12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

A Hybrid-based Clustering Approach for Fault Detection in HVAC Systems

Hesam Hassanpour * **Amir H. Hamedi** * **Prashant Mhaskar** *
John M. House ** **Timothy I. Salsbury** **

* *Department of Chemical Engineering, McMaster University, Hamilton, Ontario, Canada*
** *These results were obtained during John House and Timothy Salsbury's affiliation with the Building Management Systems Research Group, Johnson Controls International plc.*

Abstract: This paper presents a hybrid model-based fault detection strategy for heating, ventilation, and air conditioning (HVAC) systems, focusing on air handling units (AHUs). Addressing the substantial energy inefficiencies in commercial buildings due to undetected HVAC faults, this research combines first-principles knowledge with data-driven techniques to enhance fault detection accuracy. First-principles based residuals (differences between expected and observed behaviors) are integrated with data (temperature measurements in different locations of AHU) to perform principal component analysis (PCA) (pre-processing step). Pre-processed data (principal component scores) are then utilized to perform clustering analysis using K-means and DBSCAN approaches. The proposed approach is tested against two common faults in AHUs and its performance is evaluated compared to a purely data-driven method. The results indicate that the hybrid method, which synergizes residual knowledge from first-principles models with data, significantly outperforms the purely data-driven approach. This is demonstrated through performance analysis using metrics like the adjusted rand index (ARI) and normalized mutual information (NMI). The research underscores the potential of the hybrid method in improving fault diagnosis of HVAC systems, helping to conserve energy by ensuring efficient and reliable operation.

Keywords: Fault Detection, HVAC Systems, Principal Component Analysis (PCA), K-means, DBSCAN

1. INTRODUCTION

The building industry plays a substantial role in energy usage, accounting for approximately 40% of worldwide energy consumption, with an anticipated annual growth rate of 1.5%. (Mirnaghi and Haghighat (2020)). Heating, ventilation, and air conditioning (HVAC) systems, in commercial buildings, significantly contribute to this consumption, where overlooked operational faults lead to substantial energy inefficiencies and compromise occupant comfort (Yang et al. (2014)). Approximately 30% of commercial building energy is wasted due to latent system failure, making HVAC fault management critical to energy conservation (Tun et al. (2021)). Increasing energy demands and inefficiencies in HVAC systems have intensified the development of various fault detection and diagnosis (FDD) strategies to improve efficiency and comfort. These methods are categorized into history, qualitative, and quantitative-based approaches (Katipamula and Brambley (2005)).

Fault diagnosis in HVAC systems is founded on integrating domain knowledge—qualitative or quantitative, and a strategic search for faults through normal operational templates or symptom-focused methods (Venkatasubramanian et al. (2003a)). Qualitative methods rely on if-then-else logic in data-scarce situations, but lacking in fundamental system physics, they can become unwieldy with complex behaviors (Isermann (2005)). Conversely, quantitative methods use mathematical models based on physical laws, employing analytical redundancy to detect faults through deviations between model predictions and actual system behavior(Venkatasubramanian et al. (2003c)).

The shift towards data-driven methodologies in FDD for HVAC systems has been driven by the complexity and non-linearity of such systems, making the development of accurate mathematical/physical-based models challenging and time-consuming (Venkatasubramanian et al. (2003b)). Data-driven and machine-learning approaches offer a promising alternative, particularly principal component analysis (PCA) and clustering techniques such as K-means and density-based spatial clustering of applications with noise (DBSCAN) for fault detection. They leverage system data to construct precise system representations, circumventing the limitations of traditional model-based methods (Hassanpour et al. (2020); Li and Wen (2014)). PCA excels in isolating key features by identifying orthogonal components, making it instrumental in detecting faults. At the same time, K-means and DBSCAN provide robust frameworks for the unsupervised diagnosis of operational anomalies in HVAC systems, reflecting a significant evolution in the field of FDD research (Wu et al. (2010); Mu et al. (2020)).

