AI for Sustainable Building Operations: Data-Driven Anomaly Detection in Ventilation Systems

Anahid Wachsenegger, Adam Buruzs, Anabel Dautović, Miloš Šipetić, Laura Bernadó, Pedro Casas

AIT Austrian Institute of Technology, Vienna, Austria, Center of Digital Safety and Security

name.lastname@ait.ac.at

ABSTRACT

Detecting deviations in building time series data is essential for robust heating, ventilation, and air conditioning (HVAC) operation and energy-efficient facility management. In practice, however, building management system (BMS) data are often incomplete, heterogeneous, and lack reliable fault labels. This paper presents a benchmarking and feasibility study of data-driven anomaly detection on multivariate air-handling unit (AHU) time series data under realistic deployment constraints. We construct a unified dataset and define a domain-informed rule-based baseline as an interpretable operational reference and source of weak labels. We further evaluate classical unsupervised methods and representation-learning approaches using Temporal Convolutional Network (TCN) and Time Series Mixer (TSMixer) autoencoders, considering both a joint multivariate representation of all selected sensors and subsystem-based representations in which sensors are grouped by AHU function. Additionally, SHapley Additive exPlanations-based (SHAP) attribution is used to improve interpretability by identifying the sensor-level contributions to detected deviations. The results show that rule-based methods capture explicitly defined conditions, while data-driven approaches identify additional statistically unusual and temporally structured deviations, with representation-learning models flagging 1.1–1.4% of windows in the global setting and up to 4.7% in subsystem-based analyses. High-consensus events ($\approx 0.8\%$) occur during temporally localized episodes with agreement across multiple models, indicating robust, structured deviations. These detections represent candidate anomalies that require further validation. Our results show that combining rule-based, classical, and representation-learning methods provides complementary insights into AHU behavior and helps screen for relevant deviations in performance and energy use.

1. INTRODUCTION

Detecting deviations in building time series data is a prerequisite for reliable heating, ventilation, and air conditioning (HVAC) operation and for supporting energy-efficient facility management. In large facilities such as transportation hubs, air handling units (AHUs) generate multivariate sensor data under varying operational and environmental conditions. In practice, however, building management system (BMS) data are often incomplete, irregularly sampled, and heterogeneous across assets, while reliable fault labels are rarely available (Katipamula & Brambley, 2005b; Z. Chen et al., 2023; Matetić, Štajduhar, Wolf, & Ljubic, 2023; Bellanco, Fuentes, Vallès, & Salom, 2021; Ranade, Provan, El-Din Mady, & O'Sullivan, 2020).

In operational settings, rule-based methods remain dominant for anomaly detection because they encode expert knowledge through transparent thresholds and logical conditions (Katipamula & Brambley, 2005b, 2005a). However, they require substantial manual effort and cannot cover all relevant operating conditions. Data-driven methods, including classical clustering and representation-learning approaches, offer greater flexibility by capturing multivariate and temporal patterns that are difficult to define manually (Z. Chen et al., 2023; Zamanzadeh Darban, Webb, Pan, Aggarwal, & Salehi, 2024). Their practical adoption remains limited due to preprocessing requirements and the difficulty of validating outputs without reliable ground truth.

Despite extensive research, systematic comparisons of rule-based, classical, and representation-learning approaches on the same real-world HVAC datasets remain scarce, particularly under limited fault labels and heterogeneous sensor configurations (Matetić et al., 2023). It is therefore still unclear how these methods differ in the deviations they detect and how they can be combined in operational energy management workflows (Z. Chen et al., 2023).

This work addresses this gap through a staged benchmarking and feasibility study for anomaly detection in multivariate AHU time series data under realistic deployment constraints.

We combine domain-informed rule-based detection, classical clustering, and representation-learning approaches on real operational data with sparse measurements, heterogeneous sensors, and limited labels. The objective is not to confirm faults or quantify energy savings, but to establish a scalable screening layer for identifying potentially relevant deviations. Interpretability methods are used to link detections to specific sensors and time intervals, supporting expert validation and reducing manual labeling effort.

The main contributions of this work are as follows:

- **C1. Rule-based baseline and weak labeling.** Design of a domain-informed rule-based pipeline that provides an interpretable operational reference and generates weak labels reflecting expected system behavior.
- **C2. Classical unsupervised analysis.** Evaluation of clustering-based methods for identifying deviations in multivariate AHU operating regimes beyond explicitly defined rules.
- **C3. Representation-learning benchmark.** Assessment of sequence-based models for capturing temporally structured deviations in AHU time series data.

Based on these contributions, the paper addresses the following research questions (RQs):

- **RQ1.** To what extent can classical clustering methods identify deviations in AHU operation that are not captured by rule-based detection?
- **RQ2.** To what extent do representation-learning approaches detect temporally structured deviations beyond classical feature-based methods?
- **RQ3.** How do rule-based, classical, and representation-learning approaches complement each other in identifying relevant deviations in real-world AHU data?

The remainder of the paper is organized as follows. Section II reviews the relevant state of the art. Section III describes the case-study systems, data preparation, and anomaly-detection methodology. Section IV presents the experimental results. Section V concludes the paper and discusses practical implications and future work.

2. STATE OF THE ART

This section reviews key paradigms for HVAC anomaly detection, including rule-based, classical unsupervised, and representation-learning approaches. While well studied individually, systematic comparisons under real-world conditions remain limited. The review positions this study relative to existing practice and recent data-driven work.

2.1. Rule-based anomaly detection and weak labeling in building systems

Rule-based fault and anomaly detection remains the dominant approach in practical HVAC diagnostics. Expert knowledge is encoded as thresholds or logical rules describing implausible operating conditions, such as simultaneous heating and cooling or setpoint violations (Katipamula & Brambley, 2005b, 2005a; Bellanco et al., 2021; Ranade et al., 2020). These methods are widely used due to their interpretability, low data requirements, and ease of integration into building automation systems. However, rule-based approaches require significant manual effort and cannot fully capture complex multivariate behavior. Their transferability across buildings is limited by differences in system configuration and control strategies (Matetić et al., 2023; Z. Chen et al., 2023). Data-driven FDD methods learn system behavior from historical data and can capture more complex relationships (Z. Chen et al., 2023; Zhang, Saeed, & Sadeghian, 2023). In practice, however, labeled fault data are scarce, motivating semi-supervised and weakly supervised approaches that rely on heuristic or expert-derived labels (Ratner et al., 2017; Z. Chen et al., 2023). In this context, rule-based systems can serve as a source of weak labels, enabling hybrid approaches that combine rule-based screening with data-driven models (Liao, Cai, Cheng, Dubey, & Rajesh, 2021; Youssef, Guarino, Sibilio, & Rosato, 2023).

Overall, rule-based methods remain essential in practice but are limited in coverage, motivating complementary data-driven approaches.

2.2. Classical clustering methods for anomaly detection

Classical unsupervised methods, including centroid-based, probabilistic, density-based, and isolation-based approaches, are commonly used when labeled data are unavailable (Z. Chen et al., 2023; Matetić et al., 2023). In HVAC applications, these methods are typically applied to feature representations of multivariate sensor windows to identify dominant operating regimes and deviations from them. Representative techniques include K-Means (MacQueen, 1967), Gaussian Mixture Models (GMM) (Dempster, Laird, & Rubin, 1977), Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester, Kriegel, Sander, & Xu, 1996), and Isolation Forest (Matetić et al., 2023; Z. Chen et al., 2023; Bellanco et al., 2021). These methods are attractive due to their simplicity, computational efficiency, and suitability for exploratory analysis. In building systems, clusters often correspond to recurring operating modes driven by schedules, environmental conditions, or control strategies, while low-density or poorly assigned samples are interpreted as deviations. However, classical methods rely on feature engineering and struggle to capture nonlinear dependencies and temporal dynamics. Their performance is sensitive to preprocessing and parameter se-

lection, and their ability to detect subtle or context-dependent deviations is limited (Z. Chen et al., 2023; Matetić et al., 2023; Mirnaghi & Haghighat, 2020). As a result, they are most effective for identifying coarse regime changes or strongly separated patterns rather than complex temporal behavior.

