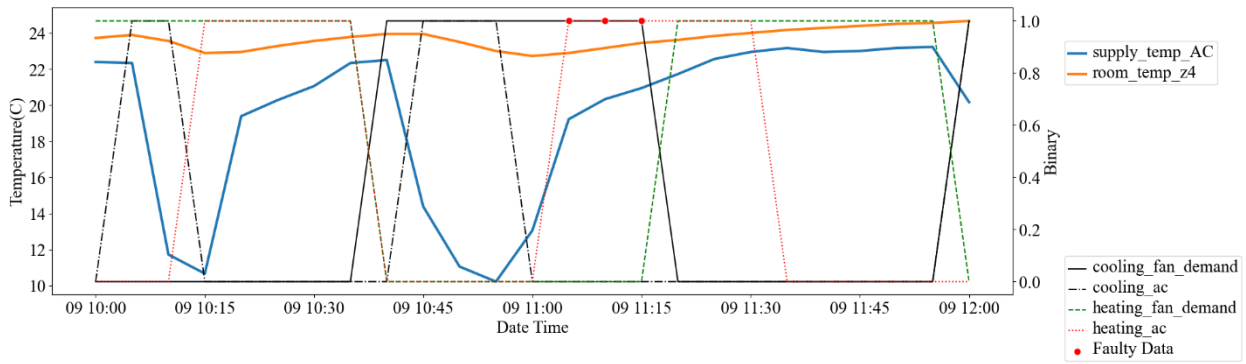


Figure 12. Fan_Stuck fault pattern during the 2022-11-03.

In Figure 13(a), a visual representation of the occurrence times of two different types of control faults is shown, namely Control_Fault1 and Control_Fault2. With the confirmation from the building HVAC operator, the issues labeled Control_Fault1 and Control_Fault2 are accountable for the sizing problem that happened during the design phase. Most of these faults occurred during the summer season on days when the OAT was relatively low. Both single-isolated samples and sequences of two to four consecutive samples (from 5 minutes to 20 minutes) are included in them. An instance of this type of fault occurred on August 9, 2022, when the system was operating in heating mode without corresponding heating demand. Rather, a consistent cooling demand persisted.



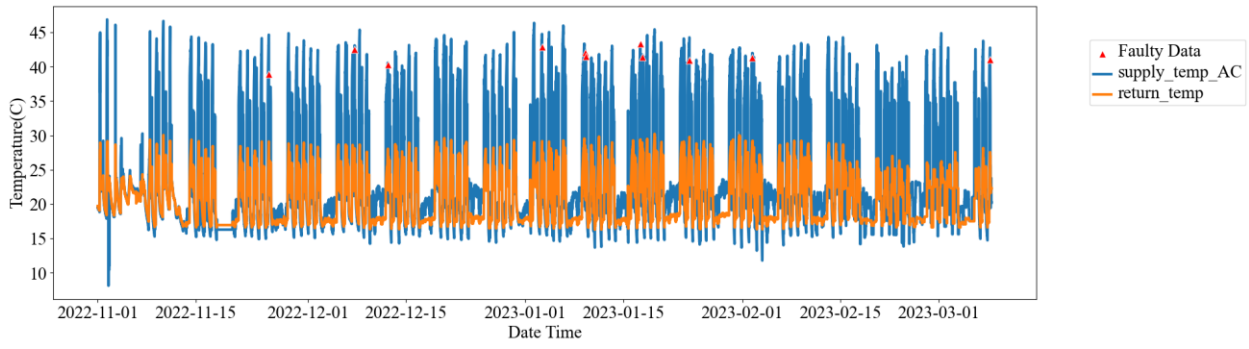
(a)



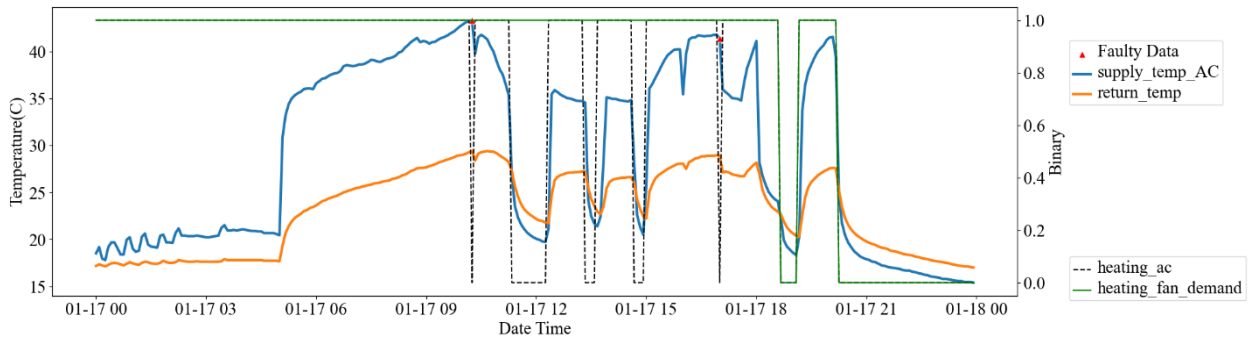
(b)

Figure 13. Control Faults occurrence, (a) whole period, (b) an example on 2022-08-09.

Lastly, the Heating_Coil_temp_high faulty pattern was identified. In this case, it refers to the protective action involved in the heating coil of the AHU. There is no repeating pattern associated with this anomaly; rather, it arises sporadically as isolated incidents. However, it can occur more than once within a single day, as illustrated in Figure 14(a). Figure 14(b) illustrates that this fault occurs when the return temperature reaches a relative maximum, and the OAT is higher than -8.44 °C (refer to the extracted rule #4). Heat is extracted less efficiently from the heating coil when the return temperature increases. To prevent potential damage or failure of the electrical heating coil, the HVAC system initiates protective actions. During this time, the heating operation is temporarily suspended to allow the coil to cool down and protect it from damage. In these situations, it may appear that the system is performing adequately, however, the highly frequent overheating of the heating coil may increase the risk of failure and contribute to a higher maintenance cost over time.



(a)



(b)

Figure 14. Coil_temp_high fault pattern (a) from 2022-11-01 to 2023-03-08, (b) on 2022-12-07.

3.5.4. AFDD results

Since the main purpose of the previous DT was rule extraction and labeling the dataset, in the initial step of the AFDD DT phase, the accuracy of DT classification was examined at varying maximum depths and consistent accuracy was observed after reaching a maximum depth of 8. As a result of the feature contribution analysis, the findings presented in Figure 15 reveal *that supply_temp_avg* exhibits the highest contribution, followed by *room_sp_z2*, *room_sp_z3*, and *supply_temp_AC* in AFDD DT classification. Besides, to evaluate the effect of SMOTE oversampling on the accuracy of the AFDD DT classification, a sensitivity analysis was conducted on the AFDD DT classification whit k_neighbors = 2, 4, 6, and 8, and the results average accuracy equal to 0.98.

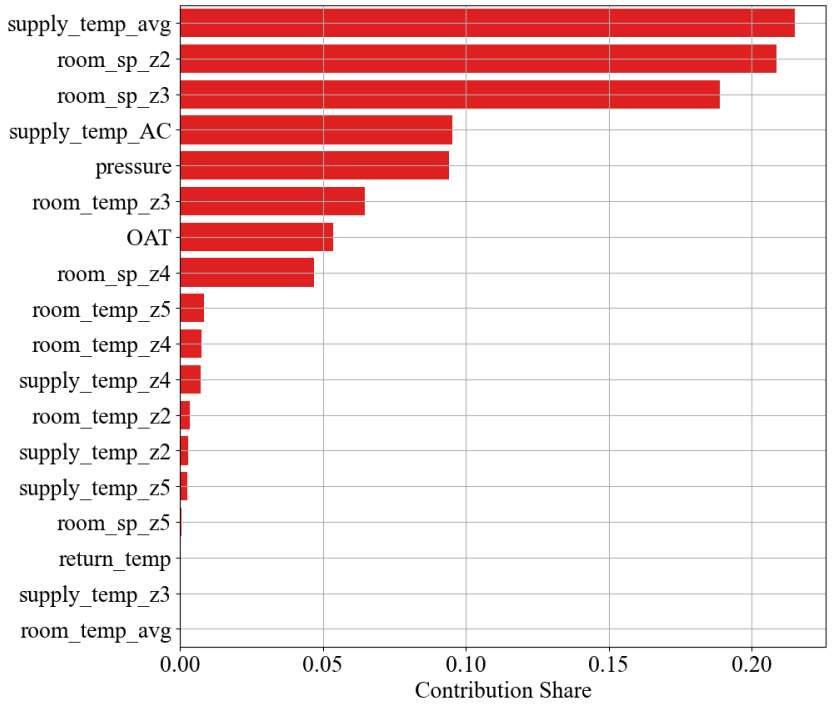


Figure 15. Feature contribution in DT classification in AFDD phase.

Table 5. AFDD performance.

	MAR (%)				FAR (%)				Detection Time
Resampling Attempt	1	2	3	4	1	2	3	4	
Heating_Coil_High_Temp	0.00	0.02	0.02	0.00	0.20	0.19	0.19	0.19	5 Min
Communication_issue	0.79	0.85	0.84	0.86	0.01	0.01	0.01	0.01	5 Min
Fan_Stuck	1.31	1.27	1.32	1.27	0.73	0.73	0.73	0.72	10 Min
Control_Fault1	7.78	7.98	7.99	7.92	2.40	2.35	2.34	2.34	5 to 15 Min
Control_Fault2	9.03	9.12	9.17	9.15	2.05	1.94	1.93	1.95	5 to 15 Min

Finally, the critical validation step involved testing the model on the raw unbalanced dataset, without any oversampling, encompassing real-world operational data from the BEMS system. This

step is essential to assess the model's performance under real operating conditions to establish a systematic measurement, four distinct SMOTE oversampling attempt phases were conducted using 50% oversampled data, followed by testing on the initial dataset to evaluate the effect of oversampling on final result. The results, including the Missing Alarm Ratio (MAR) and False Alarm Ratio (FAR), have been presented in Table 5. AFDD performance. Based on the results, it is observed that Control_Fault1 exhibits the highest False Alarm Rate (FAR), with Control_Fault2 being the second highest. Furthermore, Control_Fault2 consistently records the higher Missing Alarm Ratio (MAR) across all four trials, followed by Control_Fault1. Lowest FAR and MAR percentages related to Communication_issue, and Heating_Coil_High_Temp, respectively. Consequently, according to AFDD model test results on the unbalanced raw dataset, the model's total accuracy decreased from 95% to 89% due to inaccuracy in the detection of Control_Fault1 and Control_Fault2.

3.5.5. Discussion

With the completion of the AFDD study, it is worthwhile to mention that during the data cleansing process, sensor faults were eliminated, and previously unseen patterns were discovered that operators were not able to detect. Even though this method has proven extremely effective at detecting faulty patterns and labeling datasets, the question remains: can this labeled dataset be effectively employed to detect and diagnose faults in real-world conditions within this light commercial building? To further examine the applicability of this approach, DT classification was performed on the labeled dataset to develop an automatic method for fault detection and diagnosis. Nevertheless, there is a challenge in dealing with the features, which have a significant contribution to the labeling process. It may result in an overfitting of the classification if they are reused. Consequently, the 5 features with the highest contribution in the rule extraction phase were dropped, and the dataset was labeled according to the rules outlined in Table 3, and outliers were then removed from the dataset.

Another significant challenge encountered when dealing with massive data from a light commercial building pertains to the issue of an imbalanced dataset (due to the fact that the number of faulty samples are extremely lower than normal samples), which resulted in a low F1 score for faulty classes despite achieving a notably high accuracy rate. Within the DBSCAN analysis, the entirety of the anomalies cluster encompassed 482 data points, equivalent to only 0.4% of the entire cleansed dataset. Labeling the dataset produced classes with even lower sample sizes. To address the challenge of the imbalanced dataset in the case of the proposed framework, a SMOTE oversampling technique was implemented for the faulty classes. This rebalancing of the dataset aimed to enhance the model's ability to effectively detect and classify faults, ultimately improving the overall performance of the fault detection and diagnosis method. In this manner, a considerable amount of data was produced, which provides ample information for training and testing purposes. As part of our efforts to ensure the effectiveness of the AFDD method in handling unseen data in the future, a split of 50% was made between the training and testing portions of the dataset. This balanced approach ensures the AFDD method's robustness in practical scenarios.

It should be noted that the six faults that were studied and discussed in this study were not recognized prior to the analysis of raw data. These faults have been detected during the labeling process and have been confirmed by consulting the HVAC operators of the building, as well as the system controller designers, as a result of the labeling process. A couple of these faults have been mentioned in previous literature, such as fan_stuck, control faults, and communication issues, but Heating_Coil_High_Temp is a unique fault that only applies to this system (as this system uses electric heating coils). Other faults, such as those related to valve positions (which were not identified in this system), sensor faults, and other types of faults could also be identified and labeled if these types of faults were found in the historical data used in this study. For this reason, it may be necessary to repeat the framework on a regular basis (such as every 3 or 6 months) in order to update the labels and make the framework more applicable to a larger number of faults.

It is worth emphasizing that, in the context of fault detection and diagnosis, the consequences of missing alarms significantly outweigh those of false alarms. Verifying false alarms through manual

checks is relatively straightforward, whereas missing alarms can lead to severe damage to the HVAC system[107]. The higher rates of missing alarms in Control_Fault1 and Control_Fault2 can be attributed to their association with a sizing problem, which is fundamentally a design issue. This problem has persisted in the system's behavior since the system's inception, making it considerably more challenging to detect anomalies solely based on sensor values, particularly when employing historical data-driven methods. However, The AFDD DT classification performed very well in identifying faults in other fault classes. A noteworthy observation from the results is that false alarms across all classes correspond to normal conditions rather than other types of faults, which is very important to the diagnosis process.