Hybrid approaches, integrating first-principles knowledge with data, have developed as a powerful strategy for FDD in HVAC systems (Gálvez et al. (2021)). Building upon our previous work (Hassanpour et al. (2020)), which introduced a hybrid FDD system for HVAC systems, this paper advances the methodology by incorporating unsupervised clustering techniques for fault detection. This approach utilizes the residual knowledge from first-principles models with data-driven analytics, specifically focusing on the application of clustering for unsupervised fault detection and diagnosis, addressing the limitations of prior implementations.

2. PRELIMINARIES

2.1 Air Handling Unit: First-Principles based Residuals

Air Handling Units (AHUs) are critical in HVAC systems for regulating temperature and ventilation in large buildings. They come in two types: constant air volume (CAV) and variable air volume (VAV) air-conditioning systems.

Utilizing the methodology by Seem and House (2009), this paper employs model-based residuals for fault detection in AHUs. These residuals are crucial indicators in our hybrid FDD approach, reflecting the system across four operational states:

- **State 1 (Heating)**: Heating coil valve controlled to maintain supply air temperature and dampers positioned for minimum outdoor air.
- **State 2 (Cooling with Outdoor Air)**: Dampers modulate to maintain supply air temperature at setpoint. The heating and cooling valves are closed.
- **State 3 (Mechanical Cooling with 100% Outdoor Air)**: Cooling valve controlled to maintain supply air temperature with dampers positioned for 100% outdoor air.
- **State 4 (Mechanical Cooling with Minimum Outdoor Air)**: Cooling valve controlled to maintain supply air temperature with dampers positioned for minimum outdoor air (ventilation requirement).

Each state is associated with specific sensor readings: supply air temperature (T_s), return air temperature (T_r), outdoor air temperature (T_o), and mixed air temperature (T_m). First-principles modeling is performed to develop/calculate the residuals based on each state of the AHU. These residuals provide a comparison between measured and expected conditions and highlight any discrepancies indicating faults.

In our FDD approach, we examine two common faults with their respective residuals:

- **Fault 1**: 2C offset in the return air temperature sensor.
- **Fault 2**: 2C offset in the mixed air temperature sensor.

The residuals for State 1 and State 2 are given as:

$$r_5 = f_{\text{design}} - \frac{T_{m,1} - T_{r,1}}{T_{o,1} - T_{r,1}} \tag{1}$$

$$r_6 = T_{S,2} - T_{m,2} - \frac{\hat{\dot{W}}_{\text{fan}}}{\hat{\dot{m}}_S \hat{c}_p} \tag{2}$$

In these equations:

- f_{design} is the estimated fraction of outdoor air that should typically mix with the return air to form the mixed air under design conditions. The value of f_{design} is determined to be 0.3 based on the numerical experiments.
- $T_{m,1}$ and $T_{m,2}$ represent the mixed air temperatures in States 1 and 2, respectively.
- $T_{r,1}$ is the return air temperature in State 1.
- $T_{o,1}$ is the outdoor air temperature in State 1.
- $T_{S,2}$ is the supply air temperature in State 2.
- $\hat{\dot{W}}_{fan}$ denotes the design power of the supply fan, which is required to move the conditioned air throughout the building and is given as 7.14 kW.
- $\hat{\dot{m}}_S$ is the mass flow rate of supply air, crucial for determining the amount of air delivered to the conditioned space, with a design value of 10.53 kg/s.
- $\hat{c}_p$ is the specific heat of the moist air mixture at constant pressure, essential for calculating the energy needed to condition the air, with a design value of 1.02 kJ/(kgC).

These parameters are used in the residuals to assess whether the AHU is operating within the expected parameters or if an anomaly could indicate a fault. The correct operation should result in residuals close to zero, whereas significant deviations may point to specific faults in the system.

2.2 Data-driven methods

PCA Principal Component Analysis (PCA) is utilized as a method for dimensionality reduction, particularly addressing multicollinearity among process variables. The essence of PCA lies in representing high-dimensional data in a lower-dimensional space, especially when the data is proximal to a linear manifold within the high-dimensional space. By identifying and projecting data onto this linear manifold, PCA preserves the essential characteristics of the data with minimal variability in orthogonal directions. The principal components (PCs) are sorted such that the first component captures the maximum variation in the original real variable space, followed by each subsequent component accounting for the maximum remaining variation. The fundamental equation in PCA is as follows:

$$X = TP^T + E \tag{3}$$

where X is the original data matrix, T is the matrix of scores, P is the matrix of loadings, and E is the residual matrix. This equation decomposes the data into a product of scores and loadings, with residuals capturing the unexplained variation in X.