2.3. Representation-learning-based anomaly detection for time series data

Recent approaches shift from engineered features to learned representations of time series data. Neural sequence models are trained to encode normal behavior, and deviations are identified through reconstruction errors, prediction errors, or anomalies in latent space (Zamanzadeh Darban et al., 2024; Goodfellow, Bengio, & Courville, 2016). Common models include autoencoders, recurrent networks, temporal convolutional networks, and transformer-based architectures (El Mokhtari & McArthur, 2024; Sipetic, Schöny, & Catal, 2024; Hochreiter & Schmidhuber, 1997; Vaswani et al., 2017; Kingma & Welling, 2013). These methods can model nonlinear relationships, temporal dependencies, and cross-variable interactions more effectively than classical approaches (Zhang et al., 2023; Z. Chen et al., 2023). However, most reported results rely on simulated or laboratory data, while evidence from long-term real building operation remains limited (Zhang et al., 2023). In practical settings, their deployment is constrained by data quality, limited interpretability, and the lack of reliable validation data. A key advantage of representation learning is the ability to capture temporally structured deviations. In addition to reconstruction-based scoring, latent representations can be analyzed using clustering or density models, enabling the detection of patterns that are not apparent in static feature spaces (Zamanzadeh Darban et al., 2024). At the same time, these methods require careful tuning and should be viewed as complementary to rule-based and classical approaches rather than as replacements.

2.4. Explainable AI for time series anomaly detection

As anomaly detection models become more complex, explainability becomes essential for practical use. In HVAC systems, anomaly scores must be interpretable to support diagnosis and operational decision-making (Saeed & Omlin, 2023). This has led to increasing use of explainable AI techniques in time series analysis (Troncoso-García, Martínez-Ballesteros, Martínez-Álvarez, & Troncoso, 2023). Feature attribution methods such as permutation importance and SHapley Additive exPlanations (SHAP) provide insights into which variables contribute to a detected deviation (Breiman, 2001; Lundberg & Lee, 2017). In sequence models, attention mechanisms are sometimes used to highlight relevant time steps, although their interpretation remains limited (Vaswani et al., 2017; Saeed & Omlin, 2023). Despite these advances, explaining multivariate time series anomalies remains challenging due to temporal dependencies, high dimensionality, and coupled physical processes. In building systems, explanations must align with engineering reasoning and support actionable decisions. Consequently, explainability remains a key requirement for the practical adoption of data-driven anomaly detection methods.

3. METHODOLOGY

This section describes the case-study systems, data preparation, anomaly-detection methods, and experimental setups used to ensure comparability across approaches.

3.1. Ventilation System Description

The dataset comprises sensor data from four large buildings. In this study, the analysis focuses on the transportation terminal (main hub) as the primary case study due to its higher data completeness. In this case study, the BMS provides AHU time series data at 15-minute resolution. The system includes supply- and exhaust-air paths with speed-controlled fans, heating and cooling coils, and heat recovery. Heating is provided via district heating, and cooling via chilled water. Monitored signals include physical measurements, actuator states, setpoints, and system variables, reflecting typical real-world HVAC operation. The system follows the general configuration illustrated in Figure 1.

3.2. Data Collection and Preprocessing

The dataset consists exclusively of real-world measurements collected from operational building management systems; no synthetic or simulated data were used. The dataset contains 755 sensor streams recently collected over two years (2023-24) across 5 different buildings. Signals include temperature, pressure, humidity, CO_2, actuator positions, and control variables. Metadata from domain experts supports sensor interpretation. Sensor identifiers follow the original BMS naming convention, where prefixes indicate subsystems (e.g., ZUL: supply air, ABL: exhaust air) and suffixes denote measured variables (e.g., TE: temperature, AF: absolute humidity). Refer to Tab. 5 for further detailed explanations about the sensor names. Exploratory analysis showed that only a subset of sensors exhibits continuous temporal dynamics suitable for multivariate time series modeling. This motivates two setups: a joint multivariate representation of analog sensors and a grouped subsystem-based representation. Data quality issues include missing values, irregular sampling, and low-variability or event-driven signals. Preprocessing includes:

1. *Signal selection*: We focus our study on the sensor data stemming from a single building, where we had the most detailed description and specification of the HVAC components, corresponding also to the most modern building. This reduced the dataset to 290 sensors. With the idea to further investigate transfer learning of the models to other buildings with similar systems.
2. *Alignment*: All time series are resampled to a uniform

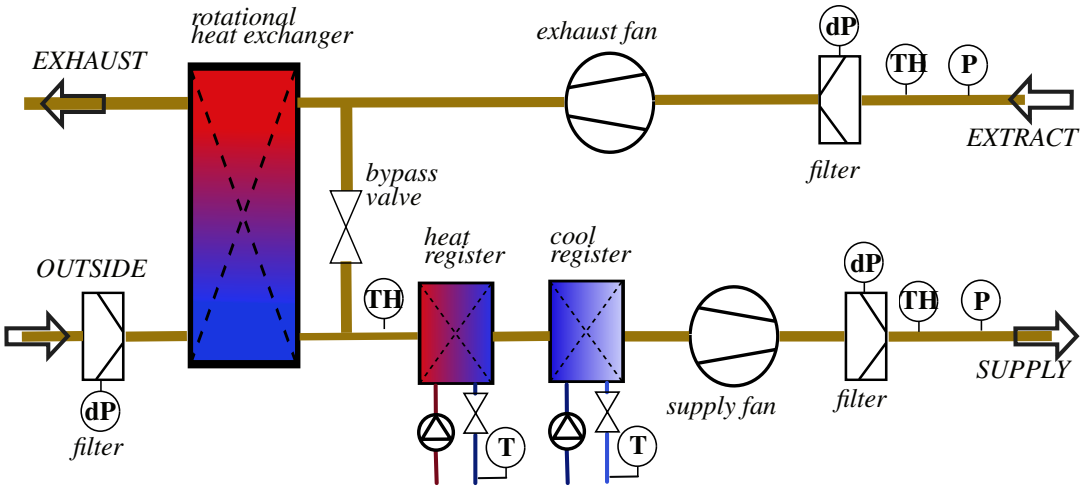

Figure 1. Schematic of the air handling unit considered in this study. Temperature, humidity, and pressure sensors are indicated in the highlighted circles.

30-minute grid, resulting in a total of 70177 samples per sensor.

3. *Training dataset split*: We split the dataset into training and test; further details will be given in Sec.3.4.
4. *Missing-value handling*: missing values were handled by applying median imputation based on training data statistics.
5. *Standardization*: we standardized the data using z-score normalization computed from the training set mean and standard deviation.

To provide a quantitative overview, the selected subset used for modeling comprises approximately N aligned time steps after preprocessing. Using a sliding window of 6 hours (24 samples at 15-minute resolution), this results in approximately M windows for analysis. A time-based split is applied, with 95% of the windows used for training and 5% for testing. For representation-learning models, an additional 10% of the training data is used as a validation set for early stopping. All preprocessing steps, including imputation and normalization, are performed using statistics computed on the training set only to avoid data leakage.

3.3. Rule-based Baseline and Weak Labeling

A domain-informed rule-based pipeline is used as an operational baseline. Existing BMS fault detection rules provide reference anomalies, which are used to remove known anomalous periods from the training data and to compare with detections from data-driven models. The rules encode typical fault conditions and implausible behavior, such as simultaneous heating and cooling or setpoint violations (Table 1). Short events are filtered by merging events within 30 minutes and discarding events shorter than one hour. The resulting rule-based outputs provide interpretable weak labels, but capture only part of energy-relevant behavior, as the current rule set focuses mainly on thermal conditions while aspects such as fan operation are assessed manually. A simple per-sensor data-driven baseline is therefore included as a complementary reference (see Appendix).