3.5.6. Conclusion

A comprehensive framework for establishing a semi-supervised Fault Detection and Diagnosis (FDD) in light commercial buildings was outlined in this study. This framework, driven by unlabeled raw data obtained from the BAS, was devised to create an adaptable methodology applicable to real-world scenarios. To achieve this goal, a series of objectives were pursued. The framework involves the creation of a sizable and synchronized raw dataset, derived from the multitude of sensor tags embedded within the BAS. A pivotal stage was the comprehensive investigation of these anomalies, where the objective was to differentiate concealed, unseen faulty patterns from outliers. This exploration ultimately led to the elucidation of a set of rules governing these anomalies, thereby facilitating the labeling of the dataset in accordance with the derived methodology.

The subsequent phase of this research delved into the implementation of an AFDD classification model on the labeled dataset. This analysis provided insights into the model's efficacy in discerning between normal and faulty classes. Notably, the model performance was assessed across various fault types, unveiling both strengths and areas for enhancement. In summary, this study makes a significant contribution to the field of fault detection and diagnosis of light commercial buildings HVAC by offering a systematic approach to handling real-world operational data. The results

underscore the potential of the proposed framework to bolster the efficiency and accuracy of FFD in light commercial buildings, ultimately elevating the performance and occupant comfort of HVAC systems. The highlighted outcomes can be summarized as follows.

- The combination of PCA (for dimensionality reduction) and DBSCAN (for unsupervised clustering) demonstrates remarkable proficiency in identifying anomalies within HVAC system datasets. This methodology showcases its effectiveness in uncovering irregularities, even when dealing with complex and unstructured data.
- Decision Tree classification emerges as a potent tool for revealing patterns within the normal and anomalous operational conditions of light commercial building`s HVAC systems. This approach provides a systematic means to unravel the underlying rules governing system behavior, contributing to more accurate fault detection and diagnosis.
- Imbalanced datasets can compromise the performance of AFDD. To mitigate this, SMOTE oversampling techniques are shown to be effective in enhancing AFDD's overall effectiveness in HVAC systems. By increasing the representation of faulty classes, the model becomes better equipped to distinguish between normal and anomalous conditions.
- Control faults related to sizing problems in HVAC systems pose a unique challenge for fault detection. These issues are deeply rooted in the historical behavior of the system and tend to recur frequently. Consequently, the detection of sizing problems may lead to an elevated false alarm rate, as the data closely resembles normal operating conditions. This underscores the importance of fine-tuning AFDD algorithms to account for such nuances in system behavior.
- Certain fault types, such as Fan_Stuck, Communication_issue, and Heating_Coil_High_Temp, exhibit robust detection and diagnostic capabilities within the studied framework. These faults, although less common and less similar to normal conditions, can be reliably identified. This highlights the potential for proactive fault management, even in cases where fault occurrences are infrequent.

- The accuracy of the proposed AFDD method based on the different faults ranges between 0.89 and 0.99. It shows better performance in detecting and diagnosis Communication_issue, Fan_Stuck, and Heating_Coil_temp_High faults, and lower performance to diagnose Control_Faults.

While this framework has demonstrated commendable performance, it is important to acknowledge certain limitations that offer avenues for further research and improvement. The first direction of improvement is related to feature availability. The primary limitation lies in the reliance on features solely derived from the available descriptions in the I/O report of the light commercial building BAS. This constraint led to the exclusion of potentially valuable data, such as airflows, damper positions, heating and cooling loads on the coils. Expanding the feature set to encompass these additional parameters could enhance the framework's fault detection and diagnosis capabilities. Classification optimization is a second area of improvement. The classification methods employed on labeled data, i.e. Decision Trees, while effective, may benefit from further optimization. Exploring alternative techniques such as Random Forests, Support Vector Machines (SVM), and Artificial Neural Networks (ANN) could potentially yield even higher performance levels. These methods can be evaluated to identify the most suitable approach for minimizing the FAR and MAR and improving overall system accuracy.

Taking everything into account, The AFDD framework demonstrated promising and repeatable results in a real-world light commercial building, highlighting its effectiveness and reliability.

Chapter 4: Unsupervised AFDD Method

The main goal of this section is to develop a generalizable unsupervised AFDD method tailored for light commercial buildings. This method aims to be generally applicable across various buildings within this class, capable of utilizing different types of historical data from the BEMS systems, even if it is unlabeled, inconsistent, or fragmented. Additionally, the AFDD method should provide minimal yet essential information about faulty conditions to the HVAC operators. Thus, the objectives of this chapter can be mainly divided into the following:

1. Develop a data cleaning process to identify the most suitable subset for training a PCA-based time series FDD and train a PCA on the selected subset.
2. Establish a PCA-based time series fault detection method using historical raw unlabeled data.
3. Develop an automated method to identify the primary sources of faults and their locations.
4. Assess the method's applicability and effectiveness in various light commercial building's HVAC configurations.

The reason for selecting PCA among other techniques lies in its ability to satisfy the main goals and objectives of the research. While newer methods like encoding also use Reconstruction Error (RE) for fault detection, they have high computational costs and are typically limited to the detection phase [108]. Besides, PCA-based fault detection, with minor adjustments and tuning, can be easily transferred between different datasets and buildings and can also be used for diagnosis as well as detection. Most existing BEM systems have limited computational and storage resources, and PCA-based AFDD can be easily implemented using simple dependencies and basic Python libraries (such as the Pandas and NumPy).

4.2. Methodology

The proposed method for AFDD in this chapter is divided into three steps based on the objectives. Each step is presented separately, and the combination of all these steps leads to the comprehensive AFDD framework as it illustrated in Figure 16. Proposed AFDD framework. The methodology starts with data cleaning for training a PCA to be used for FDD. It then continues with the proposed method for fault detection, and finally, the diagnosis part will be presented.

4.2.1. Data cleaning and PCA training

Dealing with long-term, raw, unlabeled data from existing light commercial buildings presents challenges due to the system's dynamics, varied modes of operation, maintenance breaks, changing weather conditions, and firmware updates. Such data inherently contains diverse patterns, outliers, missing data, and inconsistencies. Consequently, manually identifying a comprehensive set of samples for PCA training that encompasses all periods is nearly impossible. Several preprocessing steps are essential for PCA, including removing missing values, scaling, and centering the data.

A cleaning loop has been devised to identify an appropriate dataset for training. This loop uses PCA and SPE to remove days with significant deviations from the dataset. The variability in system behavior significantly influences the minimum SPE, where days with greater deviation exhibit higher minimum SPE values. The cleaning process begins by determining the necessary number of PCs to achieve a Cumulative Variance (CV) representing the percentage of explained variance by the selected PCs. The minimum acceptable CV can vary between 75% to 90% for different purposes [98]. This study, however, sets the minimum CV, through trial and error, at 92% as it presented promising results during the tuning phase. The proposed process continues by applying PCA to the initial dataset and generating an SPE time series. This series is then resampled daily with the minimum daily SPE values. Outlier days are flagged for removal from the dataset employing a statistical outlier detection mechanism using a threshold as defined by (5), outlier days are flagged for removal from the dataset.

$$Threshold_{SPE}^{Daily} = SPE_{min}^{Daily} + m * SPE_{std}^{Daily} \quad (5)$$

Since SPE_{min}^{Daily} may not normally be distributed in all cases, the value of m, set at 3 for this study, can be adjusted based on the dynamic characteristics of the building's HVAC operations and skewness of the SPE_{min}^{Daily}. If any outlier days are identified, all data corresponding to those days are excluded from the initial dataset. PCA and minimum daily SPE analysis are subsequently applied to the refined dataset iteratively until no outliers remain. These days were removed from the initial dataset to find an appropriate training subset for the primary training of PCA for FDD. Later, during the test steps, they can be considered as a subset to determine whether there exist any faults during those days by using them as the inputs of the fault detection part. Upon completion of the first objective, the final trained PCA model after complete cleansing is saved for use in subsequent fault detection phases. The cleaning and training processes are illustrated in Figure 16.

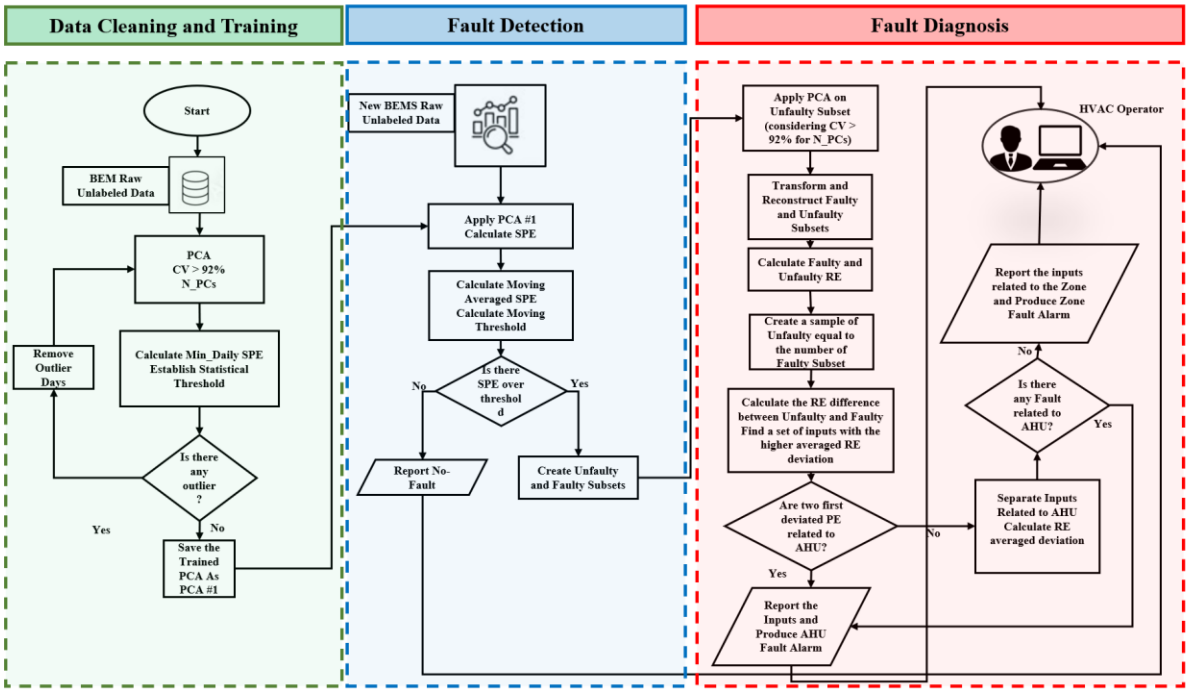


Figure 16. Proposed AFDD framework.

4.2.2. PCA-Based Fault Detection

This section's target is to detect faulty patterns in the time series data of BEM, considering the historical behavior of HVAC systems in light commercial buildings. To achieve this, a smoothing method for the SPE is employed, allowing the analysis of a specific period (window size) of previous data to identify any faulty patterns in the current time.

The detection phase begins by applying PCA #1 (the trained PCA from the previous step) and calculating the SPE of the input data. The input data can be the entire initial dataset, a specific period of the initial dataset that is of particular interest, or a new period of raw data not included in the initial training phase. This process generates an SPE time series corresponding to the original timestamps of the data. Then, a simple moving average (SMA) of SPE in the current time step (SPE_{ma}^{t}) for a window size “ws” can be calculated using equation (6):

$$SPE_{ma}^{t} = \frac{1}{ws} \sum_{i=t-ws+1}^{t} SPE_i \tag{6}$$

Besides, the Moving Standard Deviation (MSD) of SPE in the current time step (SPE_{msd}^{t}) can be calculated as shown in equation (7):

$$SPE_{msd}^{t} = \sqrt{\frac{1}{ws} \sum_{i=t-ws+1}^{t} (SPE_i - SPE_{ma}^{t})^2} \tag{7}$$

The dynamic band of the system adopted from Bollinger Bands [109] can be determined through the upper band (UB) and lower band (LB) as shown in equation (8):

$$\begin{aligned} UB &= SPE_{ma}^{t} + m_{std} * SPE_{msd}^{t} \\ LB &= SPE_{ma}^{t} - m_{std} * SPE_{msd}^{t} \end{aligned} \tag{8}$$

Where m_{std} is a tunable multiplier that can be selected based on the desired sensitivity to deviations. However, in this study, $m_{std} = 3$, which is a standard statistical selection.

Where m_{std} is a tunable multiplier that can be selected based on the desired sensitivity to deviations. However, in this study, $m_{std} = 3$, which is a standard statistical selection.

The primary threshold for fault detection is set based on SPE_{ma}^{t}. While traditional statistical methods based on confidence levels or higher quantile percentages are common, they may not effectively account for the smoothness of SPE_{ma}^{t} in this case, potentially leading to high false alarm rates during fault-free conditions. Hence, this study proposes using equation (9) to determine the main threshold ($Threshold_{SPE}$) for fault detection. It is important to note that the parameter m_{thrsh} can be adjusted during the tuning phase for each specific building regarding HVAC operators' considerations. To avoid missing alarms, m_{thrsh} should not be set too high relative to m_{std}. Ideally, it should be set between 10% to 20% of m. For this study, an acceptable value considered $m_{thrsh} = 0.5$ which is 16.7% of m_{std}.

$$Threshold_{SPE} = SPE_{ma}^{t} + m_{thrsh} * SPE_{msd}^{t} \quad (9)$$

Any SPE_{ma}^{t} above the $Threshold_{SPE}$ will be flagged as a potential fault. Then faulty and unfaulty samples will be labeled and separated. Additionally, any SPE_i outside the dynamic band of the system may be considered outliers; however, the combination of moving SPEs above the threshold and SPEs outside the dynamic band can indicate a higher level of fault severity.