K-means Clustering K-means clustering, a prominent unsupervised learning algorithm, was originally proposed by MacQueen et al. (1967). This method is noted for its simplicity and efficacy. The primary objective of K-means clustering is to partition a dataset X of size $N \times M$

into K disjoint subsets $C_1, \ldots, C_K$. The optimization aim is to minimize the clustering criterion F, often represented by the sum of squared Euclidean distances between each data sample x_i and the corresponding cluster center m_k of each subset C_K.

A crucial element in K-means clustering is the selection of the ideal number of clusters. The elbow method is frequently utilized for this purpose. This method entails graphing the explained variance as a function of cluster quantity and selecting the curve's elbow point as the cluster count. This method is heuristic but widely accepted for its simplicity and effectiveness in identifying the point where the increase in the number of clusters does not significantly improve the fitting.

DBSCAN Clustering DBSCAN algorithm, introduced by Ester et al. (1996), is a density-based spatial clustering method. In the context of DBSCAN clustering, there are several key definitions. The ϵ neighborhood of an object p in a dataset D is defined as the region with p as its center and ϵ as its radius, encompassing all objects within this distance:

$$N_\epsilon(p) = \{q \in D | Dist(p, q) \leq \epsilon\} \tag{4}$$

where D is the dataset, $Dist(p, q)$ is the distance between object p and q. In simpler terms, it encompasses all objects in dataset D that are within a distance of ϵ from object p. A core object in D is one whose ϵ neighborhood contains more objects than a threshold, MinPts:

$$|N_\epsilon(p)| \geq MinPts \tag{5}$$

DBSCAN's procedure involves conducting region queries on each object to ascertain its ϵ neighborhood. Objects with neighbors less than MinPts are marked as noise. Otherwise, they form a cluster C_1, and the process continues for each neighbor. This repeats, creating new clusters C_i until all objects are classified.

The essence of DBSCAN lies in identifying high-density regions in the dataset, marked by small distances between samples, and distinguishing these from lower-density areas. The algorithm's efficiency and effectiveness rely on the proper selection of parameters ϵ and MinPts, crucial for defining high-density thresholds.

Numerical Metrics In the assessment of clustering outcomes, two key numerical metrics are employed: the Adjusted Rand Index (ARI) (Hubert and Arabie (1985)) and Normalized Mutual Information (NMI)(Strehl and Ghosh (2002)):

$$ARI = \frac{\sum_{ij} \binom{n_{ij}}{2} - \left[\sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2}\right] / \binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{n_i}{2} + \sum_j \binom{n_j}{2}\right] - \left[\sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2}\right] / \binom{n}{2}} \tag{6}$$

where n_{ij} is the number of elements common between clusters i and j, n_i and n_j are the number of elements in clusters i and j respectively, and n is the total number of elements. ARI values range from -1 to 1, with 1 indicating a perfect match, 0 suggesting random labeling, and negative values implying dissimilarity worse than random chance.

NMI, a normalization of the Mutual Information (MI) score, ranges between 0 and 1, indicating the level of correlation between two clusterings. It is defined as follows:

$$NMI(U, V) = \frac{2I(U; V)}{H(U) + H(V)} \tag{7}$$

where U and V are two clusters, $I(U; V)$ represents mutual information, and $H(U)$ and $H(V)$ are the entropies of the clusterings. An NMI of 1 implies perfect correlation, while a value of 0 indicates no mutual information.

3. PROPOSED METHOD

The developed hybrid FDD scheme leverages model-based residuals in conjunction with data-driven analytics to enhance fault detection in HVAC systems, marking an improvement in diagnostic processes.

