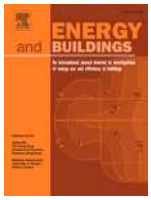
A label-free hybrid fault detection and diagnosis approach for HVAC systems using bayesian networks

Energy & Buildings 351 (2026) 116658

Contents lists available at ScienceDirect

Energy & Buildings

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



A label-free hybrid fault detection and diagnosis approach for HVAC systems using bayesian networks

Marco Paolini, Marco Savino Piscitelli *, Alfonso Capozzoli

Department of Energy (DENERG), TEBE Research Group, BAEDA Lab, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, 10129, Italy

ARTICLE INFO

Keywords:
Fault detection and diagnosis
HVAC systems
Bayesian networks
Hybrid approach
Machine learning

ABSTRACT

The building sector accounts for up to 40 % of global energy demand, with HVAC systems responsible for nearly half of this consumption. Faulty HVAC operation can result in substantial energy waste, reduced equipment lifespan, and increased operational costs. To address these challenges, the scientific community has focused on developing Fault Detection and Diagnosis (FDD) strategies that are both accurate and applicable in real conditions. Although data-driven approaches have shown strong potential, their practical deployment remains limited by the need for labeled data and variables not commonly available in real buildings.

This paper presents a hybrid FDD framework based on Bayesian Networks (BNs) that combines data-driven models with expert knowledge. A set of reference models, using Random Forest algorithms, was developed to predict key variable values and define a baseline. Deviations from this baseline, expressed as residuals, were converted into virtual evidence and combined with hard evidence derived from domain expertise. These inputs were fed into a set of BN models, one for each operational mode, whose parameters and system-level structure were informed by expert knowledge and constructed efficiently through semantic metadata schemas based on brick ontology.

The BNs performed fault detection and component isolation while the diagnosis is supported by targeted statistical analyses tailored to each system component. The main advantage of the proposed framework is that it requires only variables typically available in building management systems and does not rely on the a-priori knowledge of fault labels.

The approach was validated on a simulated dataset from the Single Duct Air Handling Unit developed by the LBNL and further tested on a Fan Coil Unit. It achieved detection and isolation accuracies of approximately 91 % and 87 %, respectively, confirming its robustness, adaptability, and practical relevance.

1. Introduction

Heating, Ventilation, and Air-Conditioning (HVAC) systems account for 50–60 % of building energy demand and 10–20 % of global energy consumption, making them crucial for enhancing energy efficiency in the building sector [1,2]. Existing literature reports that various issues may arise during the long-term operation of HVAC systems, with energy waste due to faulty operations or improper control strategies contributing to approximately 15–30 % of total energy consumption in commercial buildings [3–5]. By promptly detecting, diagnosing, and addressing these inefficiencies, buildings may improve energy performance and achieve potential energy savings of 5–30 % [6]. In this context, advanced Energy Management and Information Systems (EMIS), such as Fault Detection and Diagnosis (FDD) tools, represent a valuable solution to systematically reduce energy waste during the operation of energy systems [7].

In recent decades, HVAC systems have become increasingly equipped with diverse sensors and IoT devices, enabling easier collection and storage of extensive historical operational data [8]. This progress has greatly fostered the adoption of data-driven approaches to tackle FDD-related challenges [9,10]. These approaches employ machine learning, advanced statistical methods, or deep learning to derive models from monitored system data [11]. Compared to conventional knowledge-driven techniques, data-driven FDD offers notable advantages, requiring minimal prior system knowledge while ensuring accurate results at low computational cost. It also provides high analytical flexibility, achieving strong fault detection and diagnosis performance even with limited variables. However, data driven FDD tools face key limitations. They require large datasets of both normal and faulty operations, while fault data remain scarce and mostly rely on public sources. These tools also suffer from extrapolation issues, struggling to identify unseen faults or severity levels outside the training set.