4.2.3. PCA-Based Fault Diagnosis

After detecting a fault, the fault diagnosis process relies on three key aspects: localization, identification, and severity assessment [110]. Identifying and localizing the root cause of an occurred fault is generally more difficult than detecting the fault itself. Various faults can produce similar symptoms among different systems even in a specific HVAC system. Accurate diagnosis of the root cause usually demands an in-depth understanding of the HVAC configuration and control strategies, which are unique to each building [14]. However, as a generic AFDD framework applicable to different light commercial buildings, the fault diagnosis in the proposed method focuses on identifying problematic inputs (sensor data, control values, etc.) and pinpointing the location (level of HVAC system) of the fault within the HVAC system (either at the central system or in the zones). Additionally, each fault is classified into two levels of severity. This information is then provided to HVAC operators, enabling them to make informed decisions and accurately identify the fault.

The diagnosis process starts with applying PCA to the subset of data identified as unfaulty during the fault detection phase. By transforming and reconstructing each of the unfaulty and faulty subsets using the PCA model trained on the unfaulty samples, the reconstruction error matrix for each subset can be calculated. Using the principal concepts of Reconstruction Error (RE) illustrated in Figure 1, the RE of unfaulty and faulty samples should exhibit distinct differences. The average RE deviation between the unfaulty and faulty subsets can provide insights into the specific inputs causing the fault. If the inputs with high RE deviations belong to the central system level (e.g.,

Roof Top Unit, AHU), these inputs will be reported, and the fault location will be identified as the 'central system level, otherwise, it will be considered as 'zone level'. For inputs associated with zone-level faults, a further investigation is required. The diagnosis process will be reapplied solely to inputs belonging to the central system. The Deviated RE of inputs corresponding to the source will be calculated. If a Deviated RE input related to the central system is detected, the fault location will be reported at the central system level. Otherwise, the fault will be noted at the zone level, specifying the faulty zone based on the affected inputs.

In the end, if only the SPE_{ma}^{t} is above the threshold, the severity is considered to be low (level 1). If not only the SPE_{ma}^{t} is above the threshold but SPE_i values are also outside the dynamic range, then the fault severity is considered to be high (level 2). The reported inputs, fault location, and severity levels provided to the HVAC operators will enable them to conduct further investigations. Using their experience with the system, they can identify its specific faults.

An important consideration is the amount of data required for the method. To address this, various data durations were chosen for training and testing purposes. Based on the findings of this study testing on two different case studies, a minimum of 3 months of data is necessary for the training phase. For testing (operational purposes), the amount of data required for each execution should be at least one step larger than the window size used for the moving SPE. For instance, if a 15-day window is employed, at least 16 days of data are needed. During the initial 15 days, the method learns the system dynamics but does not detect faults within this period.

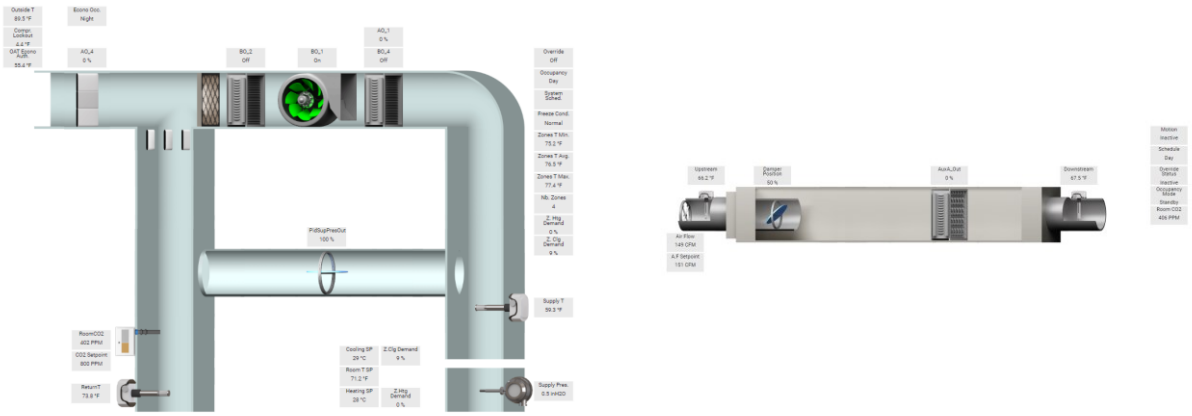
Finally, it is important to mention that based on the computational limits of the existing building system, the method is suggested to be executed once per day. The specific time for execution can be chosen based on the HVAC operator's preference and operational requirements.

4.3. Validation Case Studies

The HVAC system of a typical light commercial building in Montreal, Canada, has been selected as the primary building to validate the presented method. This choice was made because detailed information about the system and access to HVAC operators were available. The system includes

a single-duct electrical heating AHU combined with four VAV boxes with electric coil reheating capability. The AHU's electric heating coil has a maximum heating capacity of 21.2 kW. The supply fan can deliver air at a maximum rate of 2500 cubic feet per minute (CFM) and is powered by a 1.5 horsepower (HP) motor. The system's cooling capacity is comparable to 6 tons of refrigeration. The heating coil's steady-state efficiency is between 80.5% and 81.1%, and the direct expansion cooling coil's efficiency stands at 80%. In terms of energy efficiency, the system has an Energy Efficiency Ratio (EER) of 12 during cooling and a Seasonal Energy Efficiency Ratio (SEER) of 14. Given the nature of the data and the extensive period it covers, this building serves as an appropriate example for developing and evaluating AFDD methods in light commercial structures. The floor plan and the schematic of the HVAC system given from OCN+ BEMS of the building are presented in Figure 17. OCN+ is the building energy management system of the building of interest. It provides information about sensor locations and descriptions, and it can be used to monitor data trends and export data.

Floor plan of the interested light commercial building.



(b)Main AHU (c) VAV box of each zone

Figure 17. The Floor plan (a) and HVAC system (b, c) of the studied light commercial building (From OCN+ BEMS system).

Data collection was conducted utilizing the OCN+ BEM system of the building, spanning from January 8, 2022, to March 18, 2023. The selection of input data for the AFDD was informed by a detailed analysis of sensor descriptions within the OCN+ BEMS framework. This selection process was guided by consultations with HVAC system operators and the incorporation of critical variables as recommended by the ASHRAE Guideline 36 for high-performance sequence control [111].

The data frame encompasses sensor data, including analog values, as well as control values, which consist of both binary and analog types. Additionally, three calculated features have been integrated into the dataset. The first is the total airflow rate of the AHU, determined by the continuity (the conservation of mass) equation. The second and third features are the average room temperature and average supply temperature across the zones, respectively. With these features included, the total number of distinct features in the data frame is 94, which are referred to as inputs of AFDD. Table 6. The description of input data for the AFDD method.. Introduced all the inputs separated by the level (location) of each and (i) refers to the zone assigned number.

Table 6. The description of input data for the AFDD method.

Input Parameter	Feature	Type
Central system Level (AHU)		
Airflow	airflow_AC	Calculated
Supply temperature	supply_temp_AC	Sensor_Analog
Return temperature	return_temp	Sensor_Analog
Outdoor Air Temperature	OAT	Sensor_Analog
Room setpoint AC	room_sp_AC	Control value
Supply pressure	pressure	Sensor_Analog
Damper positions	dmp_pos_AC	Control value
Heating loop	heating_load_AC	Control value
Cooling loop	cooling_load_AC	Control value
Fan mode	fan_ac	Control_Binary
Cooling mode	cooling_ac	Control_Binary
Heating mode	heating_ac	Control_Binary
Cooling fan demand	cooling_fan_demand	Control_Binary
Heating fan demand	heating_fan_demand	Control_Binary
Active heating room temperature setpoint AC	act_htg_RTSP_AC	Control value
Active cooling room temperature setpoint AC	act_clg_RTSP_AC	Control value
Sandby mode room temperature setpoint cooling offset	STBmode_RTSP_clg_offset	Control value
Standby mode room temperature setpoint heating offset	STBmode_RTSP_htg_offset	Control value
Zones Level		
Airflow	airflow_z(i)	Sensor_Analog
Room Temperature	room_temp_z(i)	Sensor_Analog
Average Room Temperature	Room_temp_avg	Calculated
Supply Temperature	supply_temp_z(i)	Sensor_Analog
Average Supply Temperature	Supply_temp_avg	Calculated
Heating loop	heating_load_z(i)	Control value
Damper Position	dmp_pos_z(i)	Control value
Room temperature setpoint cooling occupied	RTSP_clg_occ_z(i)	Control value
Room temperature setpoint heating occupied	RTSP_htg_occ_z(i)	Control value
Room temperature setpoint cooling un-occupied	RTSP_clg_unocc_z(i)	Control value
Room temperature setpoint heating un-occupied	RTSP_htg_unocc_z(i)	Control value
Motion sensor	MotionSens_z(i)	Sensor_Binary
Occupied mode	OccMode_z(i)	Control value
Active heating room temperature setpoint	act_htg_RTSP_z(i)	Control value
Active cooling room temperature setpoint	act_clg_RTSP_z(i)	Control value

Active flow setpoint	act_FlowSP_z(i)	Control value
Minimum Airflow heating	MinAF_htg_z(i)	Control value
Minimum Airflow cooling	MinAF_clg_z(i)	Control value

Since one target of this study is to investigate the generalizability of the proposed method for light commercial buildings, a second case example has been chosen from the reference [19,95]] with the dataset available in [112] as the secondary dataset for validation. This dataset, collected from a small industrial facility in Ireland, serves as an excellent test case for generalizability since the building can be considered representative of a light commercial building. Additionally, the dataset, obtained from a real BMS system, includes all the typical challenges found in real data used for validation of the method, such as broken data (missing data points), poor quality data (noisy sensor data), and importantly, the dataset is unlabeled. Table 7. Comparison between validation data and generalizability analysis data.. Comparison of Data Used for Validation and Generalizability Analysis. The configuration and feature description of the dataset have been detailed in [112].

Table 7. Comparison between validation data and generalizability analysis data.

Characteristics	Industrial facility, Ireland	Office building, Montreal
Location	Ireland	Canada
Nature of data	Unlabeled, Raw data	Unlabeled, Raw Data
Data Quality	Many missing, Inconsistency	Many missing, Inconsistency
Time step	15 Minutes	5 Minutes
System Configuration	AHU + 2 VAVs	AHU + 4 VAVs
Heating loop	Water heating coils	Electric heating/reheating coils
Cooling loop	Water cooling coil	DX cooling coil
Performance characteristic	Energy Consumption available	Energy Consumption unavailable

Number of inputs	20 Inputs	94 Inputs

4.4. Results and Discussion

This section presents and discusses the results of the proposed method applied to the primary and secondary case study buildings. The section starts with the data cleaning and training phase and continues with fault detection and isolation. The results begin with the comprehensive framework outcomes for the primary building, selected due to extensive information obtained from HVAC operators. Following this, AFDD results from the secondary building are presented as an illustrative example.

4.4.1. Data Cleaning and PCA Training

The number of principal components (PCs) required for subsequent steps was primarily determined as shown in Figure 18. Cumulative Variance Explained by Number of PCs. The results indicate that a minimum of 17 PCs is necessary to meet the framework's requirements. The cumulative variance explained of 0.926 for 17 PCs signifies that these components capture 92.6% of the total variance in the dataset. Therefore, the next step involves utilizing these 17 PCs for the cleaning and training processes.

Figure 19. The results in internal steps of the cleaning process using PCA loop.. Shows the initial step, two interval steps, and the final step of the data cleaning process for training the PCA. The horizontal axis represents the days within the dataset, and the vertical axis shows the minimum daily SPE. It is evident that the threshold for the minimum daily SPE decreases with each step of the loop. The maximum acceptable of minimum daily SPE for training the PCA is calculated to be 0.23. The entire cleaning process took 9 iterations, during which 27.5% of the data was removed from the initial dataset.

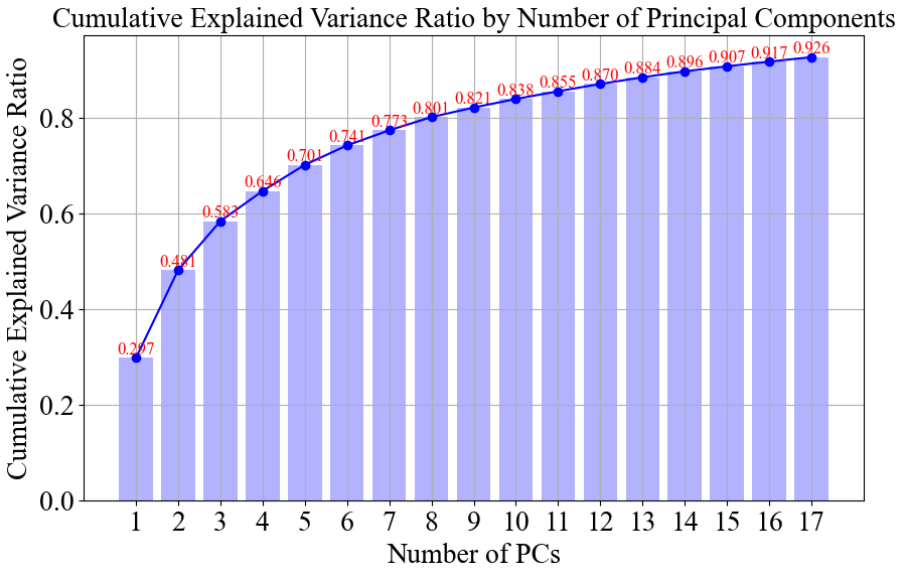


Figure 18. Cumulative Variance Explained by Number of PCs.