3.4. Experimental Setups

Two complementary setups are defined:
Scenario 1: Analog-only modeling. Continuous physical signals (e.g., temperature, pressure, humidity) are used to capture global process behavior.
Scenario 2: Grouped sensor modeling. Sensors are grouped into subsystems (e.g., supply air, exhaust air, pumps) to improve interpretability. The grouping is adopted to provide a reproducible experimental setup and to assess the ability of the selected methods to detect anomalies within subsystem-specific contexts.

In both cases, the unit of analysis is a sliding time window capturing short-term temporal dynamics. The window length used in both scenarios was chosen to be 6 hours, which captures sub-daily operational cycles and relevant HVAC dynamics, providing a balance between statistical robustness and temporal resolution. Data are filtered, imputed, standardized, and aligned across all methods. Details on model configuration are given in the following sections.

3.5. Classical Clustering-Based Analysis

Classical unsupervised methods are applied to window-level features (mean, standard deviation, minimum, maximum), optionally reduced via Principal Component Analysis (PCA) (Hotelling, 1933). The evaluated methods include K-Means, GMM, DBSCAN, and Isolation Forest (Liu, Ting, & Zhou, 2008). These methods identify statistically unusual patterns

Table 1. Rule-based fault and anomaly detection rules applied to the AHU dataset.

No.	Indicator	Conditions	Significance
2	System off	Supply fan frequency signal $<$ 10% **OR** Exhaust fan frequency signal $<$ 10%	Indicator
3	System off	Supply fan frequency signal $<$ 5 Hz **OR** Exhaust fan frequency signal $<$ 5 Hz	Indicator
4	System off	Supply air pressure $<$ 50 Pa **OR** Exhaust air pressure $<$ 50 Pa	Indicator
5	CO_2 implausible	(Value $<$ 350 ppm **OR** Value $>$ 1000 ppm) **AND** Ventilation system status = ON	Fault
6	Heating on a hot day	Heating valve status $>$ 10% **AND** Outside temperature measurement $>$ 25°C **AND** Ventilation system status = ON	Fault
7	Cooling on a warm day	Cooling valve status $>$ 10% **AND** Outside temperature measurement $<$ 10°C **AND** Dehumidification status = OFF **AND** Ventilation system status = ON	Fault
8	Cooling and heating simultaneously	Cooling valve status $>$ 10% **AND** Heating valve status $>$ 10%	Fault
9	Heating and reheating simultaneously	Heating valve status $>$ 10% **AND** Reheating valve status $>$ 10%	Fault
10	Overheating	Heating valve status $>$ 10% **AND** Supply temperature measurement $>$ Supply temperature setpoint + 2°C	Fault
11	Overcooling	Cooling valve status $>$ 10% **AND** Supply temperature measurement + 2°C $<$ Supply temperature setpoint	Fault
12	Wasting heat	Heat recovery signal $<$ 90% **AND** Heating valve $>$ 10% **AND** Extract air temperature measurement $>$ Outside temperature measurement **AND** Ventilation system status = ON	Waste
13	Wasting cool	Heat recovery signal $<$ 90% **AND** Cooling valve $>$ 10% **AND** Extract air temperature measurement $<$ Outside temperature measurement **AND** Ventilation system status = ON	Waste

in the feature space based on distance, likelihood, density, or isolation criteria. The results are interpreted as candidate deviations rather than confirmed anomalies. Model configurations and all hyperparameters are summarized in Table 4.

3.6. Representation-Learning-Based Analysis

Sequence models are used to learn representations of multivariate time series windows. Two autoencoder architectures are evaluated: a Temporal Convolutional Network (TCN) (Bai, Kolter, & Koltun, 2018) and a Time Series Mixer (TSMixer) (S.-A. Chen, Li, Arik, Yoon, & Pfister, 2023). These models provide complementary inductive biases: TCN captures temporal dependencies through convolutional structures, while TSMixer models temporal and channel interactions through mixing operations.

Both models are trained using mean squared reconstruction error with the Adam optimizer. Candidate deviations are identified from reconstruction error and latent-space clustering on learned embeddings, resulting in six anomaly views per setup. This multi-view formulation captures temporally structured deviations, supports cross-validation across model assumptions, and requires operational validation for interpretation.

3.7. Explainability Approach

To support interpretation, reconstruction-based anomaly scores are analyzed using both model-based contributions and SHAP-based attribution. First, sensor-level contributions are derived directly from the reconstruction error of the autoencoders. For a given time window, the mean squared reconstruction error is computed per sensor, providing a model-specific measure of how strongly each variable contributes to the detected deviation. Second, SHAP is applied to the reconstruction score by wrapping the autoencoder to output a scalar reconstruction error per window. SHAP values are computed at the sensor–time level, attributing the reconstruction error to individual inputs within the time window. This yields local explanations that highlight which sensors and time steps contribute most strongly to elevated anomaly scores. The resulting explanations are analyzed at three levels: (i) sensor-level importance by aggregating SHAP values over time, (ii) timestep-level importance by aggregating over sensors, and (iii) full sensor–time attribution using heatmaps. In addition, SHAP values are aggregated across windows to obtain global sensor importance, identifying variables that consistently contribute to elevated anomaly scores.

4. EXPERIMENTAL RESULTS

This section presents results for the rule-based baseline, classical clustering methods, and representation-learning approaches. All methods were implemented in `Python 3.10` and evaluated on the same preprocessed dataset to ensure comparability. Sensor streams were resampled, aligned into fixed-length windows, and standardized; PCA-based feature compression was applied in the classical pipeline. The results are interpreted as an initial assessment of candidate deviations in AHU operation rather than confirmed faults, requiring further validation.

4.1. Results for the Rule-based Baseline

The rule-based pipeline serves as an operational baseline and as a source of weak labels. Figure 2 summarizes the rule-triggered events over the analyzed period.

For the analyzed period, only one rule-triggered event group was detected. It corresponded to Rule 5 and comprised 43 readings. This rule is triggered when the ventilation system is on, and the measured CO_2 concentration is below 350 ppm or above 1100 ppm. Here, the values were below 350 ppm, suggesting implausible sensor behavior, since atmospheric CO_2 is around 420 ppm. These results show that the rule

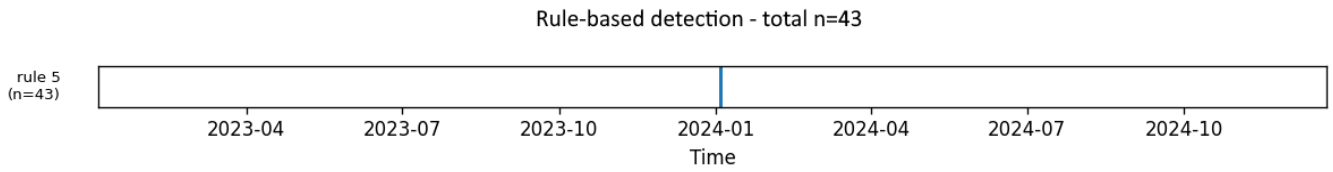

Figure 2. Prevalence of rule-triggered events over the analyzed period. Only one incident group involving rule 5 was detected.

set can clearly identify implausible conditions, but its limited number of detections highlights its restricted coverage. The rule-based baseline captures only explicitly defined conditions and therefore serves as a partial operational reference rather than complete ground truth. To provide an additional comparison, a simple per-sensor data-driven baseline is included in the Appendix (Section 7).

4.2. Results for Classical Clustering (RQ1)

This section evaluates whether classical unsupervised methods identify candidate deviations beyond those covered by the rule-based baseline. The rule-based outputs are treated as a limited reference, not as exhaustive truth labels. The classical results are therefore interpreted in terms of detector agreement, temporal concentration of flagged windows, and the structure of the identified operating regimes.