* Corresponding author.
E-mail address: marco.piscitelli@polito.it (M.S. Piscitelli).

https://doi.org/10.1016/j.enbuild.2025.116658
Received 17 July 2025; Received in revised form 16 October 2025; Accepted 28 October 2025
Available online 3 November 2025

Moreover, their transferability across HVAC systems is limited by the lack of standardized metadata schemas, while transfer learning, though promising, still needs validation in real environments [12–14]. Baseline establishment from historical data is also challenging, as persistent faults may be misinterpreted as normal behavior without ground truth validation. Moreover, the lack of physical interpretability can reduce user trust and represent a barrier to adoption. Lastly, their implementation requires advanced knowledge in ML and AI, which many HVAC professionals may not have [15,16]. For these reasons, a promising opportunity is to develop hybrid FDD approaches that combine the advantages of data-driven processes with knowledge-driven ones [17,18]. The latter leverage domain expertise especially in the diagnosis phase, which makes them well suited to replicate diagnostic reasoning across various systems [19,20]. In the literature, among all the algorithms available for performing FDD in HVAC systems, Bayesian Networks (BNs) have proven valuable in integrating expert knowledge within data-driven processes, demonstrating superior capabilities in modeling complex variable dependencies, effectively handling uncertainties, and exploiting prior knowledge of the system configuration [21]. However, knowledge-driven models often require manual design of network structures, which limits their scalability. To this purpose, semantic metadata standards (such as Brick, Project Haystack) provide machine-readable descriptions of building systems, components, and sensor data that can be systematically exploited to support FDD tasks. In this context, this study presents a BN-based FDD framework that integrates data-driven and knowledge-driven approaches for improved reliability and interpretability in HVAC fault analysis. A robust baseline for normal operation is established using Random Forest (RF) models, predicting key system variables such as air temperatures, fan power demand, and control signals. Deviations from predicted values serve as fault symptoms, distinguishing between virtual evidence (deviations from baselines) and hard evidence (deterministic deviations from setpoints or physics-based rules), which update BN probabilities for structured fault detection. An ontology-based approach leveraging the Brick Schema [22] streamlines the definition of the BN structure, mapping HVAC system components, relationships, and sensor data.

The model includes system fault nodes for overall HVAC malfunctions, component nodes for failure sources, and evidence nodes for symptoms. Further refinement integrates domain expertise, optimizing node relationships and incorporating expert-based rules. Differently from traditional BN-based models that define individual fault nodes, the proposed framework groups faults at the component level, reducing complexity while maintaining diagnostic accuracy. To reduce false alarms, fault detection results are evaluated daily rather than at fixed short time intervals. The model accumulates fault predictions every 15 min but applies expert-driven post-processing at the end of each day to refine them, considering information across operational modes. The proposed framework was validated on two case studies [23], to demonstrate its generalizability. The first case study focused on a Single-Duct Air Handling Unit (SDAHU), which regulates supply air temperature using a cooling coil and an economizer. The second case study involved a Fan Coil Unit (FCU) with a four-pipe hydronic system and multi-speed fan control, maintaining room air temperature based on thermostat setpoints. To highlight the novelties introduced in this study, the following section reviews the literature on the use of BNs for FDD in HVAC systems, examining both their advantages and widespread adoption. Additionally, it explores the key challenges that remain unaddressed when approaching FDD as a fully data-driven problem.

2. Related works

In recent decades, BNs have been widely applied in HVAC FDD [24] for their ability to represent cause-effect relationships. They effectively diagnose faults and trace root causes in complex systems where faults propagate and interact. By structuring dependencies probabilistically, BNs improve diagnostic accuracy and enable more targeted maintenance.

Building on this capability, Huang et al. [25] proposed a BN-based framework for root cause isolation using Eigen Entropy Causal Learning, a data-driven approach that infers fault symptom relationships without expert input. By analysing synchronicity in time series, the model constructs a BN for diagnosis and achieves accuracy comparable to expert informed models with a simpler structure.

Another key advantage of BNs in FDD is their ability to integrate multiple diagnostic approaches within a unified probabilistic framework. This capability enhances fault detection by leveraging complementary diagnostic strategies, balancing automation with expert-driven insights, and improving reliability in real-world applications.