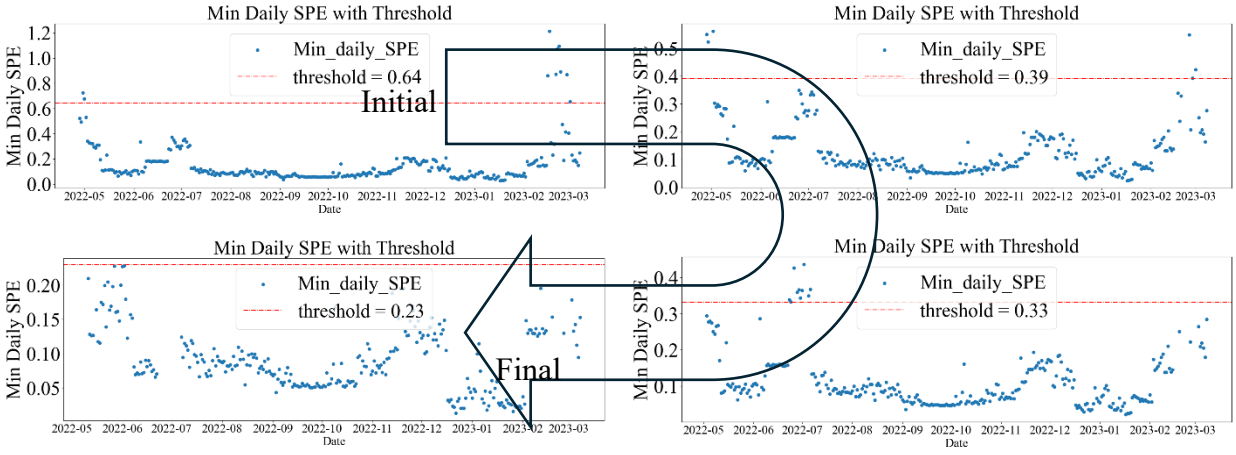


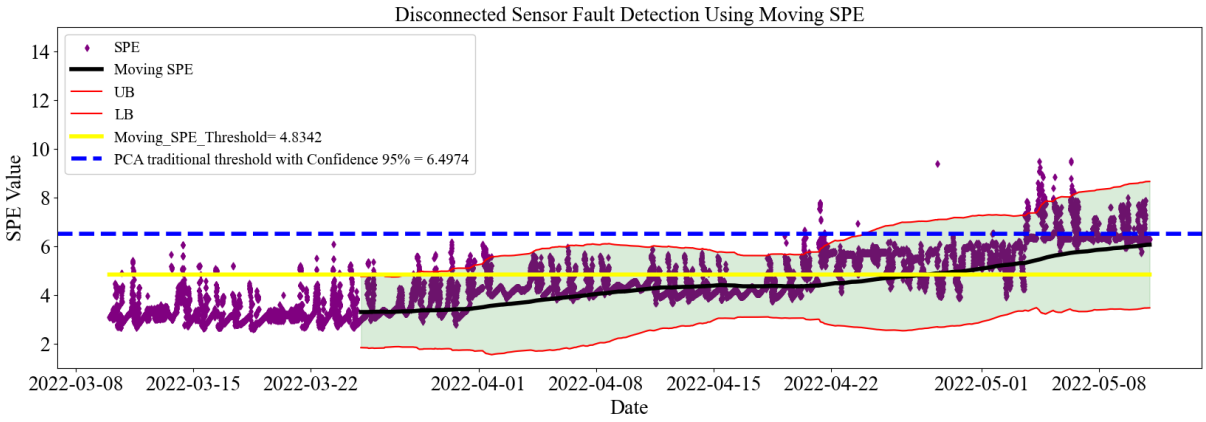
Figure 19. The results in internal steps of the cleaning process using PCA loop.

4.4.2. Fault Detection and Isolation in Primary Building

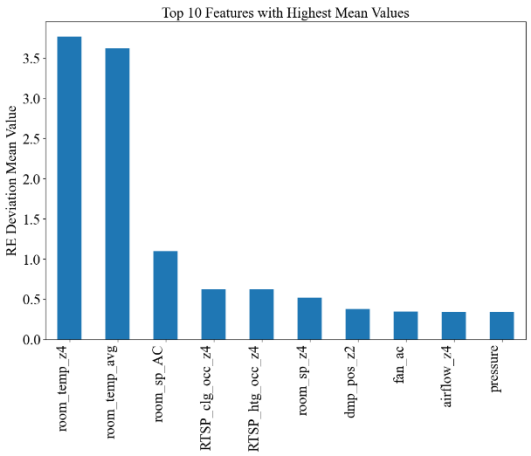
This section provides the results of fault detection and isolation using a dataset spanning two months. It discusses two examples of faults within the initial dataset and an additional example of an unseen fault outside the initial data. In all results figures, the green band indicates the dynamic

range of the system, with the black line representing the moving SPE and the yellow line indicating the moving SPE threshold. Additionally, the traditional SPE threshold using a 95% confidence level is shown by the dashed blue line.

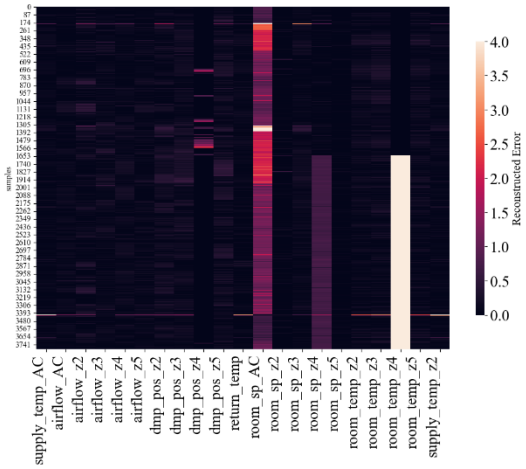
The first example of a faulty pattern was detected on April 27, 2022, as shown in Figure 20(a). The two inputs most contributing to this fault were room_temp_z4 in the level of the zones, as indicated by the mean value of reconstructed errors in Figure 20(b). Figure 20 (c) shows that the deviation of room_temp_z4 was constant as depicted in Figure 20(d). With this provided information, system operators confirmed this condition as a sensor disconnection for room_temp_z4, where the sensor data reached its maximum allowable limit in the system. A comparison between the date of sensor disconnection and fault detection revealed that the crossing of the moving SPE and threshold occurred before this incident, which is attributed to the fluctuating behavior of room_sp_AC during a system reconfiguration period. Moreover, the method successfully detected and isolated the fault related to room_temp_z4 disconnection, situated at the zone level. Furthermore, the analysis indicated that following the sensor disconnection, this fault significantly impacted the system. This finding was corroborated by the presence of outliers outside the dynamic system band concurrent with the sensor disconnection fault, thereby elevating the fault severity to level 2, as detailed in the methodology section. Compared to traditional PCA, the proposed approach can detect deviations in room_sp_AC, indicating sensor disconnection.



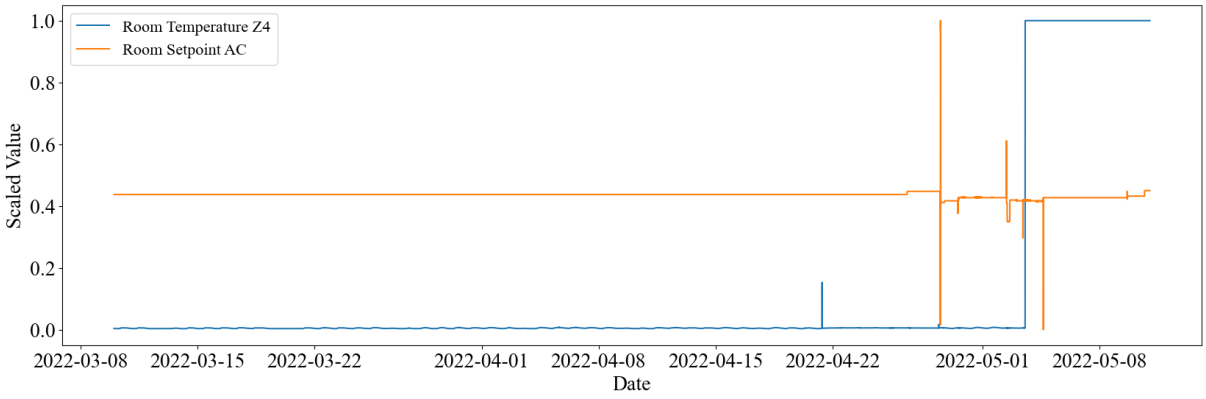
(a) Detection Fault, April 27, 2022



(b) Averaged RE deviation of inputs



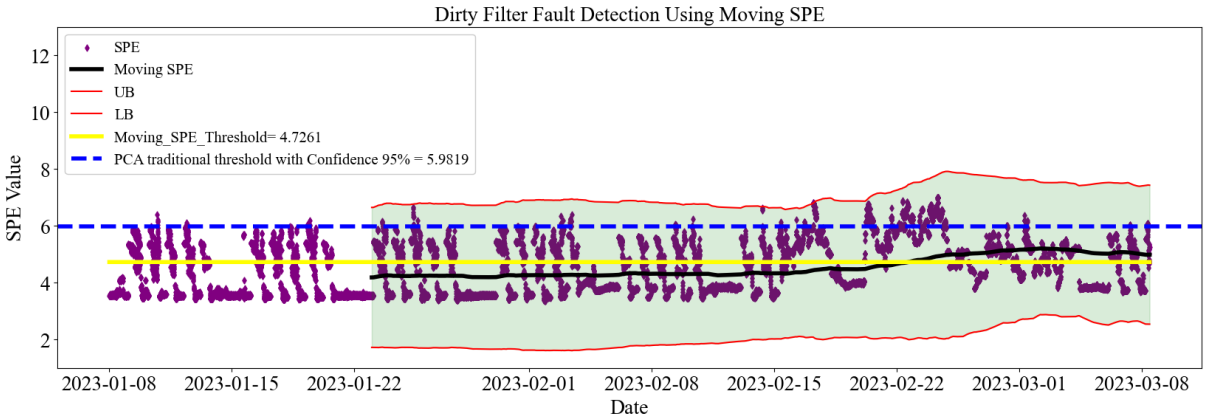
(c) RE deviation of samples



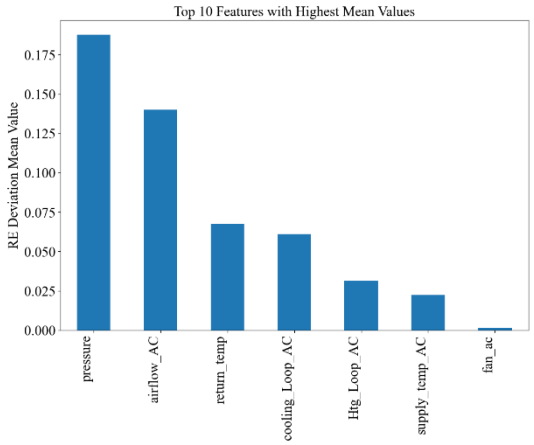
(d) Behaviors of inputs with the highest averaged RE deviation

Figure 20. Disconnected sensor fault on 2022-04-27.

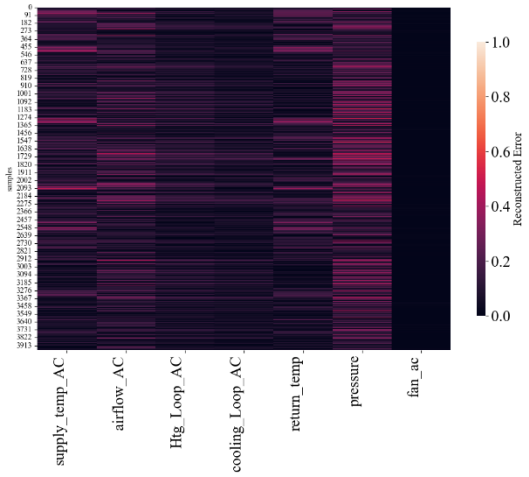
As another example of a different source, the fault triggered on 2023-02-22 is shown in Figure 21 (a). The initial inputs contributing to the reconstruction error deviation were the room temperature setpoints for zone 2 and zone 5, prompting further analysis of the AHU. The AHU analysis also indicated a fault, with significant deviations in pressure and airflow as shown in Figure 21 (b) and (c). Based on the results, the location of the fault was reported at the level of the AHU, with pressure and airflow identified as the two inputs with the highest deviation. This led HVAC operators to confirm a dirty filter fault. As shown in Figure 7(d), the deviation of the setpoints is much higher than that of the pressure and airflow due to the normalization process before PCA analysis. This indicates that the dirty filter significantly impacted the thermal comfort of the occupants by decreasing the airflow, leading to a setpoint override to improve comfort. This override had a substantial effect on the entire system, making fault detection easier. Additionally, it's important to note that all four highly contributing inputs showed decreased values, demonstrating that the method is effective in identifying faults in both increasing and decreasing input behaviors. Furthermore, since there are not any outliers beyond the dynamic band (when the moving SPE is above the threshold), the severity of fault is reported as level 1. Finally, although the method shows a delay in detecting the dirty filter fault, it is much more effective than traditional PCA, which almost missed this fault entirely.