4.2.1. Scenario 1: Analog-Only Sensor Modeling

In the first scenario, classical unsupervised methods were applied to aligned windows constructed from analog sensor data during System-On operation (i.e., ventilation system operating state) in the main hub. After filtering, the dataset contained 76,676 System-On rows, of which 36,804 (48.0%) were used for training and 39,872 (52.0%) for testing. This split was chosen to ensure that both training and test sets capture the full seasonal variability present in the two-year dataset, thereby avoiding seasonal bias in the learned representations. All variables were standardized using z-score normalization estimated on the training set. Model-specific criteria were used to flag candidate deviations: K-Means identifies windows with large distances to cluster centroids, GMM flags low-likelihood observations, DBSCAN treats points outside dense regions as noise, and Isolation Forest detects observations that are easily isolated. The resulting detections differ in coverage, with K-Means producing the most and DBSCAN the fewest, indicating that the methods respond to different structures in the feature space.

Figure 3 shows that the flagged windows are temporally localized and only partially shared across models. Figure 4 shows that most windows are concentrated in a small number of dominant regimes, while flagged windows tend to occur near regime boundaries or outside dense regions. In practical terms, this setup captures unusual combinations of continuous variables such as temperature, pressure, airflow, and humidity. A typical example would be a window in which temperature and airflow jointly evolve in a way that differs from the dominant operating regimes, even though no explicit threshold violation is present. This differs from rule-based detection, which only flags conditions that have been predefined. The classical methods, therefore, broaden the analysis from known rule violations to statistically unusual operating patterns. These results reflect deviations at the level of the global system state, where anomalies correspond to unusual combinations of variables across the entire AHU. As a result, detections tend to occur in concentrated bursts when the overall operating regime deviates from typical patterns.

4.2.2. Scenario 2: Grouped Sensor Modeling

In the second scenario, the same methods were applied separately to windows constructed for physically meaningful sensor groups in the main hub during System-On operation. In contrast to the analog-only setup, this formulation focuses on subsystem behavior by grouping related analog and selected digital signals according to their functional role. Candidate deviations were again defined using the model-specific criteria introduced above. The grouping does not change the detection principle itself, but changes the notion of normality by restricting the representation to smaller and more homogeneous subsystems. This makes the resulting detections easier to associate with functional subsystems such as supply air, exhaust air, pumps, or temperature-related behavior.

Figure 5 shows that the detections remain localized in time and only partially shared across methods. The anomaly detections are sparse and unevenly distributed over time, shown as isolated flags, which suggests sensitivity to localized deviations rather than persistent anomalies. However, we observe short periods of vertical alignment across multiple methods, indicating possible temporally coherent events, pointing to potentially meaningful subsystem-level anomalies.

Figure 6 shows the PCA projection of the `abluft_all` windows and reveals that most windows remain concentrated in a limited number of subsystem operating regimes, while flagged windows again tend to lie near regime transitions or outside dense cluster cores, indicating that anomalies might correspond to deviations from typical subsystem states, which often occur during transitions rather than within steady operation.

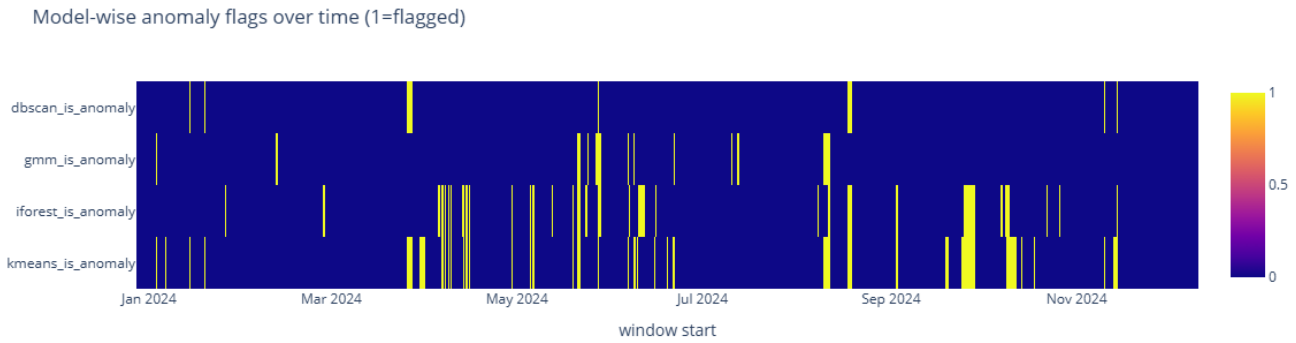

Figure 3. Temporal agreement of flagged windows for the analog-only setup. Each row corresponds to one model, and each column to one aligned time window. Flags occur in localized bursts and overlap only partially.

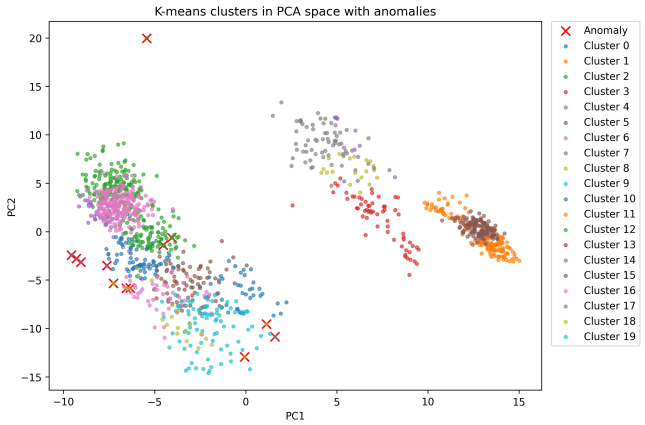

Figure 4. Two-dimensional PCA projection of the analog-only window representations colored by K-Means cluster assignment. Flagged windows are highlighted separately. Most windows fall into a few dominant regimes, while flagged windows tend to lie near regime boundaries or outside dense cluster cores.

Compared with the analog-only setup, the grouped formulation offers a more direct physical interpretation. For example, a window may be flagged within a supply-air or temperature-related group even when the global feature space does not show a strong deviation. This makes grouped modeling useful for identifying candidate subsystem inconsistencies that are not explicitly covered by the available rule set. Compared with the analog-only setup, the grouped formulation captures deviations at the subsystem level. This results in more distributed detections over time, as localized inconsistencies within individual subsystems can be identified even when the global system state appears normal. Consequently, grouped modeling provides a more fine-grained and physically interpretable view of system behavior.

Summary with respect to RQ1. The classical clustering results show that unsupervised methods detect candidate deviations beyond rule-based detection by identifying statistically unusual operating regimes. The analog-only setup captures global system-state deviations, while grouped modeling reveals localized subsystem inconsistencies. These representations are complementary: global modeling highlights coherent system-level events, whereas grouped modeling improves sensitivity to local effects and supports physical interpretation. Combined, they enable structured screening, where global detections indicate high-confidence events and subsystem-level detections support localization and early fault identification.

4.3. Results for Representation Learning (RQ2)

This section evaluates whether sequence-based models reveal candidate deviations beyond those captured by static feature-based methods. Unlike classical clustering, representation-learning models explicitly account for temporal structure within the input windows. Two autoencoder models were evaluated: TCN and TSMixer. For each representation, deviations were derived from reconstruction error and latent-space structure, resulting in six complementary views: TCN reconstruction, TCN latent K-Means, TCN latent GMM, TSMixer reconstruction, TSMixer latent K-Means, and TSMixer latent GMM. This multi-view setup captures temporal and distributional deviations while supporting cross-method validation.