This integration was explored in Wang et al. [26], where a Hybrid BN combined data driven and expert based approaches, maintaining diagnostic accuracy even when one component was missing. Tested on two experimental chillers, it outperformed standard methods. Similarly, Wu et al. [27] proposed a fusion driven BN integrating a Broad Learning System for structure generation and fuzzy logic for expert knowledge refinement. Expert input was encoded as evidence nodes, and Noisy MAX was used for parameter learning, achieving high interpretability and performance on a chiller system. Likewise, Behravan et al. [28] employed fuzzy logic to simplify diagnosis, reducing expert effort while preserving accuracy and interpretability.

Considering that HVAC systems operate under diverse conditions influenced by external temperature, system load, and control strategies, maintaining the accuracy of FDD models across these varying operational modes presents a significant challenge. BNs provide a structured way to address this complexity by explicitly modeling dependencies between system components and dynamically adapting to changing conditions.

To enhance system level HVAC fault diagnosis, authors in Verbert et al. [29] proposed a multiple model approach using separate BNs for each operational mode. The models captured component interactions, conservation laws, and virtual sensors, with conditional probabilities derived from labeled simulations, improving accuracy under limited monitoring. Similarly, the authors in Pradhan et al. [30] developed a dynamic BN that incorporates operational modes defined in ASHRAE Guideline 36. By accumulating fault beliefs over time, it achieved accurate diagnosis in a Modelica-based HVAC simulation. Unlike a static BN, where the probability of a fault node depends solely on its corresponding symptom nodes, the Dynamic BN also considers the node state at the previous time step, thereby capturing temporal relationships between faults and symptoms.

Managing uncertainty is also crucial in FDD, as sensor noise, missing data, and ambiguous symptoms can reduce diagnostic accuracy. BNs address this through a probabilistic framework that quantifies uncertainty, enabling more reliable fault classification.

Using this principle, authors in Xiao et al. [31] developed a DBN for FDD of VAV terminals, modeling fault symptom relationships with data from building management systems and manual tests. Two probability-based rules were applied for fault isolation, effectively handling uncertainty and incomplete data and accurately diagnosing ten common faults.

Sensor reliability also contributes to uncertainty, as biased measurements can distort system behavior. To address this, Ng et al. [32] introduced a Bayesian model to detect and evaluate biases in water flow and temperature sensors of a central chiller plant. The method proved effective with both synthetic and real data, identifying biases as small as ±0.5 °C.

By optimizing probabilistic inference and conditional probabilities, BNs reduce false alarms and adapt to varying system conditions.

In [33], a BN classifier for chillers was developed to balance False Alarm Rate (FAR) and Missed Detection Rate (MDR). A Conditional Gaussian Network modeled continuous variables from labeled data, while probabilistic thresholds and expert input minimized errors. The

method reduced FAR from 22.7 % to 4.7 % and improved detection of refrigerant overcharge and leakage. Further, authors in Xu et al. [34] enhanced BN-based FDD by integrating orthogonal defect classification for statistical analysis of labeled air conditioning faults. A four layer BN optimized with a Genetic Algorithm refined conditional probabilities within expert defined ranges, increasing diagnostic accuracy from 81.1 % to 92.8 %.

Finally, A key advantage of BNs is their ability to automate parameter learning, reducing reliance on expert defined rules and assumptions. BNs mitigate this issue by learning network parameters from historical labeled data, allowing for greater flexibility.

In [35] and [36], Distance Rejection and Principal Component Analysis were combined with BNs to enhance FDD performance, validated on chiller data from the ASHRAE RP 1043 dataset. Building on this, Wang et al. [37] introduced a discretization-based BN that autonomously determined network parameters from labeled data, removing the need for expert input or distributional assumptions.

The reviewed studies highlight the wide use of BNs in HVAC FDD. However, most approaches still depend on labeled historical or simulated data to train networks and define probabilistic dependencies. Although domain knowledge can enhance performance, the scarcity of labeled fault data remains a major limitation. Collecting such data is costly and time consuming, despite common assumptions of its availability. To address this, recent work has explored knowledge-driven BN methods that incorporate expert knowledge, physical principles, and heuristic rules, enabling effective diagnosis without exhaustive fault datasets and improving applicability in real systems.