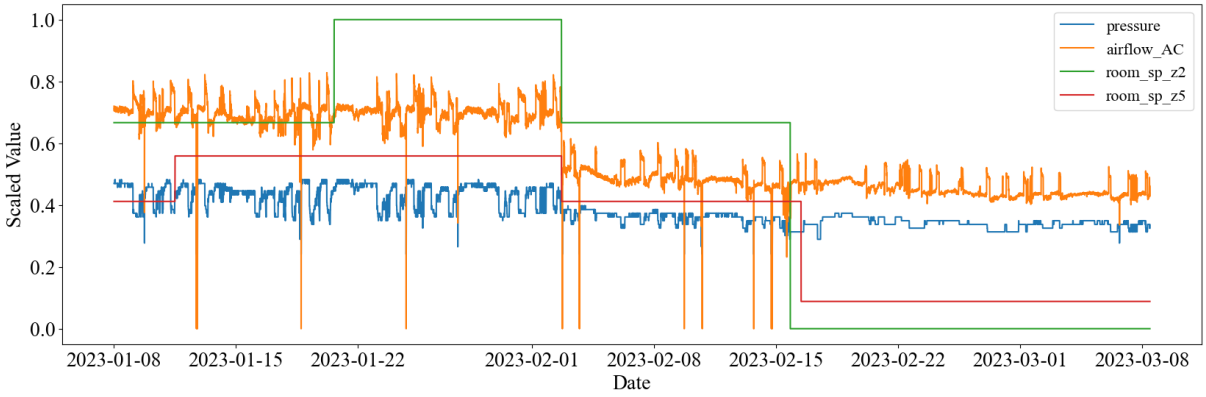
(a) Detection Fault, on 2023-02-22



(b) Averaged RE deviation of inputs



(c) RE deviation of samples

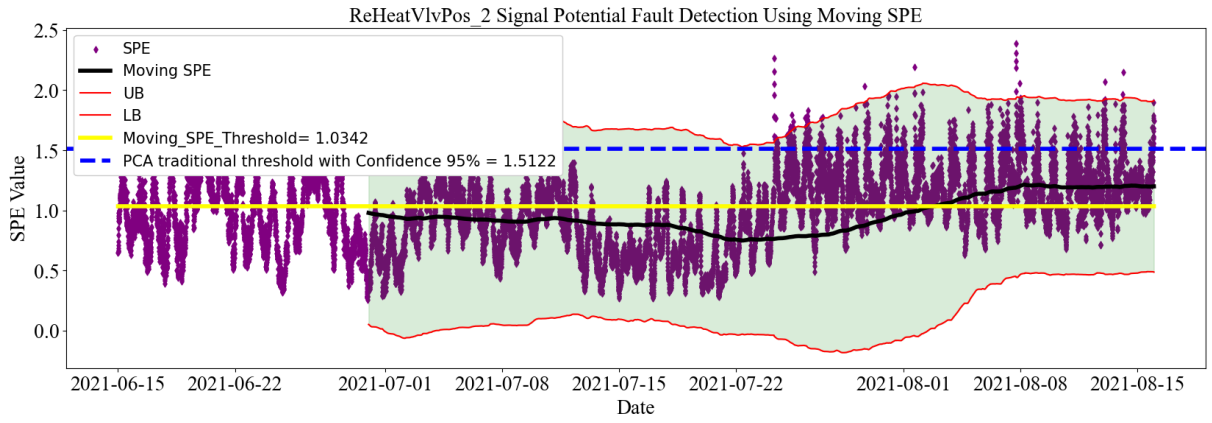


(d) Behaviors of inputs with the highest averaged RE deviation

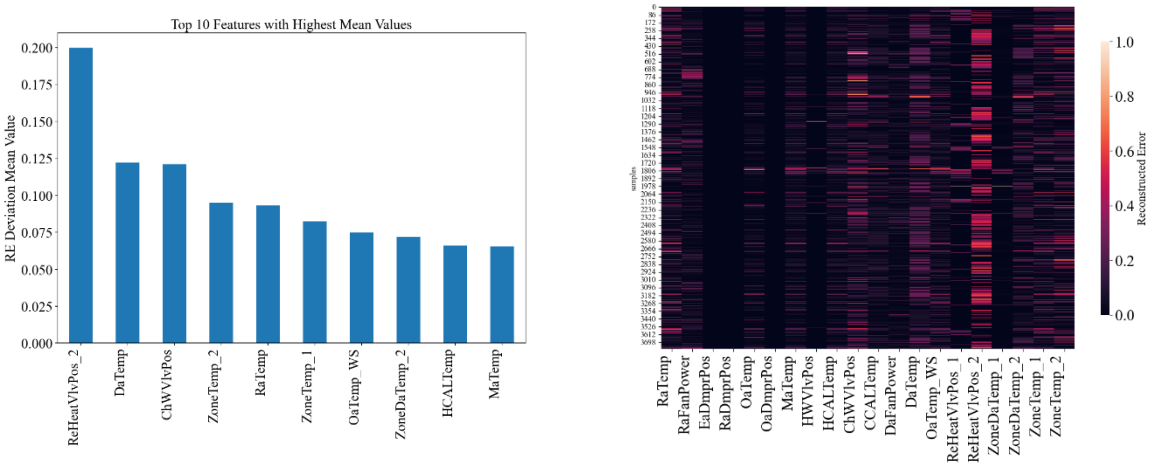
Figure 21. Dirty Filter Fault detected on 2023-02-22.

4.4.3. Fault Detection and Isolation in Secondary Building

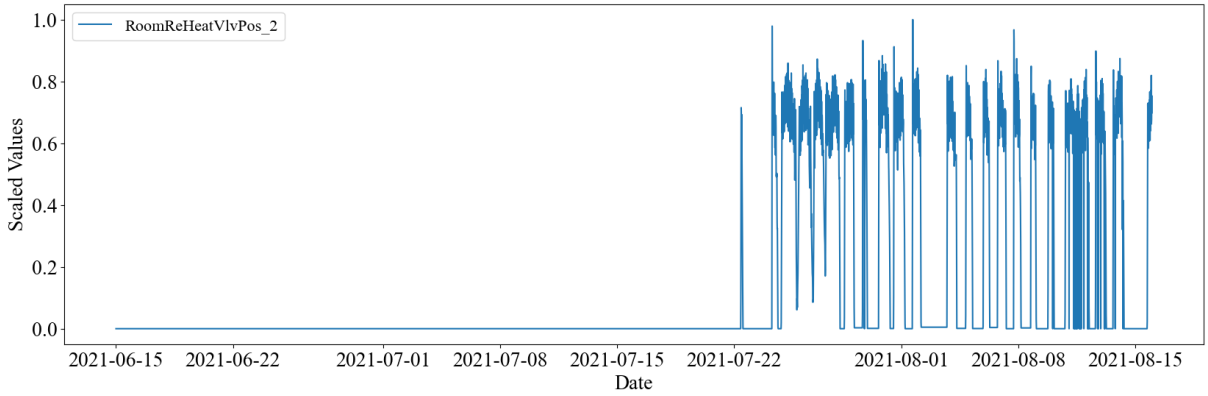
The entire method was applied to the secondary dataset, and the results have been obtained. The testing of the technique showed promising outcomes. A good example of the model's performance is when a fault was triggered on 2021-08-02, as shown in Figure 22 (a). The most contributing input was reported as ReHeatVlvPos_2, which is the reheat valve position signal for zone 2, based on Figure 22(b). The RE has fluctuations in Figure 22 (c), indicating that it is not stuck. Additionally, there is no deviation in the inputs related to the AHU, the fault is determined to be at zone 2, with the problematic input being ReHeatVlvPos_2. Although the location and inputs should be reported to the HVAC system operators to decide if there was a fault at that time, the method shows that it can detect deviations, isolate the problematic input, and determine the level of divergence.



(a) Detection Fault, on 2021-08-02



(b) Averaged RE deviation of inputs (c) RE deviation of samples



(d) Behaviors of inputs with the highest averaged RE deviation

Figure 22. Generalizability analysis of the method.

Regarding AFDD results from both the primary and secondary buildings, the method demonstrated promising effectiveness in detecting and isolating faults using the same tuning parameters, including window sizes, m, m_{std}, and m_{thresh}. It should be noted that the secondary building's data was resampled using the forward filling method to match the primary building's data time step of 5 minutes. Furthermore, the results indicated that different data from various HVAC configurations and components of light commercial buildings can be effectively used as inputs for the method, showcasing its adaptability to different types and volumes of data. However, the method still

requires HVAC operators to confirm faults based on the information provided by the AFDD method. In both buildings, AFDD exhibited a specific limitation when dealing with faults of lower severity. Since the method is designed to be executed once a day, the detection time can vary. For instance, immediate detection occurred with the first execution of the method for the Disconnected Sensor fault in the primary building, while detection took up to six executions for the ReHeatVlvPos_2 signal fault in the secondary building. This indicates that the method can detect condition-based faults when they begin to significantly impact the behavior of the HVAC system. For both buildings, the diagnosis results were highly precise. As soon as the fault was detected, the method was able to accurately isolate the source of the detected faults (or potential faults in the secondary building) without exception.

4.5. Conclusion

In conclusion, this study has developed an unsupervised PCA-based AFDD tailored for application in light commercial buildings. The method is designed to detect and diagnose faults using raw, unlabeled data from Building Energy Management Systems (BEMS), showcasing its ability to work with datasets from existing buildings without any further pre-investigation and labeling. The data can be exported from BEMS and directly used in the AFDD method. Validation was conducted on a light commercial building in Montreal, Canada, and its transferability and generalizability were confirmed by successfully testing it on another building in Ireland. The method's transferability and generalizability allow it to be applied to different configurations of HVAC systems with varying numbers of inputs, ensuring effective performance across diverse HVAC setups. Furthermore, the method demonstrated promising performance in detecting and isolating faults both in initial and unseen datasets. Different faults related to sensors and system conditions have been detected and isolated using this method, illustrating its capability beyond sensor faults. Generally, any fault that affects the system's behavior can be identified and isolated using this approach.

The most important aspect of using this method is choosing the appropriate tuning parameters. A good set of tuning parameters can be selected by HVAC consultants or operators through initial

trials on the dataset to optimize fault detection. While the method automates fault detection, isolates faulty inputs, and identifies fault severity levels, some limitations can be listed as follows:

- The involvement of an HVAC operator is always necessary for comprehensive diagnosis and final confirmation.
- The method requires at least three months of data to begin functioning and is not applicable during the initial running period of the system or after major system changes until sufficient data accumulates. However, additional operational data can always be appended to the initial dataset, providing a more comprehensive dataset for the training phase.
- During the fault detection phase, a single day of data is insufficient; at least n+1 (which n is the window size in days) consecutive days of data are needed when the method is detecting faults. The method cannot detect faults within the first n days, as it relies on a window size of n days to capture the dynamics of the system.
- If there is a continuous fault in the testing dataset that started before the date of the first sample and persisted until the test date, the method will not be able to detect it, as the system's dynamics will remain unchanged.

However, this method offers a straightforward AFDD solution with minimal computational overhead, suitable for diverse HVAC systems and configurations in light commercial buildings. One of the significant advantages of this AFDD method is its ability to utilize an unlabeled raw dataset without any prior information about the faults and their severity, making it highly adaptable to different scenarios. Unlike previous approaches that require a set of unfaulty samples for PCA training, this method operates without pre-known unfaulty samples, enabling execution with minimal information about the system condition. This flexibility ensures that the method can be readily applied in various situations, even when detailed system information is lacking. Furthermore, while earlier methods focused solely on steady-state conditions, which increases false alarms during dynamic changes, this method incorporates historical dynamic information of the system, utilizing both dynamic and steady-state conditions. This comprehensive approach

significantly enhances fault detection accuracy, reducing the likelihood of false alarms and ensuring reliable performance in real-world applications.

In summary, this method presents an efficient AFDD solution that leverages minimal information to deliver accurate and reliable AFDD. Its applicability across various HVAC configurations, combined with its potential to reduce operational costs and improve system efficiency, makes it a valuable tool for HVAC operators in light commercial buildings. All light commercial building HVAC operators can benefit from the ability of this method to detect and isolate faults efficiently not only enhances system performance but also contributes to energy savings and cost reductions, making it an economically beneficial solution. Additionally, this method can facilitate dataset labeling for developing supervised AFDD methods and supports semi-automatic commissioning of HVAC systems.

Chapter 5: Inverse Modeling

The primary goal of this chapter is to develop a physics-based inverse model for AHU systems to establish a baseline for heating load to find the anomalies (potential faults). This overarching goal can be divided into several specific objectives:

- Defining a physical model using variables associated with AHU operation.
- Establishing baselines for steady-state and transient conditions using curve fitting techniques like Genetic Algorithm(GA) optimization, ANN optimization with callbacks, and Physics-based Neural Network (PBNN) models.
- Employing the established baseline to detect potential faults presented as anomalies for further investigations.

5.1. Methodology

AHUs operation can mostly be analyzed as steady-state processes. However, applying the physical meaning of steady-state conditions to sensory data is challenging. Therefore, various statistical methods have been developed to detect operational conditions close to steady state. This study employed the steady-state detector introduced by Lee W. et al [104]. Equation (10) has been used to calculate the slopes of the variables. The slope value equal to zero indicates a fully steady-state condition, although values close to zero are also considered acceptable. In this study, the steady-state threshold is set at 0.25 times the standard deviation of the sum of slopes, ensuring continuous steady-state periods. Increasing it results in more isolated single samples as steady-state conditions. Also, the window size is considered 15 minutes.

$$S = (V_{max} - V_{min}) / V_{avg} \quad (10)$$

S is the slope of variable V in a specific window size. The AHU represents a typical light commercial building in Montreal, Canada, with an electrical heating coil (21.2 kW). Heat transfer by the fan was disregarded during modeling. Applying conservation laws yields equation (11) for

steady-state conditions, with variables X and f unknown. Due to unavailable humidity sensor data, air density and heat capacity were treated as unknown parameters.

$$\%htg\_load = X.\ LPS.\ (\ T_{sup} - [f\ .\ T_{OA} + (1-f)\ .\ T_{ret}]) \tag{11}$$

f presents the fresh air fraction, and X represents the amount of heating load percentage that should be added for 1 LPS of supply airflow when there is a 1 degree of Celsius of change between supply and mixed air temperature. %htg_load represents the heating load on the AHU. LPS, T_{sup}, $T_{OA,}$ and T_{ret} are AHU airflow in liter per second, supply temperature of AHU and return temperature, respectively. The primary challenge involves an optimization problem: determining the values of X and f to minimize the difference between the estimated and measured percentage of heating load on the AHU coil. To address this, the fitness function has been defined as the mean square error (MSE) to facilitate minimization.