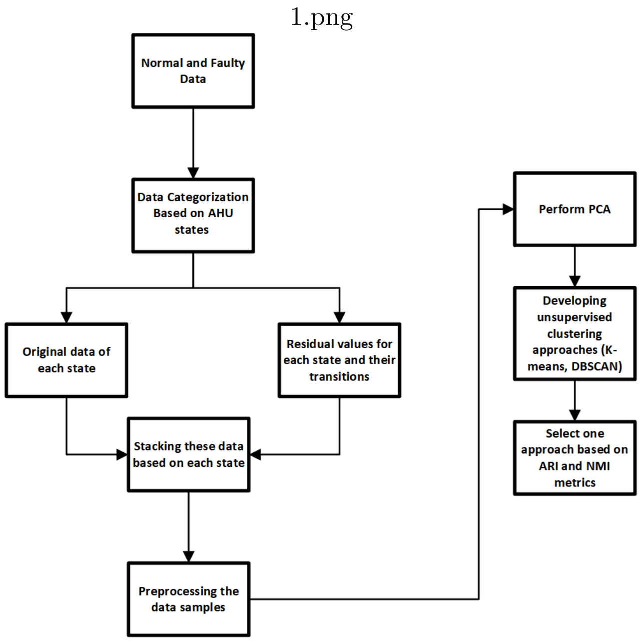

Fig. 1. Flow Diagram of the proposed hybrid approach

Referring to Fig. 1, the process begins by collecting data (normal and faulty data). This data is then categorized into different states based on the operating state of AHU. Following the methodology established by Seem and House (2009), residuals for each state and transitional phase are computed. For the scope of our study, which focuses on specific faults, only states S_1 and S_2 are used, hence only residuals r_5 and r_6 are needed.

These residuals are then stacked with the original data for each state. Standard scaling is applied to normalize the dataset. PCA is performed on the scaled data to determine the principal components. Unsupervised clustering algorithms, K-means, and DBSCAN are applied to the PCA scores. The selection of the clustering method is based on ARI and NMI metrics, ultimately achieving unsupervised fault detection in the system.

4. RESULTS AND DISCUSSION

As mentioned, two common faults (Fault 1 and Fault 2) are used, in our analysis, to evaluate the performance of different techniques. In addition, it should be noted that the occurrence of these faults in States 1 and 2 is considered. As shown in the flow diagram (Figure 1), the dataset is categorized based on the AHU states. Each state is further divided into three scenarios: the occurrence of Fault 1, Fault 2, or the presence of both faults in the dataset.

Initially, the efficacy of data-driven models in detecting Fault 1 in State 1 is assessed. Subsequently, the performance of our proposed hybrid method is evaluated, demonstrating its capability to facilitate unsupervised fault detection within the AHU of the HVAC system.

For the implementation of unsupervised data-driven clustering fault detection, the datasets are mean-centered and scaled to unit variance. PCA is then employed, on the standardized data, to derive component scores. The required number of principal components is determined by the cumulative explained variance, three for the purely data-driven method and four for the hybrid approach. The PCA scores are utilized to perform clustering using K-means and DBSCAN algorithms, which enables the distinction between normal and faulty operational clusters by comparison with the space of true labels.

Despite the slight differentiation provided by the K-means algorithm, its predetermined cluster count limits its diagnostic precision, particularly in areas of score space overlap. The DBSCAN algorithm, without a predefined cluster number, categorizes nine clusters, including noise (C_i, $i \in \{1, \ldots, 8\}$). This method of identification of noise underscores the challenge of distinguishing between normal and faulty data; as such, a portion of data points may be classified as noise. However, based on ARI and NMI metrics presented in Table 1, DBSCAN exhibits better performance, because it is influenced by the exclusion of noise in the cluster-to-true label comparison.