In this context, Chen et al. [38] proposed a two-level discrete BN for diagnosing cross-level faults in HVAC systems. The model identified 11 of 14 faults and ranked the top three causes for the others. BN parameters combined expert knowledge with machine learning on normal operation data, reducing reliance on fault data. The framework integrated a knowledge based structure with Noisy MAX nodes, multi level evidence, and weather and schedule-based pattern matching, emphasizing the value of expert informed reasoning. Similarly, Zhao et al. [39,40] developed a DBN based framework for air handling units. Part I focused on fans, dampers, ducts, filters, and sensors, while Part II extended to coils and secondary water systems. Built on first principles, fault patterns, and expert rules, the framework remained effective also under uncertain or incomplete conditions.

Reinforcing the value of expert-driven FDD, Dey and Dong [41] integrated AHU Performance Assessment Rules (APAR) with BNs to improve fault diagnosis. Validated with real-time data from a campus building, the BN accurately prioritized faults, with manual inspections confirming its reliability. The study demonstrated how predefined expert rules strengthen BN frameworks, ensuring robustness when labeled fault data are limited.

Despite advances in knowledge driven BN methods, many models remain overly complex, with numerous nodes representing fault types, severities, and symptom relations. While this improves diagnostic precision, it limits practical use due to the scarcity of labeled data and the wide variability of faults in real HVAC systems. These approaches often require extensive inputs, such as damper or valve positions, making them data intensive and poorly transferable. Few studies have explored more generalizable BN structures that do not rely on fault labels. One example is reported in Gao et al. [42], where authors moved from system-level to building-level FDD by automatically generating BN structures from semantic descriptions of HVAC systems, enabling scalable root cause diagnosis. This approach is conceptually aligned with the ontology-based BNs, where domain ontologies are used to construct BN structures by mapping concepts to nodes and semantic relations to edges. In [43] it was introduced an approach to derive BN structure and conditional probability tables from ontologies, ensuring consistency with existing knowledge. Other studies extend this by integrating ontologies with BN learning to support causal inference, as shown in additive manufacturing [44], assembly diagnostics [45], and bio-based process modeling [46]. A systematic framework for deriving causal graphs from ontologies was also proposed in Pfaff-Kastner et al. [47], highlighting the role of ontologies as semantic blueprints for explainable, transferable reasoning under uncertainty.

The studies above discussed collectively emphasize the potential of BNs in enhancing different aspects pertaining data-driven FDD in HVAC systems. However, several critical gaps remain in the literature, limiting the effectiveness and reliability of current methods. The following research gaps highlight key areas where advancements are needed:

- **Limited focus on unlabeled faulty data and commonly monitored variables**. Most data-driven FDD methods still rely on extensive labeled fault datasets, which are often scarce and difficult to obtain in real HVAC operations. Furthermore, existing studies rarely focus on models that use only frequently monitored variables, which are the most accessible in practical applications.
- **Insufficient integration of data-driven and expert knowledge-based approaches**. The lack of integration between machine learning techniques and expert knowledge within a unified diagnostic framework limits the interpretability and diagnostic reliability of FDD processes.
- **Inadequate handling of persistent faults and their impact on baseline learning**. Many existing data-driven FDD approaches assume that frequently occurring operational patterns represent normal behavior, which can lead to persistent faults being misclassified as normal conditions. This issue is particularly problematic in self-learning models, where the continued presence of an undetected fault may corrupt the baseline reliability.
- **Challenges in generalizing the automation and structure of FDD processes**. Many data-driven FDD processes require extensive manual effort to define structures, fault dependencies, and symptom relationships, making them difficult to scale and adapt to different HVAC systems.

In this perspective, the following section outlines the primary contributions and the novel elements this research seeks to bring.