4.3.1. Scenario 1: Analog-Only Sensor Modeling

In the first representation-learning scenario, detection was performed on a joint multivariate representation of all selected analog sensors, corresponding to the analog-only setup. The resulting feature set, denoted as `analog_sensors`, comprised 18 analog HVAC measurements and was modeled jointly as one aligned multivariate time series representation. Aligned windows were processed using a TCN autoencoder and a TSMixer autoencoder. In addition to reconstruction-based scores, K-Means and GMM were applied in the latent spaces of both models. Table 2 summarizes the fraction of flagged

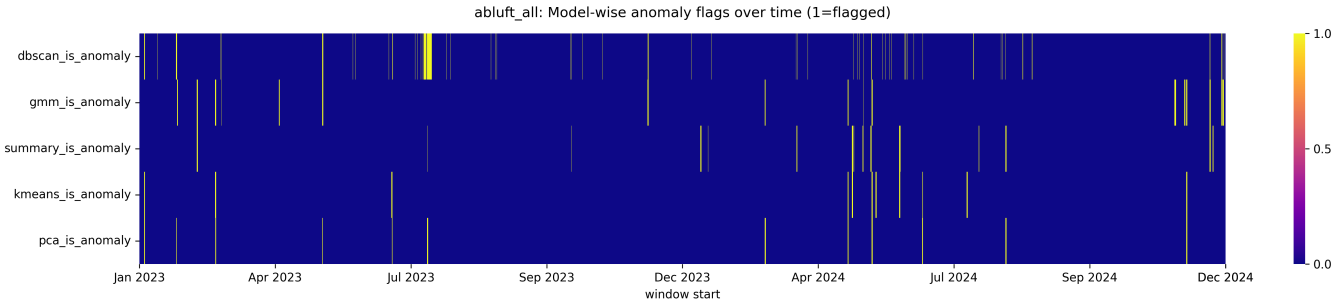

Figure 5. Model-wise flag heatmap for the `abluft_all` group over time. Rows represent the individual methods and columns the ordered time windows. Flagged windows are shown in yellow.

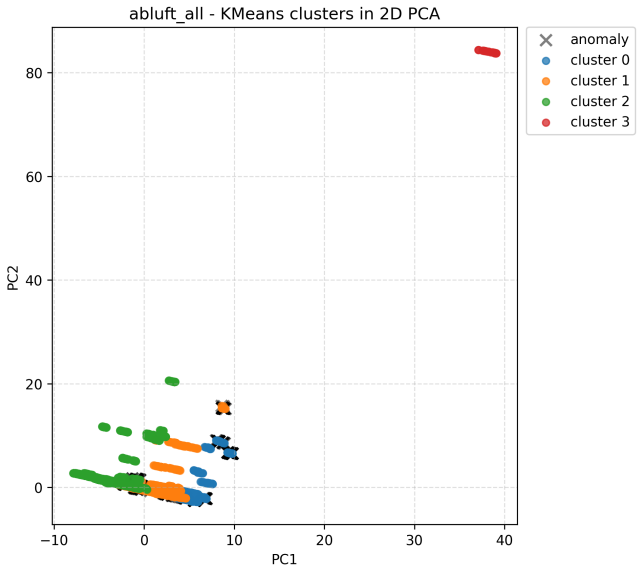

Figure 6. Two-dimensional PCA visualization of the `abluft_all` windows under K-Means clustering. Points are colored by cluster membership, while flagged windows are marked separately.

windows for the `analog_sensors` feature set. Across all six views, the rates range from 1.14% to 1.36%. No single view dominates the result, and all six views indicate a similar level of rare and localized deviations.

The temporal flag map for the `analog_sensors` feature set (Fig. 7) shows that flagged windows are concentrated in short episodes rather than evenly distributed. The strongest episode occurs around 2023-12-03 to 2023-12-04, where all six views flag consecutive windows, indicating a robust system-wide deviation rather than a detector-specific artifact. Sensor-level reconstruction errors for a representative window (Fig. 8) identify `L_01 HRG_RL_M` and `L_01 ZUL_TE2M` as dominant contributors, followed by `ZUL_TE3M` and `NHR_RL_M`. This suggests that the event is mainly driven by return-air and supply-air temperature dynamics within a specific thermal subsystem. A meteorological check indicates that early December 2023 coincided with heavy snowfall and low temperatures, which may have shifted HVAC operating regimes or disturbed sensor behavior. Thus, the anomaly episode may reflect a weather-related operational response rather than a confirmed fault.

Compared with the classical methods, the main difference here is temporal modeling. A classical method may flag a window because its aggregated statistics differ from the dominant clusters. By contrast, the sequence models can flag a window whose temporal evolution is unusual even when static summary statistics remain less distinctive. In practice, this is useful for capturing gradual drifts, oscillatory control behavior, or coordinated temporal changes across multiple variables.

4.3.2. Scenario 2: Grouped Sensor Modeling

In the second representation-learning scenario, detection was performed separately on physically meaningful sensor groups. Five groups were evaluated: `zuluft_all`, `abluft_all`, `pumps`, `zuluft_temps`, and `abluft_temps`. For each group, aligned multivariate windows were modeled using a TCN autoencoder and a TSMixer autoencoder. As in the analog-only case, K-Means and GMM were also applied in the latent spaces of both models. Table 3 summarizes the resulting flag rates. Most groups exhibit values close to 1%, while `zuluft_temps` stands out with rates between 2.47% and 4.73%. This indicates that the strongest temporally structured deviations are concentrated in the supply-air temperature subsystem.

The consensus statistics reinforce this result. `zuluft_temps` shows the highest mean vote count, consensus rate, and strong-consensus rate, clearly above the other groups.

Figure 9 shows short episodes of concentrated flags rather than uniformly distributed detections. While `abluft_all` exhibits mostly scattered detections, `zuluft_all` shows more pronounced temporal clustering and stronger agreement across models. In particular, `zuluft_all` exhibits a pronounced event around 2024-08-11 to 2024-08-12, during which multiple consecutive windows are flagged by all 6 views.

Table 2. Representation-learning flag rates (%) for the `analog_sensors` feature set across all six views, together with consensus statistics.

Feature set	TCN Recon.	TCN Latent K-Means	TCN Latent GMM	TSMixer Recon.	TSMixer Latent K-Means	TSMixer Latent GMM
analog_sensors	1.36	1.16	1.14	1.28	1.20	1.15

Consensus statistics across representation-learning views

Feature set	Mean vote count	Consensus rate (%)	Strong consensus rate (%)
analog_sensors	0.0730	2.18	0.84

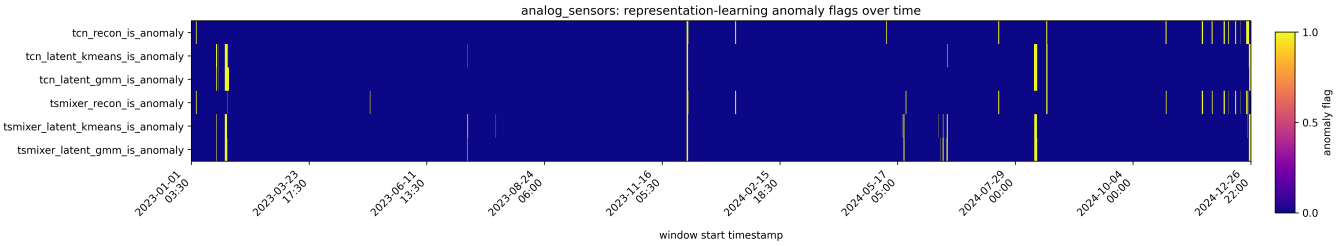

Figure 7. Temporal flag map for the `analog_sensors` feature set using six representation-learning views. Flags remain limited over most of the observation period but form a pronounced temporally coherent episode around early December 2023.