In the hybrid approach, residuals are computed for each state of the AHU and integrated into the dataset. Having applied PCA on these datasets, the resultant component scores are utilized to perform clustering. Figure 3 illustrates the effectiveness of this method. The top row, color-coded by the true operational labels, demonstrates a clear separation of normal and faulty conditions. This separation is leveraged by clustering algorithms to accurately cluster data points. The middle row, in this figure, presents the K-means clustering results, where two clusters are predefined (the number of clusters are determined using the elbow method). The bottom figure depicts the DBSCAN clustering, which includes Noise and clusters C_1 and C_2. Compared to the pure data-driven approach, the hybrid method significantly reduces noise points and accurately captures the two primary operational datasets. The efficacy of this method is further substantiated by the ARI and NMI metrics presented in Table 1, illustrating the superior performance of this clustering method compared with the data-driven approach.

The performance of data-driven and hybrid clustering methods, for all scenarios, is compared using the ARI and NMI metrics, and the results are listed in Table 1. Notably, in state 1, the hybrid method demonstrates a significant improvement over the data-driven approach, particularly with the DBSCAN clustering algorithm, where it achieves remarkably high ARI and NMI values. This represents a significant improvement in fault detection accuracy and reliability.

2.png

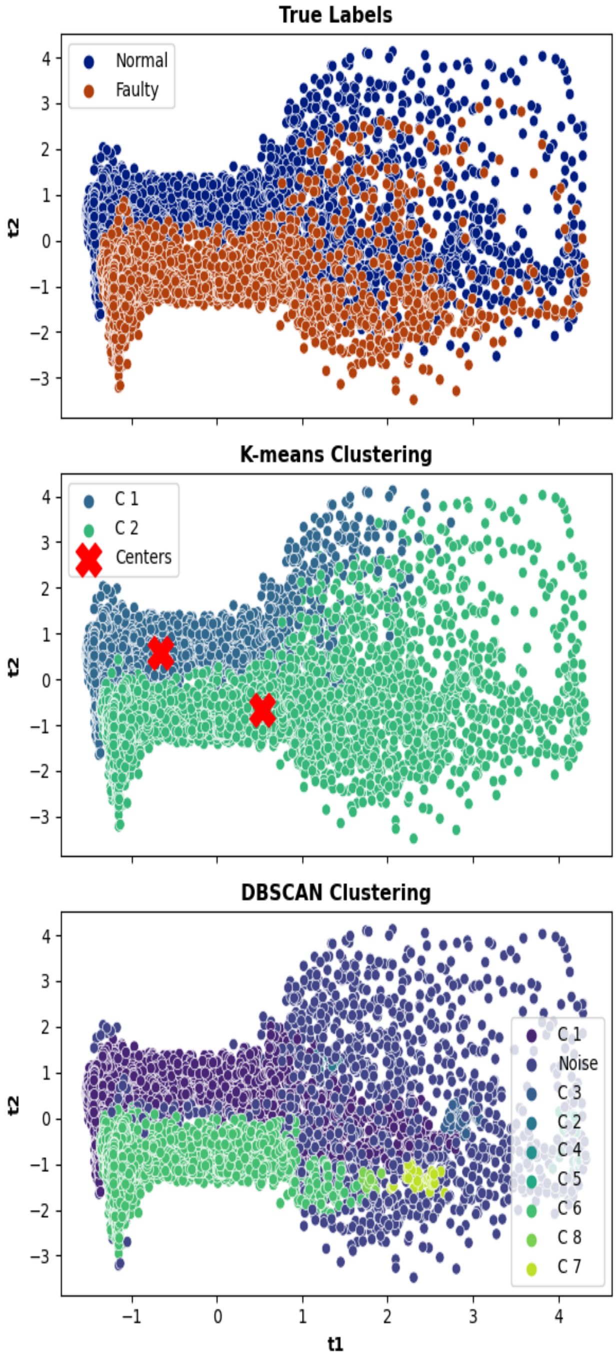

Fig. 2. Data-driven based clustering approach for Fault 1 in State 1: (Top) Data samples are color-coded based on the true labels, (Middle) K-means clustering performance with two predefined clusters, and (Bottom) DBSCAN clustering performance.