3. Novelty and contributions

In response to the challenges identified in the literature, this study presents a BN-based FDD framework that integrates data-driven and knowledge-driven approaches to improve HVAC fault analysis. The proposed study introduces the following key innovations:

- It introduces a novel hybrid FDD methodology relying on unlabeled faulty data and frequently monitored operational variables in real HVAC systems. Few studies in the literature consider both of these aspects.
- It combines data-driven models, such as Random Forest, used for baseline generation, with expert knowledge into the BN-based model to improve the identification of the faulty component in the analysed system.
- It analyses the extent to which an accurate understanding of normal operation conditions affects fault detection and isolation performance, addressing the gap in the literature on quantitatively assessing baseline reliability.
- It leverages semantic metadata descriptions from the building HVAC system to streamline the construction of the BN structure. A system-level BN structure is employed to reduce complexity and improve interpretability, in contrast to more complex structures commonly used in the literature.
- It performs a component-level fault isolation. Differently from traditional DBNs, which typically focus on fault diagnosis by modeling causal relationships between symptoms and specific fault types or severities, the proposed system-level BN adopts a more holistic approach. It explicitly represents HVAC components and their associated variables across the entire system, enabling fault isolation at the component level.

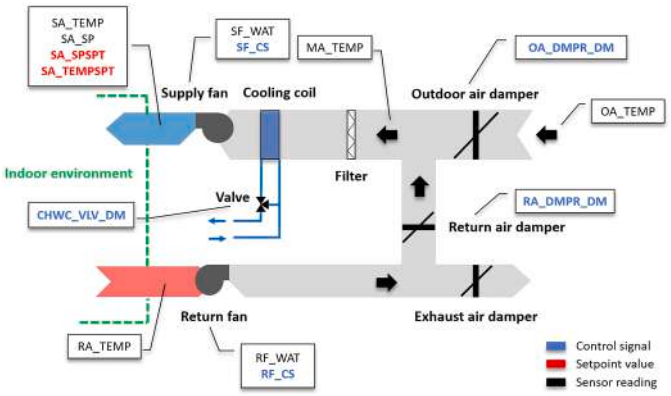


Fig. 1. Single duct air handling unit layout.

- It performs a temporal aggregated fault assessment and expert-driven post-processing. Instead of evaluating faults at fixed short intervals, the model accumulates fault predictions over time and applies expert-driven post-processing on a daily basis. This approach considers operational mode switches and transition phases, helping reduce false alarms.
- It validates the framework generalizability by testing it on multiple case studies, specifically a SDAHU and a FCU. By applying the methodology to HVAC systems with different control strategies and operational complexities, the study demonstrates its adaptability and effectiveness in real-world scenarios.

The rest of the paper is organized as follows: Section 4 provides an overview of the two case studies considered in the analysis. In Sections 5 and 6, the methods and overall methodology employed in this study are detailed. Sections 7 and 8 present the results for each of the two case studies. Finally, Sections 9 and 10 offer a comprehensive discussion of the findings and outline the next steps in this field of research.

4. Case study

The methodological framework described in Section 6 was applied to two different case studies, both developed and provided by LBNL [23].

Case Study 1 involves a Single Duct Air Handling Unit, where the main components include supply and return fans with variable frequency drives, a cooling coil, a cooling coil control valve, and outdoor air, exhaust air, and return air dampers. The layout of the system is illustrated in Fig. 1 while the considered variables are listed in Section 7. The SDAHU is controlled to maintain the Supply Air Temperature (SA_TEMP) set-point when the cooling coil is active. Otherwise, it maintains the Mixed Air Temperature (MA_TEMP) set-point by modulating the economizer dampers. The faults studied in this case are related to a stuck cooling coil valve, a stuck outdoor air damper, and a biased SAT sensor. Each fault category has 4 different severity levels, resulting in 12 different faults.

On the other hand, Case Study 2 focuses on a Fan Coil Unit (Fig. 21), which has a vertical four-pipe hydronic configuration (with both heating and cooling coils), an outdoor air damper, and a fan that operates at three possible speeds: high, medium, and low. The FCU is controlled to maintain the room air temperature according to the thermostat heating and cooling set-points. The faults considered in this case are 14 with various severity levels, resulting in 45 different fault labels. The system layout, the list of considered variables and a Table listing the analysed faults are provided in Appendix A.

5. Materials and methods

This section provides a concise overview of the data analytics methods underlying the proposed hybrid FDD process. Rather than offering an exhaustive description, it highlights the key aspects of the algorithms relevant to this study objectives. Specifically, it presents the fundamental concepts of BNs alongside the theory behind Random Forest, which was employed as the data-driven algorithm for establishing the normal operation baseline.