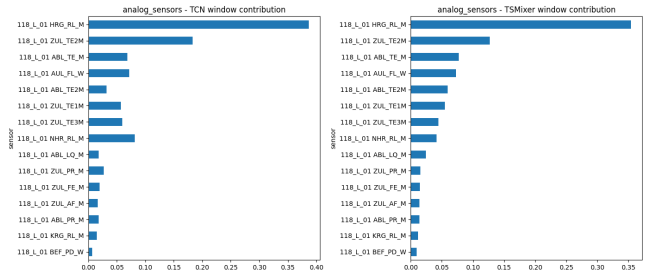

Figure 8. Sensor-level reconstruction-error contributions for a representative high-consensus window in `analog_sensors`. Both the TCN and TSMixer identify `HRG_RL_M` and `ZUL_TE2M` as dominant contributors.

Sensor-level reconstruction errors for a high-consensus `zuluft_all` window indicate that the event was mainly driven by supply-air temperature variables, with secondary contributions from airflow and fan-control signals. This suggests a temperature-dominated subsystem deviation with associated airflow effects, highlighting the value of grouped modeling for isolating subsystem-specific behavior. A meteorological check suggests that the August 2024 episode may be partly related to elevated summer temperatures around 11–12 August. Although no strong weather fluctuations were observed, the pattern is consistent with a possible HVAC response to external conditions rather than random variation.

Summary with respect to RQ2. The representation-learning results show that sequence-based models provide a complementary view to classical clustering by directly modeling temporal structure. This enables the detection of candidate deviations that evolve, such as coordinated temperature dynamics or control-related patterns, rather than only static differences in aggregated features. From a practical perspective, these models can support deeper screening of subsystem behavior, including sustained irregularities that do not violate fixed thresholds. While the current analysis does not prove operational or energy impact, it provides a basis for prioritizing cases for expert inspection.

4.4. Cross-Method Interpretation (RQ3)

The three methodological paradigms provide complementary views of system behavior: rule-based methods capture predefined conditions, classical clustering identifies unusual operating regimes, and representation-learning methods detect temporally structured deviations such as drift or oscillations. In the analyzed dataset, flagged windows were often concentrated in limited periods rather than uniformly distributed; this observation is system-specific and should not be generalized. Overall, the approaches are complementary rather than interchangeable, and their outputs should be interpreted as candidate deviations. The next step is to assess which detections are operationally valuable for improving performance or reducing energy use. The comparison between global and subsystem-based representations further shows that anomaly detection operates at multiple levels: global models capture system-wide deviations, while grouped models reveal localized effects that support fault localization.

4.4.1. Explainability via SHAP Analysis

To improve interpretability of the detected candidate deviations, SHAP was applied to the reconstruction-based anomaly scores of the TCN and TSMixer models. Instead of explain-

Table 3. Representation-learning flag rates (%) for each sensor group and view, together with consensus statistics across the six views.

Group	TCN Recon.	TCN Latent K-Means	TCN Latent GMM	TSMixer Recon.	TSMixer Latent K-Means	TSMixer Latent GMM
zuluft_temps	4.66	3.43	4.67	4.73	2.47	3.39
zuluft_all	1.03	1.11	1.10	1.09	1.13	1.15
abluft_temps	1.29	1.20	1.08	1.08	1.22	1.55
abluft_all	1.08	1.05	1.05	1.05	0.98	1.02
pumps	0.96	0.98	0.95	0.96	0.95	0.98

Consensus statistics across representation-learning views

Group	Mean vote count	Consensus rate (%)	Strong consensus rate (%)
abluft_all	0.0624	1.79	0.94
abluft_temps	0.0742	1.70	0.65
pumps	0.0579	1.54	0.64
zuluft_all	0.0661	1.69	0.97
zuluft_temps	0.2336	5.05	3.49

ing binary anomaly flags, SHAP attributed the continuous reconstruction error of each time window to individual sensor–time inputs, indicating which variables and temporal segments contributed to elevated anomaly scores. SHAP values were computed with a gradient-based explainer using background samples from normal training data and visualized as sensor–time heatmaps.

The explanations indicate that elevated anomaly scores are typically driven by localized rather than system-wide contributions. In the August case study (Figure 11), the dominant contribution originates from a single sensor–time point, suggesting a localized and transient deviation. Such explanations can help operators narrow the analysis to specific sensors and time intervals. However, SHAP identifies statistical contributions to the model output rather than causal faults, so domain validation is required to assess operational relevance.

5. CONCLUSION, DISCUSSION, AND FUTURE WORK

5.1. Limitations

This paper presented a benchmarking and feasibility study for data-driven analysis of multivariate AHU time series data under realistic deployment constraints. A rule-based baseline, classical clustering methods, and representation-learning approaches were evaluated to identify candidate deviations in operational building management system data. Several limitations should be considered when interpreting the results. The analysis is based on data from a single building system, which may limit generalizability to other configurations and operational contexts. The dataset is proprietary and not publicly available, restricting reproducibility and benchmarking against external methods. In addition, no ground-truth fault labels were available, so evaluation relied on indirect indicators such as detector agreement and temporal consistency rather than verified faults. The study also does not quantify the impact of detected deviations on energy consumption or operational performance. Accordingly, the presented results should be viewed as a feasibility study for identifying candidate deviations rather than a validated fault detection system.

5.2. Computational Considerations

From a practical deployment perspective, the evaluated methods differ in computational requirements. Classical clustering methods such as K-Means and GMM are computationally efficient and scale well with dataset size, making them suitable for large-scale or near real-time screening. Representation-learning approaches require higher computational effort during training due to neural network optimization, but inference remains efficient once models are trained. Overall, the methods provide a scalable and interpretable screening layer for prioritizing potentially relevant deviations for expert analysis.

5.3. Conclusion and Future Work

The results show that the methodological paradigms provide complementary views of system behavior. Rule-based methods capture explicitly defined conditions, classical clustering reveals statistically unusual operating regimes, and representation-learning methods identify temporally structured deviations. These detections should therefore be interpreted as candidate deviations rather than confirmed faults. A key finding is that data representation strongly influences the results: global models capture system-wide deviations, while subsystem-based grouping improves interpretability by linking deviations to physically meaningful AHU components and supporting fault localization. Interpretability was further enhanced using SHAP-based attribution, which provided local explanations of individual deviations and global feature importance across sensors. Future work will focus on validating the operational relevance of detected deviations, quantifying their impact on energy consumption and system performance, refining sensor grouping and data representation strategies, and improving interpretability. Transfer-learning approaches will also be investigated to improve generalization across buildings and reduce the need for site-specific training data. The ultimate objective