Initially, a GA with different hyperparameters was employed to examine the sensitivity of the results to hyperparameters. For the final results, a generation size of 500 and a population size of 100 were selected to examine the unknown X and f. Secondly, an optimization using ANNs' callbacks was employed to calculate the values of X and f. Lastly, the ANN model that was utilized for optimization was also employed to establish the baseline. It is important to note that the first and second methods examined the unknown parameters and replaced them in equation (2) to define a baseline, while the last method utilized the ANN model directly for baseline definition. Since the physical information based on equation (2) has been applied to the ANN, it can be considered a Physics-Based Neural Network (PBNN) [6]. The ANN utilizes X and f, which are calculated based on optimization, and replaces them into equation (2). Subsequently, the model is trained to minimize the MSE. AHU data was collected from the Building Energy Management System during the period of January 18, 2022, to March 31, 2023. Before analysis, inconsistencies related to sensor faults were removed from the dataset. Additionally, the accuracy of outdoor sensor records was validated using data from the nearest weather station. Since the airflow of the AHU was not

directly measured, the airflow rate was estimated by summing the airflow rates of connected Variable Air Volume (VAV) systems and converted to Liter per minute unit for more consistency.

5.2. Results and Discussion

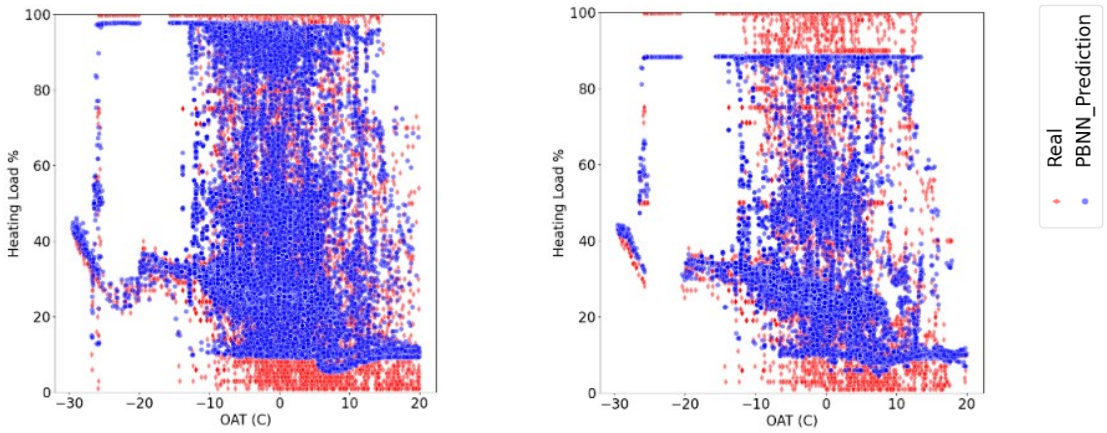
This section presents the results of the methodology and discusses them. By applying equation (10) to the data, steady-state conditions were detected. While reaching steady-state conditions depends on various parameters such as outdoor conditions, schedules, and sequence of operation, the results indicate that it mostly occurred between 10 AM to 3:30 PM when the AHU was operating in heating mode. Regarding the results of the inverse modeling presented in Table 8, the PBNN model exhibited superior performance in terms of R^2 and Mean Averaged Error (MAE). Additionally, it was found that considering the steady-state data led to a decrease in the model's performance, except for ANN Optimization. The degradation in model performance can be observed in Figure 23. which presents the expected heating load calculated by PBNN relative to measured heating loads. Although the model used different input variables to make the results more sensible it has been plotted versus OAT which is common in the commissioning stage of HAVC systems. The samples with high values of heating load primarily correspond to unsteady conditions when the temperature gradients are high. By removing this valuable information (in steady-state conditions) from the data, the models lack sufficient information about these samples and cannot be adequately trained for situations with high values of heating load as shown in Figure 23(b).

Table 8. Inverse Modeling Results.

	X	f	R^2	MAE
Whole Data				
ANN_Optimization	0.0693	0.2043	0.6663	10.2596
GA Optimization	0.0578	0.2708	0.6859	10.3337
PBNN			0.9211	4.2121
Steady-state				
ANN_Optimization	0.0374	0.56195	0.7324	9.6225

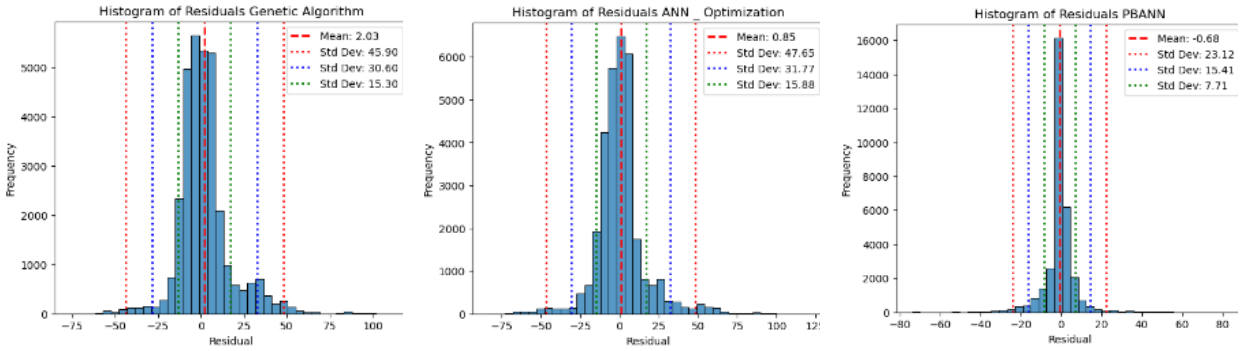
GA Optimization	0.0731	0.19119	0.6899	11.6578
PBNN			0.9124	4.7964

Anomaly detection can be conducted using an established baseline. An anomaly is defined as a specific deviation from the expected (baseline) value. The threshold for this deviation can be set in various ways. Gunay et al. [59], suggest that building operators can define a threshold based on their expertise and knowledge of the system. However, in this study, the threshold is determined using the mean plus three times the standard deviation of residuals, which is a standard statistical method. Figure 24 illustrates the distribution of residuals for each method and potential statistical thresholds. The anomalies detected by the method represent potential faults, most likely related to programming logic or issues with the sequence of operation (other soft and hard faults are also possible). Figure 25. Shows an example period for anomaly detection using the PBNN baseline. These anomalies should be investigated according to the rules provided by ASHRAE Guideline 36 and confirmed by building operators to establish an effective retrofitting or reprogramming plan. As an example, Section 5.14.8.3 ASHRAE Guideline 36 can be utilized to examine programming logic for heating supply air temperature setpoint reset.



(a) (b)

Figure 23. Heating load vs TOA for (a) whole data and, (b) Steady-state conditions.



(a) (b) (d)

Figure 24. Residual distribution: (a)GA Optimization, (b)ANN Optimization, (c)PBNN; whole data.

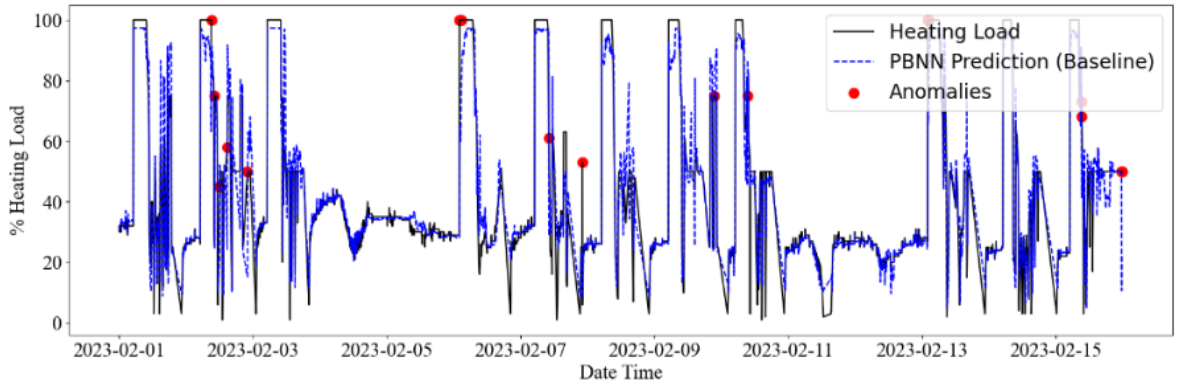


Figure 25. Anomaly detection using PBNN during two first week of Feb 2023.

5.3. Conclusion

This study employed three different techniques to establish a baseline to detect anomalies in the AHU heating load of light commercial buildings. Among the methods, the physics-based neural network demonstrated superior performance in predicting the load and establishing a more fitted baseline. Subsequently, anomalies were detected using the established baseline, representing potential faults in the AHU system. Investigating these anomalies in conjunction with ASHRAE Guideline 36 and system knowledge can reveal their underlying causes, providing robust information on whether these anomalies can be considered faults. In summary, while the physics-based neural network exhibited higher accuracy in load prediction, equations derived from ANN-Optimization and GA-Optimization can also be employed for anomaly detection with their

respective thresholds. In addition, the simple equations extracted from optimization methods are easy to use and implement in HVAC control systems, making them highly practical and valuable for real-world applications in systems with lower computational and memory capacity.

6. Conclusion

In this study, three different methods have been developed for Automated Fault Detection and Diagnosis (AFDD) of light commercial buildings' HVAC systems and validated using data extracted from this class of buildings. The first method starts with data cleansing, followed by dimensional reduction using Principal Component Analysis (PCA), applying Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for anomaly detection, Decision Tree (DT) rule extraction for separating faults from outliers, and finally, classification to produce an AFDD model. The second method is a fully unsupervised PCA-based technique for AFDD, which can operate effectively without any pre-defined unfaulty conditions. The third method includes various inverse modeling techniques for modeling the Air Handling Unit (AHU) and detecting anomalies related to energy consumption and potential fault candidates for the sequence of operations.

To conclude all results, the semi-supervised method provides valuable insights into the system's operating modes. Using this framework enables us to label the dataset for future studies, ultimately producing a supervised AFDD that effectively detects and diagnoses faults that have occurred at least once in the data history. However, to identify and label new faults, the entire process must be re-executed. This method is highly effective for condition-based fault detection but performs less effectively with control faults, which are often related to the sequence of operation and programming errors. Additionally, since the data cleansing process automatically removes outliers and inconsistent sensor values, this method cannot detect sensor faults. A summarized comparison of different methods has been presented in Table 9.

Table 9. Comparison of AFDD methods

	High
	Medium
	Low (or not)

	Unsupervised	Semi-supervised	Inverse Modeling
Condition-based faults FDD performance			
Behavior-based faults FDD performance			
Outcome-based faults FDD performance			
Control faults FDD performance			
Sensor faults FDD performance			
Transparency in fault diagnosis			
Transferability, Scalability, and Generalizability			
Human interference			
A long period of training needs			

The unsupervised method presents a novel approach to fault detection and diagnosis using PCA. It works effectively with various types of data and different HVAC sizes and configurations, making it a scalable and transferable AFDD method applicable to any light commercial building. Although it provides useful information about fault sources, it cannot fully diagnose faults and requires an HVAC operator to observe and diagnose the problematic sources (inputs). It is straightforward to use, without the complexities of labeling and establishing a supervised AFDD as required in the semi-supervised method. This method is capable of detecting and diagnosing faults in historical data as well as new, unseen faults that may occur in future operations. This method can detect and diagnose all the faults that change the historical behavior of the system including sensor faults and conditions-based faults.

The third proposed method uses an inverse model and focuses solely on the AHU as well as its related inputs. It demonstrates that a simple equation derived from PBNN can effectively model the dynamics of AHUs in light commercial buildings and detect anomalies in energy consumption. While further investigation showed some of these anomalies are not related to the sequence of operations, they can still be considered anomalies warranting further investigation and commissioning.

In summary, although each of the proposed methods has limitations, all can be applied to light commercial buildings' HVAC systems either individually or as an integrated package. For instance, the dirty filter fault was not detectable using the semi-supervised method but was effectively detected and diagnosed using the unsupervised method. The dirty filter samples identified by the unsupervised method can then be used in the supervised AFDD method.

Applying these techniques for AFDD offers multiple significant benefits for HVAC systems in light commercial buildings. Firstly, they can reduce energy consumption significantly, which directly lowers the operational costs of running HVAC systems. This reduction in energy use also leads to a decrease in the air pollution and carbon production generated by the energy conversion process, contributing to a more sustainable and environmentally friendly building operation. The validation of these methods in different buildings confirmed their generalizability and

effectiveness. The AFDD methods detected and isolated faults consistently, even with varying HVAC configurations and data inputs. For example, the dirty filter fault was undetectable using the semi-supervised but was successfully identified and diagnosed using the unsupervised method. The detected samples can then be utilized to enhance the supervised AFDD model.

The proposed methods in this thesis can be applied together to form a robust, integrated AFDD package, capable of handling a wide range of fault types and operational conditions. They not only improve energy efficiency and reduce emissions but also enhance occupant comfort by ensuring the HVAC systems operate optimally. Furthermore, the ability to work with raw, unlabeled data from BEMS without extensive preprocessing makes these methods practical for real-world applications, facilitating broader adoption in light commercial building control systems.

Lastly, ongoing testing with new data and firmware updates confirms the methods' adaptability and effectiveness in maintaining reliable AFDD, thereby ensuring continuous improvement in HVAC system performance and energy efficiency.

References

[1] F. Xiao, S. Wang, Progress and methodologies of lifecycle commissioning of HVAC systems to enhance building sustainability, Renewable and Sustainable Energy Reviews 13 (2009) 1144–1149. https://doi.org/10.1016/j.rser.2008.03.006.

[2] M.S. Mirnaghi, F. Haghighat, Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review, Energy Build 229 (2020). https://doi.org/10.1016/j.enbuild.2020.110492.