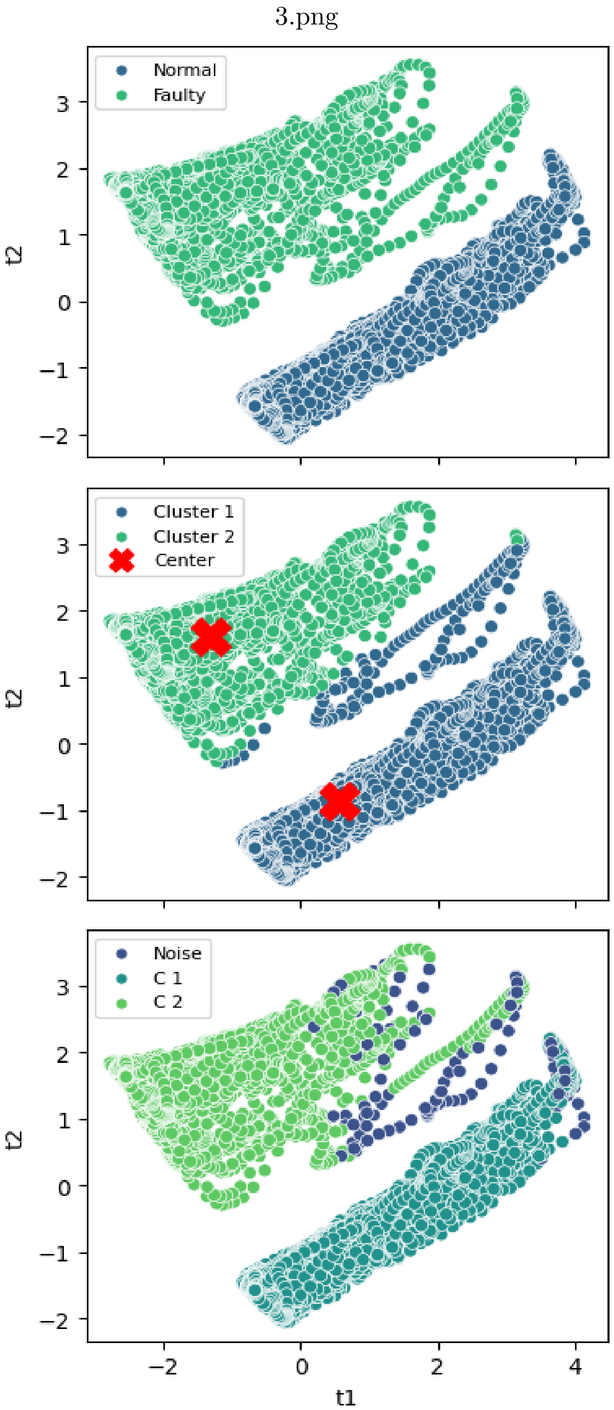

Fig. 3. Hybrid-based clustering approach for Fault 1 in State 1: (Top) Data samples are color-coded based on the true labels, (Middle) K-means clustering performance with two predefined clusters, and (Bottom) DBSCAN clustering performance.

In contrast, for state 2, both methods show comparable performance, suggesting specific scenarios where the data-driven approach remains equally viable. However, using a hybrid approach allows better differentiation of different (normal and faulty) datasets using both clustering techniques. In addition, across other scenarios, the hybrid method enhances detection accuracy, reinforcing its robustness and adaptability in varied conditions. The overall

Table 1. Performance Metrics of Fault Detection Methods

State	**Fault**	Data-driven			
		K-means		DBSCAN	
		ARI	NMI	ARI	NMI
1	1	0.5231	0.4768	0.6548	0.5994
	2	0.0209	0.0049	0.0160	0.0120
	1 and 2	0.3744	0.4451	0.4028	0.4832
2	1	0.8513	0.8966	0.9974	0.9931
	2	0.3391	0.5522	0.9895	0.9763
	1 and 2	0.5839	0.7258	0.9971	0.9976
State	**Fault**	Hybrid			
		K-means		DBSCAN	
		ARI	NMI	ARI	NMI
1	1	0.9313	0.8852	0.9765	0.9569
	2	0.9998	0.9999	0.9756	0.9434
	1 and 2	0.9716	0.9447	0.9716	0.9447
2	1	0.9857	0.9725	0.9992	0.9977
	2	0.8274	0.8161	0.9991	0.9971
	1 and 2	0.9993	0.9981	0.9992	0.9980

analysis indicates that DBSCAN, whether employed in data-driven or hybrid configurations, works better than the K-means algorithm, underscoring its superiority in this application domain. The hybrid method, with its integration of multiple techniques, emerges as not only efficient but also reliable, adapting to various fault conditions with improved precision. This adaptability is particularly valuable in HVAC systems, where accurate fault detection is crucial for maintaining system performance and efficiency.