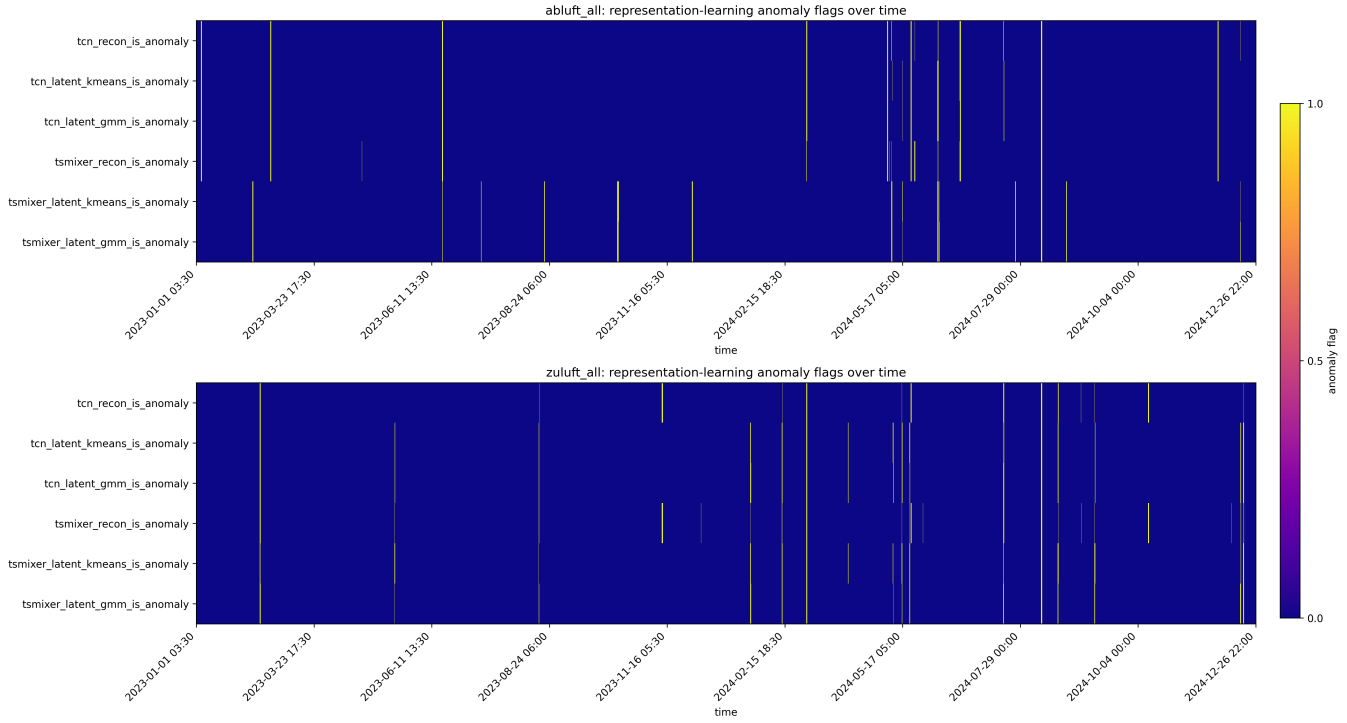

Figure 9. Temporal flag maps for the `abluft_all` and `zuluft_all` groups using six representation-learning views. Overlapping flags are more pronounced in `zuluft_all`, indicating a clearer temporally coherent subsystem-level deviation.

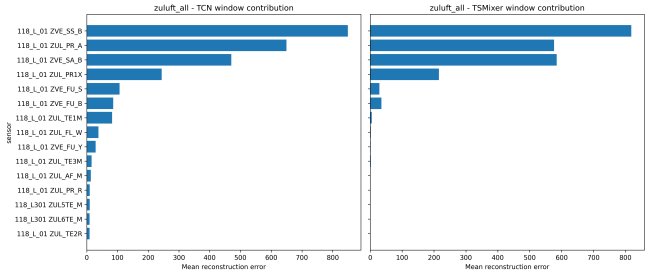

Figure 10. Sensor-level reconstruction-error contributions for a representative high-consensus window in `zuluft_all`. Supply-air temperature variables dominate the contribution profile.

is to identify detections that support actionable improvements in AHU performance and energy use.

REFERENCES

Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.

Bellanco, I., Fuentes, E., Vallès, M., & Salom, J. (2021). A review of the fault behavior of heat pumps and measurements, detection and diagnosis methods including virtual sensors. *Journal of Building Engineering*, *39*, 102254. Retrieved from `https://www.sciencedirect.com/science/article/pii/S2352710221001108` doi: https://doi.org/10.1016/j.jobe.2021.102254

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi: 10.1023/A:1010933404324

Chen, S.-A., Li, C.-L., Arik, S. O., Yoon, J., & Pfister, T. (2023). Tsmixer: An all-mlp architecture for time series forecasting. *arXiv preprint arXiv:2303.06053*.

Chen, Z., O'Neill, Z., Wen, J., Pradhan, O., Yang, T., Lu, X., ... Herr, T. (2023). A review of data-driven fault detection and diagnostics for building hvac systems. *Applied Energy*, *339*, 121030. doi: 10.1016/j.apenergy.2023.121030

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–38.

El Mokhtari, K., & McArthur, J. (2024, October). Autoencoder-based fault detection using building automation system data. *Adv. Eng. Inform.*, *62*(PC). Retrieved from `https://doi.org/10.1016/j.aei.2024.102810` doi: 10.1016/j.aei.2024.102810

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the second international conference on knowledge discovery*

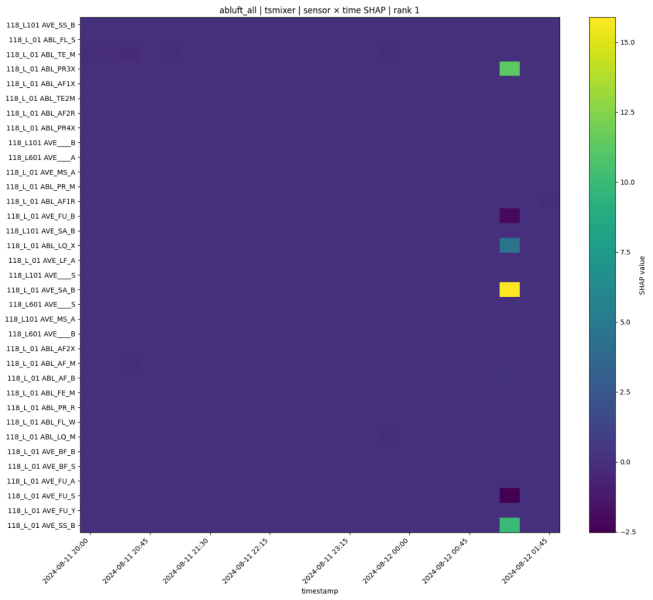

Figure 11. Sensor–time SHAP heatmap for a representative high-scoring window in the `abluft_all` group (August case). The explanation highlights a localized contribution pattern, where a small subset of sensor–time points dominates the reconstruction error.

and data mining (kdd-96) (pp. 226–231). AAAI Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.

Hochreiter, S., & Schmidhuber, J. (1997, November). Long short-term memory. *Neural Comput.*, *9*(8), 1735–1780. Retrieved from https://doi.org/10.1162/neco.1997.9.8.1735 doi: 10.1162/neco.1997.9.8.1735

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*(6), 417–441. doi: 10.1037/h0071325

Katipamula, S., & Brambley, M. R. (2005a). Methods for fault detection, diagnostics, and prognostics for building systems—a review, part ii. *HVAC&R Research*, *11*(2), 169–187. doi: 10.1080/10789669.2005.10391133

Katipamula, S., & Brambley, M. R. (2005b). Review article: Methods for fault detection, diagnostics, and prognostics for building systems—a review, part i. *HVAC&R Research*, *11*(1), 3–25. Retrieved from https://doi.org/10.1080/10789669.2005.10391123 doi: 10.1080/10789669.2005.10391123

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Liao, H., Cai, W., Cheng, F., Dubey, S., & Rajesh, P. B. (2021). An online data-driven fault diagnosis method for air handling units by rule and convolutional neural networks. *Sensors*, *21*(13), 4358. doi: 10.3390/s21134358

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *Proceedings of the 2008 eighth ieee international conference on data mining* (pp. 413–422). IEEE. doi: 10.1109/ICDM.2008.17

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in neural information processing systems* (Vol. 30).

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297). University of California Press.

Matetić, I., Štajduhar, I., Wolf, I., & Ljubic, S. (2023). A review of data-driven approaches and techniques for fault detection and diagnosis in hvac systems. *Sensors*, *23*(1), 1. doi: 10.3390/s23010001

Mirnaghi, M. S., & 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. doi: 10.1016/j.enbuild.2020.110492

Ranade, A., Provan, G., El-Din Mady, A., & O'Sullivan, D. (2020). A computationally efficient method for fault diagnosis of fan-coil unit terminals in building heating ventilation and air conditioning systems. *Journal of Building Engineering*, *27*, 100955. Retrieved from https://www.sciencedirect.com/science/article/pii/S2352710219304930 doi: https://doi.org/10.1016/j.jobe.2019.100955

Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & Ré, C. (2017, November). Snorkel: rapid training data creation with weak supervision. *Proc. VLDB Endow.*, *11*(3), 269–282. Retrieved from https://doi.org/10.14778/3157794.3157797 doi: 10.14778/3157794.3157797

Saeed, W., & Omlin, C. (2023, March). Explainable ai (xai): A systematic meta-survey of current challenges and future opportunities. *Know.-Based Syst.*, *263*(C). Retrieved from https://doi.org/10.1016/j.knosys.2023.110273 doi: 10.1016/j.knosys.2023.110273

Sipetic, M., Schöny, M., & Catal, J. (2024). Application of autoencoders on multivariate anomaly detection in building automation systems with variable selection based on semantic metadata of the facility. In *Proceedings of the 7th international conference on efficiency, cost, optimization, simulation and environmental impact of energy systems (ecos 2024)*.