[3] M.S. Piscitelli, D.M. Mazzarelli, A. Capozzoli, Enhancing operational performance of AHUs through an advanced fault detection and diagnosis process based on temporal association and decision rules, Energy Build 226 (2020). https://doi.org/10.1016/j.enbuild.2020.110369.

[4] D. Dey, B. Dong, A probabilistic approach to diagnose faults of air handling units in buildings, Energy Build 130 (2016) 177–187. https://doi.org/10.1016/j.enbuild.2016.08.017.

[5] S. Frank, G. Lin, X. Jin, R. Singla, A. Farthing, J. Granderson, A performance evaluation framework for building fault detection and diagnosis algorithms, Energy Build 192 (2019) 84–92. https://doi.org/10.1016/j.enbuild.2019.03.024.

[6] A. Abid, M.T. Khan, J. Iqbal, A review on fault detection and diagnosis techniques: basics and beyond, Artif Intell Rev 54 (2021) 3639–3664. https://doi.org/10.1007/s10462-020-09934-2.

[7] B. Gunay, J. Bursill, B. Huchuk, S. Shillinglaw, Inverse model-based detection of programming logic faults in multiple zone VAV AHU systems, Build Environ 211 (2022). https://doi.org/10.1016/j.buildenv.2021.108732.

[8] H.B. Gunay, Z. Shi, G. Newsham, R. Moromisato, Detection of zone sensor and actuator faults through inverse greybox modelling, Build Environ 171 (2020). https://doi.org/10.1016/j.buildenv.2020.106659.

[9] Y. Yu, D. Woradechjumroen, D. Yu, A review of fault detection and diagnosis methodologies on air-handling units, Energy Build 82 (2014) 550–562. https://doi.org/10.1016/j.enbuild.2014.06.042.

[10] N. Torabi, H. Burak Gunay, W. O'Brien, R. Moromisato, Inverse model-based virtual sensors for detection of hard faults in air handling units, Energy Build 253 (2021). https://doi.org/10.1016/j.enbuild.2021.111493.

[11] A. Hosseini Gourabpasi, M. Nik-Bakht, Knowledge Discovery by Analyzing the State of the Art of Data-Driven Fault Detection and Diagnostics of Building HVAC, CivilEng 2 (2021) 986–1008. https://doi.org/10.3390/civileng2040053.

[12] W. Kim, S. Katipamula, A review of fault detection and diagnostics methods for building systems, Sci Technol Built Environ 24 (2018) 3–21. https://doi.org/10.1080/23744731.2017.1318008.

[13] A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, M. Rampazzo, F. Simmini, Data-driven Fault Detection and Diagnosis for HVAC water chillers, Control Eng Pract 53 (2016) 79–91. https://doi.org/10.1016/j.conengprac.2016.04.018.

[14] Z. Chen, Z. O'Neill, J. Wen, O. Pradhan, T. Yang, X. Lu, G. Lin, S. Miyata, S. Lee, C. Shen, R. Chiosa, M.S. Piscitelli, A. Capozzoli, F. Hengel, A. Kührer, M. Pritoni, W. Liu, J. Clauß, Y. Chen, T. Herr, A review of data-driven fault detection and diagnostics for building HVAC systems, Appl Energy 339 (2023). https://doi.org/10.1016/j.apenergy.2023.121030.

[15] S. Katipamula, M.R. Brambley, Review article: Methods for fault detection, diagnostics, and prognostics for building systems—A review, part I, HVAC and R Research 11 (2005) 3–25. https://doi.org/10.1080/10789669.2005.10391123.

[16] S.P. Melgaard, K.H. Andersen, A. Marszal-Pomianowska, R.L. Jensen, P.K. Heiselberg, Fault Detection and Diagnosis Encyclopedia for Building Systems: A Systematic Review, Energies (Basel) 15 (2022). https://doi.org/10.3390/en15124366.

[17] Ph.D.; S.L.Ph.D. Jin Wen, RP-1312 -- TOOLS FOR EVALUATING FAULT DETECTION AND DIAGNOSTIC METHODS FOR AIR-HANDLING UNITS, 2012.

[18] Jessica Granderson, Guanjing Lin, Yimin Chen, Armando Casillas, Sen Huang, Draguna Vrabie, LBNL Fault Detection and Diagnostics Data Sets: Single Duct Air Handling Unit, 2022.

[19] M. Ahern, D.T.J. O'sullivan, K. Bruton, Development of a Framework to Aid the Transition from Reactive to Proactive Maintenance Approaches to Enable Energy Reduction, Applied Sciences (Switzerland) 12 (2022). https://doi.org/10.3390/app12136704.

[20] J. Huang, J. Wen, H. Yoon, O. Pradhan, T. Wu, Z. O'Neill, K. Selcuk Candan, Real vs. simulated: Questions on the capability of simulated datasets on building fault detection for energy efficiency from a data-driven perspective, Energy Build 259 (2022). https://doi.org/10.1016/j.enbuild.2022.111872.

[21] Z. Du, B. Fan, X. Jin, J. Chi, Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis, Build Environ 73 (2014) 1–11. https://doi.org/10.1016/j.buildenv.2013.11.021.

[22] C. Miller, Z. Nagy, A. Schlueter, Automated daily pattern filtering of measured building performance data, Autom Constr 49 (2015) 1–17. https://doi.org/10.1016/j.autcon.2014.09.004.

[23] D. Li, G. Hu, C.J. Spanos, A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis, Energy Build 128 (2016) 519–529. https://doi.org/10.1016/j.enbuild.2016.07.014.

[24] G. Li, Y. Hu, Improved sensor fault detection, diagnosis and estimation for screw chillers using density-based clustering and principal component analysis, Energy Build 173 (2018) 502–515. https://doi.org/10.1016/j.enbuild.2018.05.025.

[25] C. Fan, F. Xiao, C. Yan, A framework for knowledge discovery in massive building automation data and its application in building diagnostics, Autom Constr 50 (2015) 81–90. https://doi.org/10.1016/j.autcon.2014.12.006.

[26] M. Dey, S.P. Rana, S. Dudley, A case study based approach for remote fault detection using multi-level machine learning in a smart building, Smart Cities 3 (2020) 401–419. https://doi.org/10.3390/smartcities3020021.

[27] M. Dey, S.P. Rana, S. Dudley, Smart building creation in large scale HVAC environments through automated fault detection and diagnosis, Future Generation Computer Systems 108 (2020) 950–966. https://doi.org/10.1016/j.future.2018.02.019.

[28] H.B. Gunay, Z. Shi, Cluster analysis-based anomaly detection in building automation systems, Energy Build 228 (2020). https://doi.org/10.1016/j.enbuild.2020.110445.

[29] J. Aguilar, D. Ardila, A. Avendaño, F. Macias, C. White, J. Gomez-Pulido, J.G. De Mesa, A. Garces-Jimenez, An autonomic cycle of data analysis tasks for the supervision of HVAC systems of smart building, Energies (Basel) 13 (2020). https://doi.org/10.3390/en13123103.

[30] Y. Xu, C. Yan, J. Shi, Z. Lu, X. Niu, Y. Jiang, F. Zhu, An anomaly detection and dynamic energy performance evaluation method for HVAC systems based on data mining, Sustainable Energy Technologies and Assessments 44 (2021). https://doi.org/10.1016/j.seta.2021.101092.

[31] A. Rosato, M.S. Piscitelli, A. Capozzoli, Data-Driven Fault Detection and Diagnosis: Research and Applications for HVAC Systems in Buildings, Energies (Basel) 16 (2023). https://doi.org/10.3390/en16020854.

[32] R. Yan, Z. Ma, Y. Zhao, G. Kokogiannakis, A decision tree based data-driven diagnostic strategy for air handling units, Energy Build 133 (2016) 37–45. https://doi.org/10.1016/j.enbuild.2016.09.039.

[33] G. Li, H. Chen, Y. Hu, J. Wang, Y. Guo, J. Liu, H. Li, R. Huang, H. Lv, J. Li, An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators, Appl Therm Eng 129 (2018) 1292–1303. https://doi.org/10.1016/j.applthermaleng.2017.10.013.

[34] A. Capozzoli, M.S. Piscitelli, S. Brandi, D. Grassi, G. Chicco, Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings, Energy 157 (2018) 336–352. https://doi.org/10.1016/j.energy.2018.05.127.

[35] M.S. Piscitelli, S. Brandi, A. Capozzoli, F. Xiao, A data analytics-based tool for the detection and diagnosis of anomalous daily energy patterns in buildings, Build Simul 14 (2021) 131–147. https://doi.org/10.1007/s12273-020-0650-1.

[36] X. Liu, Y. Ding, H. Tang, F. Xiao, A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data, Energy Build 231 (2021). https://doi.org/10.1016/j.enbuild.2020.110601.

[37] R. Chiosa, M.S. Piscitelli, A. Capozzoli, A data analytics-based energy information system (EIS) tool to perform meter-level anomaly detection and diagnosis in buildings, Energies (Basel) 14 (2021). https://doi.org/10.3390/en14010237.

[38] R. Chiosa, M.S. Piscitelli, C. Fan, A. Capozzoli, Towards a self-tuned data analytics-based process for an automatic context-aware detection and diagnosis of anomalies in building

energy consumption timeseries, Energy Build 270 (2022). https://doi.org/10.1016/j.enbuild.2022.112302.

[39] S. Wang, J. Qin, Sensor fault detection and validation of VAV terminals in air conditioning systems, Energy Convers Manag 46 (2005) 2482–2500. https://doi.org/10.1016/j.enconman.2004.11.011.

[40] Z. Du, X. Jin, Detection and diagnosis for sensor fault in HVAC systems, Energy Convers Manag 48 (2007) 693–702. https://doi.org/10.1016/j.enconman.2006.09.023.

[41] Z. Du, X. Jin, Multiple faults diagnosis for sensors in air handling unit using Fisher discriminant analysis, Energy Convers Manag 49 (2008) 3654–3665. https://doi.org/10.1016/j.enconman.2008.06.032.

[42] Y. Hu, H. Chen, J. Xie, X. Yang, C. Zhou, Chiller sensor fault detection using a self-Adaptive Principal Component Analysis method, Energy Build 54 (2012) 252–258. https://doi.org/10.1016/j.enbuild.2012.07.014.

[43] S. Li, J. Wen, A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform, Energy Build 68 (2014) 63–71. https://doi.org/10.1016/j.enbuild.2013.08.044.

[44] S. Li, J. Wen, Application of pattern matching method for detecting faults in air handling unit system, Autom Constr 43 (2014) 49–58. https://doi.org/10.1016/j.autcon.2014.03.002.

[45] N. Cotrufo, R. Zmeureanu, PCA-based method of soft fault detection and identification for the ongoing commissioning of chillers, Energy Build 130 (2016) 443–452. https://doi.org/10.1016/j.enbuild.2016.08.083.

[46] Y. Hu, H. Chen, G. Li, H. Li, R. Xu, J. Li, A statistical training data cleaning strategy for the PCA-based chiller sensor fault detection, diagnosis and data reconstruction method, Energy Build 112 (2016) 270–278. https://doi.org/10.1016/j.enbuild.2015.11.066.

[47] Y. Guo, G. Li, H. Chen, Y. Hu, H. Li, L. Xing, W. Hu, An enhanced PCA method with Savitzky-Golay method for VRF system sensor fault detection and diagnosis, Energy Build 142 (2017) 167–178. https://doi.org/10.1016/j.enbuild.2017.03.026.

[48] S. Shi, G. Li, H. Chen, Y. Hu, X. Wang, Y. Guo, S. Sun, An efficient VRF system fault diagnosis strategy for refrigerant charge amount based on PCA and dual neural network model, Appl Therm Eng 129 (2018) 1252–1262. https://doi.org/10.1016/j.applthermaleng.2017.09.117.

[49] L. Burgas, J. Colomer, J. Melendez, F.I. Gamero, S. Herraiz, Integrated unfold-pca monitoring application for smart buildings: An ahu application example, Energies (Basel) 14 (2021). https://doi.org/10.3390/en14010235.

[50] Z. Zhou, H. Chen, G. Li, H. Zhong, M. Zhang, J. Wu, Data-driven fault diagnosis for residential variable refrigerant flow system on imbalanced data environments, International Journal of Refrigeration 125 (2021) 34–43. https://doi.org/10.1016/j.ijrefrig.2021.01.009.

[51] M. Dey, S.P. Rana, S. Dudley, Smart building creation in large scale HVAC environments through automated fault detection and diagnosis, Future Generation Computer Systems 108 (2020) 950–966. https://doi.org/10.1016/j.future.2018.02.019.

[52] A. Alghanmi, A. Yunusa-Kaltungo, A whole-building data-driven fault detection and diagnosis approach for public buildings in hot climate regions, Energy and Built Environment (2023). https://doi.org/10.1016/j.enbenv.2023.07.005.

[53] A. Liang, Y. Hu, G. Li, The impact of improved PCA method based on anomaly detection on chiller sensor fault detection, International Journal of Refrigeration 155 (2023) 184–194. https://doi.org/10.1016/j.ijrefrig.2023.09.002.

[54] T. Zhao, B. Zhang, M. Li, G. Liu, P. Wang, Handling fault detection and diagnosis in incomplete sensor measurements for BAS based HVAC system, Journal of Building Engineering 80 (2023). https://doi.org/10.1016/j.jobe.2023.108098.

[55] X. Yang, J. Chen, X. Gu, R. He, J. Wang, Sensitivity analysis of scalable data on three PCA related fault detection methods considering data window and thermal load matching strategies, Expert Syst Appl 234 (2023). https://doi.org/10.1016/j.eswa.2023.121024.

[56] C. Fan, Q. Wu, Y. Zhao, L. Mo, Integrating active learning and semi-supervised learning for improved data-driven HVAC fault diagnosis performance, Appl Energy 356 (2024). https://doi.org/10.1016/j.apenergy.2023.122356.