5. CONCLUSION

The research demonstrates that the hybrid-based clustering method, integrating first-principles knowledge with data, significantly outperforms the purely data-driven clustering approach in the context of HVAC systems. This method exhibits better diagnostic accuracy and reliability, particularly in complex operational scenarios involving various fault conditions. The integration of model-based residuals into the dataset and the use of clustering techniques like DBSCAN and K-means enable the precise identification and classification of both normal and faulty operational states. The findings of this study highlight the importance of combining multiple diagnostic methodologies for enhanced fault detection in HVAC systems, offering valuable insights for future developments in energy-efficient building management and maintenance practices.

REFERENCES

Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, 226–231.

Gálvez, A., Seneviratne, D., and Galar, D. (2021). Hybrid model development for hvac system in transportation. *Technologies*, 9(1), 18.

Hassanpour, H., Mhaskar, P., House, J.M., and Salsbury, T.I. (2020). A hybrid modeling approach integrating first-principles knowledge with statistical methods for fault detection in hvac systems. *Computers & Chemical Engineering*, 142, 107022.

Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of classification*, 2, 193–218.

Isermann, R. (2005). Model-based fault-detection and diagnosis–status and applications. *Annual Reviews in control*, 29(1), 71–85.

Katipamula, S. and Brambley, M.R. (2005). Methods for fault detection, diagnostics, and prognostics for building systems—a review, part i. *Hvac&R Research*, 11(1), 3–25.

Li, S. and Wen, J. (2014). A model-based fault detection and diagnostic methodology based on pca method and wavelet transform. *Energy and Buildings*, 68, 63–71.

MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, 281–297. Oakland, CA, USA.

Mirnaghi, M.S. and Haghighat, F. (2020). Fault detection and diagnosis of large-scale hvac systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings*, 229, 110492.

Mu, Z., Wu, Y., Yin, H., Liu, Z., and Liu, C. (2020). Study on single-phase ground fault location of distribution network based on mds and dbscan clustering. In *2020 39th Chinese Control Conference (CCC)*, 6146–6150. IEEE.

Seem, J.E. and House, J.M. (2009). Integrated control and fault detection of air-handling units. *HVAC&R Research*, 15(1), 25–55.

Strehl, A. and Ghosh, J. (2002). Cluster ensembles—a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, 3(Dec), 583–617.

Tun, W., Wong, J.K.W., and Ling, S.H. (2021). Hybrid random forest and support vector machine modeling for hvac fault detection and diagnosis. *Sensors*, 21(24), 8163.

Venkatasubramanian, V., Rengaswamy, R., and Kavuri, S.N. (2003a). A review of process fault detection and diagnosis: Part ii: Qualitative models and search strategies. *Computers & chemical engineering*, 27(3), 313–326.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., and Yin, K. (2003b). A review of process fault detection and diagnosis: Part iii: Process history based methods. *Computers & chemical engineering*, 27(3), 327–346.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S.N. (2003c). A review of process fault detection and diagnosis: Part i: Quantitative model-based methods. *Computers & chemical engineering*, 27(3), 293–311.

Wu, D., Yang, Q., Tian, F., and Zhang, D.X. (2010). Fault diagnosis based on k-means clustering and pnn. In *2010 Third International Conference on Intelligent Networks and Intelligent Systems*, 173–176. IEEE.

Yang, H., Zhang, T., Li, H., Woradechjumroen, D., and Liu, X. (2014). Hvac equipment, unitary: Fault detection and diagnosis. *Encyclopedia of Energy Engineering and Technology, 2nd ed.; CRC Press: Boca Raton, FL, USA*, 854–864.