Troncoso-García, A. R., Martínez-Ballesteros, M., Martínez-Álvarez, F., & Troncoso, A. (2023). A new approach based on association rules to add explainability to time series forecasting models. *Information Fusion*, *94*, 169–180. doi: 10.1016/j.inffus.2023.01.021

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is

all you need. In *Proceedings of the 31st international conference on neural information processing systems* (p. 6000–6010). Red Hook, NY, USA: Curran Associates Inc.

Youssef, M. E., Guarino, F., Sibilio, S., & Rosato, A. (2023). Experimental assessment of a preliminary rule-based data-driven method for fault detection and diagnosis of coils, fans and sensors in air-handling units. In *Sustainability in energy and buildings 2022* (Vol. 336, pp. 359–370). Singapore: Springer. doi: 10.1007/978-981-19-8769-4_34

Zamanzadeh Darban, Z., Webb, G. I., Pan, S., Aggarwal, C., & Salehi, M. (2024, October). Deep learning for time series anomaly detection: A survey. *ACM Comput. Surv.*, *57*(1). Retrieved from https://doi.org/10.1145/3691338 doi: 10.1145/3691338

Zhang, F., Saeed, N., & Sadeghian, P. (2023). Deep learning in fault detection and diagnosis of building hvac systems: A systematic review with meta analysis. *Energy and AI*, *12*, 100235. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666546823000071 doi: https://doi.org/10.1016/j.egyai.2023.100235

ACKNOWLEDGMENT

This project is funded by the Austrian Research Promotion Agency (FFG) under the grant agreement number: 54497923.

APPENDIX

6. REPRODUCIBILITY CONFIGURATION SETTINGS

7. BASELINE RESULTS

As an initial reference, simple unsupervised anomaly-detection methods were applied to analog ventilation signals of the main hub. The objective of this baseline is not performance optimization, but to provide a lightweight pointwise reference for comparison with the multivariate and sequence-based methods in the main paper. In contrast to the main experiments, which operate on aligned multivariate windows, this baseline evaluates each sensor independently and captures only single-sensor deviations without temporal or cross-variable context.

Only continuous analog AHU signals were considered. Each time series was resampled, smoothed using a 30-minute rolling window, downsampled to 30-minute intervals, and standardized. This results in one scalar observation per sensor and timestamp, and each detector is applied independently per sensor.

Table 5 shows clear differences between the methods. K-Means produces the highest fraction of flagged observations, DBSCAN is highly conservative, and Isolation Forest lies between these extremes. These values should not be interpreted as fault prevalence, but as an indication of how sensitive each method is to deviations from the dominant distribution of a single sensor.

Higher flag rates are observed for humidity-related variables such as ZUL_AF_M, ZUL_FE_M, ABL_AF_M, and ABL_FE_M. These signals are more variable due to environmental conditions and control adjustments. Temperature sensors show moderate flag rates, often associated with operational transitions, while some return-loop signals such as HRG_RL_M appear comparatively stable.

Agreement between methods is limited. Only a small fraction of samples is flagged jointly, while many detections are unique to a single method. This reflects that each method captures a different notion of deviation: DBSCAN identifies isolated low-density observations, K-Means highlights distance from dominant clusters, and Isolation Forest detects statistically rare patterns.

Overall, this baseline provides a simple pointwise reference for single-sensor deviation detection. However, it does not capture multivariate dependencies or temporal structure and is therefore primarily useful as a sanity-check comparison for the more structured methods presented in the main paper.

Table 4. Configuration of classical clustering-based anomaly detection across scenarios. Data-driven parameters vary depending on feature dimensionality and data distribution.

Component	Scenario 1: Analog Sensors	Scenario 2: Grouped Sensors
1. Data filtering		
Input data	Analog signal selection (18 sensors)	System-On operation only organized in sensor groups (98 sensors)
2. Resampling, missing values and normalization		
Window length	3 hours (12 steps @ 15 min)	6 hours (24 steps @ 15 min)
Window stride	1	Same
Missing values	Median imputation (train statistics)	
Feature construction	Window features (mean, std, min, max per sensor)	Same per group
Normalization	Z-score (train mean and std)	Same
3. Train/test split		
Train/test split	48% train / 52% test (time-based)	95% train / 5% test (time-based)
random state	0	42
4. Modeling		
PCA components	Data-driven (95% variance)	Data-driven per group
PCA scoring	SPE + Hotelling T^2	Same
K-Means	$k = 20$	$k = 4$
DBSCAN	$\epsilon = 0.8$, min. samples = 10	Estimated per group from train k-distance (95%)
GMM	$n = 5$, diagonal covariance, max iterations = 500	$n = 4$, full covariance
Isolation Forest	5 estimators, contamination=0.01, random state = 0	300 estimators, contamination=0.01
5. Anomaly detection		
Thresholding	99th percentile of training scores for K-means, DBSCAN an GMM, 95th percentile of training scores for K-means, DBSCAN an GMM	Same
Aggregation	Model agreement (consensus / strong consensus)	Same

Table 5. Summary of detected anomalies across ventilation sensors using simple per-sensor baseline methods (Scenario 1). Sensor identifiers follow the original BMS naming convention, explained in *Description*.

Sensor ID	Description	KMeans (%)	DBSCAN (%)	Isolation Forest (%)	All Three (%)	Two Methods (%)	Unique (%)	Total (%)
ABL_AF_M	exhaust absolute humidity	45.07	0.36	17.17	0.32	8.36	44.92	53.60
ABL_FE_M	exhaust relative humidity	46.49	0.33	17.28	0.25	8.29	46.77	55.31
ABL_FL_W	exhaust filter pressure drop	21.41	0.06	11.19	0.03	7.59	17.37	25.00
ABL_LQ_M	exhaust air quality	48.13	0.83	13.80	0.03	3.76	55.15	58.94
ABL_PR_M	exhaust pressure	25.31	0.07	13.26	0.03	4.47	29.63	34.12
ABL_TE2M	exhaust temperature	27.99	0.24	14.90	0.13	5.95	30.84	36.92
ABL_TE_M	exhaust temperature	40.78	0.48	12.03	0.28	7.29	37.87	45.44
AUL_FL_W	fresh air filter pressure drop	46.32	0.04	21.65	0.04	12.01	43.88	55.93
BEF_PD_W	humidifier pressure drop	43.93	0.10	15.92	0.07	4.93	49.88	54.88
HRG_RL_M	heating register return temperature	10.29	0.21	13.71	0.13	10.23	3.36	13.72
KRG_RL_M	cooling register return temperature	21.09	0.46	15.99	0.36	8.92	18.61	27.89
NHR_RL_M	post-heater return temperature	33.11	0.32	15.82	0.30	10.10	28.15	38.55
ZUL_AF_M	supply absolute humidity	48.68	0.22	16.95	0.16	8.45	48.48	57.08
ZUL_FE_M	supply relative humidity	45.71	0.09	18.51	0.07	7.98	48.13	56.18
ZUL_PR_M	supply pressure	25.34	0.04	14.63	0.02	6.91	26.14	33.06
ZUL_TE1M	supply temperature	24.30	0.03	22.07	0.02	11.30	23.75	35.06
ZUL_TE2M	supply temperature	43.58	0.02	19.34	0.01	3.58	55.75	59.34
ZUL_TE3M	supply temperature	24.44	0.07	21.51	0.03	11.56	22.80	34.39