[57] N. Torabi, H. Burak Gunay, W. O'Brien, R. Moromisato, Inverse model-based virtual sensors for detection of hard faults in air handling units, Energy Build 253 (2021). https://doi.org/10.1016/j.enbuild.2021.111493.

[58] D. Darwazeh, B. Gunay, J. Duquette, Development of Inverse Greybox Model-Based Virtual Meters for Air Handling Units, IEEE Transactions on Automation Science and Engineering 18 (2021) 323–336. https://doi.org/10.1109/TASE.2020.3005888.

[59] B. Gunay, J. Bursill, B. Huchuk, S. Shillinglaw, Inverse model-based detection of programming logic faults in multiple zone VAV AHU systems, Build Environ 211 (2022). https://doi.org/10.1016/j.buildenv.2021.108732.

[60] S. Li, J. Wen, A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform, Energy Build 68 (2014) 63–71. https://doi.org/10.1016/j.enbuild.2013.08.044.

[61] J.Edward. Jackson, A user's guide to principal components, Wiley, 1991.

[62] Y. Hu, H. Chen, J. Xie, X. Yang, C. Zhou, Chiller sensor fault detection using a self-Adaptive Principal Component Analysis method, Energy Build 54 (2012) 252–258. https://doi.org/10.1016/j.enbuild.2012.07.014.

[63] S. Wang, F. Xiao, Detection and diagnosis of AHU sensor faults using principal component analysis method, Energy Convers Manag 45 (2004) 2667–2686. https://doi.org/10.1016/j.enconman.2003.12.008.

[64] S. Wang, F. Xiao, AHU sensor fault diagnosis using principal component analysis method, Energy Build 36 (2004) 147–160. https://doi.org/10.1016/j.enbuild.2003.10.002.

[65] S. Wang, J. Qin, Sensor fault detection and validation of VAV terminals in air conditioning systems, Energy Convers Manag 46 (2005) 2482–2500. https://doi.org/10.1016/j.enconman.2004.11.011.

[66] S. Wang, J. Cui, A robust fault detection and diagnosis strategy for centrifugal chillers, HVAC and R Research 12 (2006) 407–428. https://doi.org/10.1080/10789669.2006.10391187.

[67] S. Wang, F. Xiao, Sensor fault detection and diagnosis of air-handling units using a condition-based adaptive statistical method, HVAC and R Research 12 (2006) 127–150. https://doi.org/10.1080/10789669.2006.10391171.

[68] F. Xiao, S. Wang, J. Zhang, A diagnostic tool for online sensor health monitoring in air-conditioning systems, Autom Constr 15 (2006) 489–503. https://doi.org/10.1016/j.autcon.2005.06.001.

[69] Z. Du, X. Jin, Detection and diagnosis for sensor fault in HVAC systems, Energy Convers Manag 48 (2007) 693–702. https://doi.org/10.1016/j.enconman.2006.09.023.

[70] Z. Du, X. Jin, Multiple faults diagnosis for sensors in air handling unit using Fisher discriminant analysis, Energy Convers Manag 49 (2008) 3654–3665. https://doi.org/10.1016/j.enconman.2008.06.032.

[71] F. Xiao, S. Wang, X. Xu, G. Ge, An isolation enhanced PCA method with expert-based multivariate decoupling for sensor FDD in air-conditioning systems, Appl Therm Eng 29 (2009) 712–722. https://doi.org/10.1016/j.applthermaleng.2008.03.046.

[72] S. Wang, Q. Zhou, F. Xiao, A system-level fault detection and diagnosis strategy for HVAC systems involving sensor faults, Energy Build 42 (2010) 477–490. https://doi.org/10.1016/j.enbuild.2009.10.017.

[73] S. Li, J. Wen, Application of pattern matching method for detecting faults in air handling unit system, Autom Constr 43 (2014) 49–58. https://doi.org/10.1016/j.autcon.2014.03.002.

[74] M. Padilla, D. Choinière, A combined passive-active sensor fault detection and isolation approach for air handling units, Energy Build 99 (2015) 214–219. https://doi.org/10.1016/j.enbuild.2015.04.035.

[75] R. Yan, Z. Ma, G. Kokogiannakis, Y. Zhao, A sensor fault detection strategy for air handling units using cluster analysis, Autom Constr 70 (2016) 77–88. https://doi.org/10.1016/j.autcon.2016.06.005.

[76] Y. Hu, H. Chen, G. Li, H. Li, R. Xu, J. Li, A statistical training data cleaning strategy for the PCA-based chiller sensor fault detection, diagnosis and data reconstruction method, Energy Build 112 (2016) 270–278. https://doi.org/10.1016/j.enbuild.2015.11.066.

[77] Y. Guo, G. Li, H. Chen, Y. Hu, H. Li, L. Xing, W. Hu, An enhanced PCA method with Savitzky-Golay method for VRF system sensor fault detection and diagnosis, Energy Build 142 (2017) 167–178. https://doi.org/10.1016/j.enbuild.2017.03.026.

[78] Y. Guo, G. Li, H. Chen, Y. Hu, H. Li, J. Liu, M. Hu, W. Hu, Modularized PCA method combined with expert-based multivariate decoupling for FDD in VRF systems including indoor unit faults, Appl Therm Eng 115 (2017) 744–755. https://doi.org/10.1016/j.applthermaleng.2017.01.008.

[79] G. Li, Y. Hu, Improved sensor fault detection, diagnosis and estimation for screw chillers using density-based clustering and principal component analysis, Energy Build 173 (2018) 502–515. https://doi.org/10.1016/j.enbuild.2018.05.025.

[80] A. Montazeri, S.M. Kargar, Fault detection and diagnosis in air handling using data-driven methods, Journal of Building Engineering 31 (2020). https://doi.org/10.1016/j.jobe.2020.101388.

[81] Y. Guo, H. Chen, Fault diagnosis of VRF air-conditioning system based on improved Gaussian mixture model with PCA approach, International Journal of Refrigeration 118 (2020) 1–11. https://doi.org/10.1016/j.ijrefrig.2020.06.009.

[82] L. Burgas, J. Colomer, J. Melendez, F.I. Gamero, S. Herraiz, Integrated unfold-pca monitoring application for smart buildings: An ahu application example, Energies (Basel) 14 (2021). https://doi.org/10.3390/en14010235.

[83] X. Yang, R. He, J. Wang, X. Li, R. Liu, Using thermal load matching strategy to locate historical benchmark data for moving-window PCA based fault detection in air handling units, Sustainable Energy Technologies and Assessments 52 (2022). https://doi.org/10.1016/j.seta.2022.102238.

[84] A. Liang, Y. Hu, G. Li, The impact of improved PCA method based on anomaly detection on chiller sensor fault detection, International Journal of Refrigeration 155 (2023) 184–194. https://doi.org/10.1016/j.ijrefrig.2023.09.002.

[85] S. Wen, W. Zhang, Y. Sun, Z. Li, B. Huang, S. Bian, L. Zhao, Y. Wang, An enhanced principal component analysis method with Savitzky–Golay filter and clustering algorithm for sensor fault detection and diagnosis, Appl Energy 337 (2023). https://doi.org/10.1016/j.apenergy.2023.120862.

[86] X. Yang, J. Chen, X. Gu, R. He, J. Wang, Sensitivity analysis of scalable data on three PCA related fault detection methods considering data window and thermal load matching strategies, Expert Syst Appl 234 (2023). https://doi.org/10.1016/j.eswa.2023.121024.

[87] Q. Ma, C. Yue, M. Yu, Y. Song, P. Cui, Y. Yu, Research on fault diagnosis strategy of air-conditioning system based on signal demodulation and BPNN-PCA, International Journal of Refrigeration 158 (2024) 124–134. https://doi.org/10.1016/j.ijrefrig.2023.12.008.

[88] G. Li, C. Xiong, J. Gao, H. Zhu, C. Wang, J. Xiao, Fault detection, diagnosis and calibration of heating, ventilation and air conditioning sensors by combining principal component analysis and improved bayesian inference, Journal of Building Engineering 82 (2024) 108230. https://doi.org/10.1016/j.jobe.2023.108230.

[89] Mathew C. Comstock, James E. Braun, RP-1043 -- Fault Detection And Diagnostic (FDD) Requirements And Evaluation Tools For Chillers, 2006.

[90] K. Bruton, D. Coakley, P. Raftery, D.O. Cusack, M.M. Keane, D.T.J. O'Sullivan, Comparative analysis of the AHU InFO fault detection and diagnostic expert tool for AHUs with APAR, Energy Effic 8 (2015) 299–322. https://doi.org/10.1007/s12053-014-9289-z.

[91] K. Bruton, P. Raftery, B. Kennedy, M.M. Keane, D.T.J. O'Sullivan, Review of automated fault detection and diagnostic tools in air handling units, Energy Effic 7 (2014) 335–351. https://doi.org/10.1007/s12053-013-9238-2.

[92] Z. Shi, W. O'Brien, Development and implementation of automated fault detection and diagnostics for building systems: A review, Autom Constr 104 (2019) 215–229. https://doi.org/10.1016/j.autcon.2019.04.002.

[93] Y. Zhao, T. Li, X. Zhang, C. Zhang, Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future, Renewable and Sustainable Energy Reviews 109 (2019) 85–101. https://doi.org/10.1016/j.rser.2019.04.021.

[94] G. Lin, H. Kramer, J. Granderson, Building fault detection and diagnostics: Achieved savings, and methods to evaluate algorithm performance, Build Environ 168 (2020). https://doi.org/10.1016/j.buildenv.2019.106505.

[95] M. Ahern, D.T.J. O'Sullivan, K. Bruton, Implementation of the IDAIC framework on an air handling unit to transition to proactive maintenance, Energy Build 284 (2023) 112872. https://doi.org/10.1016/j.enbuild.2023.112872.

[96] C. Fan, F. Xiao, C. Yan, C. Liu, Z. Li, J. Wang, A novel methodology to explain and evaluate data-driven building energy performance models based on interpretable machine learning, Appl Energy 235 (2019) 1551–1560. https://doi.org/10.1016/j.apenergy.2018.11.081.

[97] J. Chen, L. Zhang, Y. Li, Y. Shi, X. Gao, Y. Hu, A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems, Renewable and Sustainable Energy Reviews 161 (2022). https://doi.org/10.1016/j.rser.2022.112395.

[98] M. Babadi Soultanzadeh, M.M. Ouf, M. Nik-Bakht, P. Paquette, S. Lupien, Fault detection and diagnosis in light commercial buildings' HVAC systems: A comprehensive framework, application, and performance evaluation, Energy Build 316 (2024) 114341. https://doi.org/10.1016/j.enbuild.2024.114341.

[99] M. Babadi Soutanzadeh, M. Ouf, N.-B. Nik-Bakht, P. Paquette, S. Lupien, A Framework for Automated Fault Detection in Light Commercial Buildings HVAC System, in: ASHRAE Transactions 130, 2024: pp. 590–599. https://www.scopus.com/record/display.uri?eid=2-s2.0-85198914362&origin=resultslist (accessed July 28, 2024).

[100] J.Edward. Jackson, A user's guide to principal components, Wiley, 1991.

[101] M. Ester, H.-P. Krigel, Sander, Xiaowei. Jorg. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, in: (KDD-96, 1996: pp. 226–231.

[102] A. Capozzoli, F. Lauro, I. Khan, Fault detection analysis using data mining techniques for a cluster of smart office buildings, Expert Syst Appl 42 (2015) 4324–4338. https://doi.org/10.1016/j.eswa.2015.01.010.

[103] G. Li, Y. Hu, H. Chen, J. Wang, Y. Guo, J. Liu, J. Li, Identification and isolation of outdoor fouling faults using only built-in sensors in variable refrigerant flow system: A data mining approach, Energy Build 146 (2017) 257–270. https://doi.org/10.1016/j.enbuild.2017.04.041.

[104] W.Y. Lee, J.M. House, N.H. Kyong, Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks, Appl Energy 77 (2004) 153–170. https://doi.org/10.1016/S0306-2619(03)00107-7.

[105] Canadian Historical Climate Data, (n.d.). https://climate.weather.gc.ca/ (accessed August 31, 2023).

[106] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, 1987.

[107] K. Yan, L. Ma, Y. Dai, W. Shen, Z. Ji, D. Xie, Cost-sensitive and sequential feature selection for chiller fault detection and diagnosis, International Journal of Refrigeration 86 (2018) 401–409. https://doi.org/10.1016/j.ijrefrig.2017.11.003.

[108] M.A.F. Abdollah, R. Scoccia, M. Aprile, Transformer encoder based self-supervised learning for HVAC fault detection with unlabeled data, Build Environ 258 (2024). https://doi.org/10.1016/j.buildenv.2024.111568.

[109] John Bollinger, www.bollingerbands.com, (n.d.).

[110] P. Barandier, M. Mendes, A.J. Marques Cardoso, Comparative analysis of four classification algorithms for fault detection of heat pumps, Energy Build 316 (2024) 114342. https://doi.org/10.1016/j.enbuild.2024.114342.

[111] M.M. Hydeman, G. Cmar, J.J. Ouellette, S. Katipamula, V.-C. Brent, R. Eubanks, G. Paliaga, J.J. Coogan, M. Galler, M.A. Pouchak, C.R. Amundson, N. Hampton, P. Raftery, J.G. Boldt, R.N. Hume, J.J. Santos, R. Bristol, E. Koeppel, S.T. Taylor, A. Bruno, K. Li, L. Tully, C.A. Callaway, K. Ng, X. Zhou, ASHRAE Standing Guideline Project Committee 36 Cognizant TC: 1.4, Control Theory and Application SPLS Liaison: Julie M. Ferguson, 2018. www.ashrae.org/technology.

[112] M. Ahern, D.T.J. O'sullivan, K. Bruton, Specifications Table Value of the Data, 109208 Energy Build 48 (2023) 2023. https://doi.org/10.17632/8.