5.1. Random forest

RF is a widely used machine learning method for classification and regression [48]. It builds multiple decision trees, each trained on a random subset of data and features, enhancing generalization by reducing correlation between trees [49]. In classification, the final output is based on majority voting, while in regression, it is the average prediction [50]. In this study, multiple RF models were trained, each predicting the numerical value of a specific key variable. These models establish the baseline characterizing the normal operation of the system necessary for the FDD process. The RF hyperparameters were set with a maximum tree depth of 6 and 100 trees to ensure reliable predictions.

5.2. Bayesian network

BN is a probabilistic graphical model [51,52]. BN relies on Bayes theorem, which uses conditional probability to measure the likelihood of an event, given that another event has occurred [53]. Specifically, if event A (representing a fault) occurs when event B (representing a symptom) is known or assumed to have occurred, the probability of A given B can be expressed as reported in Barber [54]:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

Here, $P(AB)$ is the joint probability of events A and B. Additionally, if $A_1, A_2, \ldots, A_n$ represent a set of random variables that satisfy the following conditions: (i) $\sum_{i=1}^{n} A_i = S$, where S is a certain event; (ii) they are mutually independent; and (iii) $P(A_i) > 0$ for all i, the marginal probability can be obtained as:

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i) \tag{2}$$

Therefore, Eq. (1) can be rewritten as:

$$P(A_i|B) = \frac{P(AB)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} \tag{3}$$

Bayesian inference aims to identify the cause of an event based on its effects by calculating the posterior probability [55,56]. Bayes' theorem provides a method to calculate the posterior probability $P(A_i|B)$ (left side of Eq. (3)) from the prior probabilities $P(A_i)$ and the conditional probabilities $P(B|A_i)$ (right side of Eq. (3)). Both prior probabilities and conditional probabilities, called *parameters* of the BN, could be assigned based on domain expertise or data-driven approaches [57]. The second

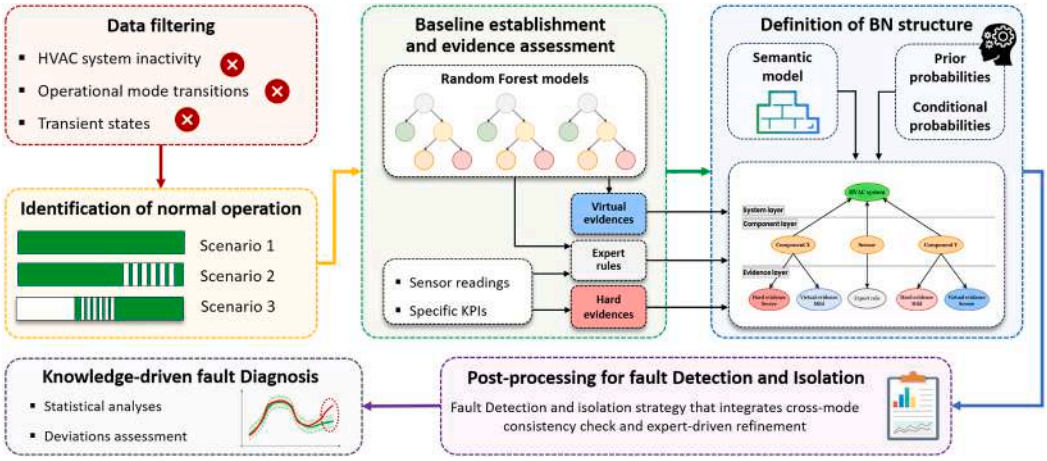


Fig. 2. Methodological framework of the analysis.

important element of a BN is the *structure*, which consists of nodes and arcs that describe events and their causal relationships [58]. Specifically, arcs represent direct probabilistic dependencies between nodes, while nodes represent the variables in the domain [59]. In this study, two different types of nodes were used: *fault nodes*, which indicate the presence or absence of faults in each component, and *evidence nodes*, which represent observable symptoms.

6. Methodology

The methodology implemented in this work aims to define a novel hybrid strategy for the implementation of FDD in HVAC systems. The framework unfolds over multiple steps as shown in Fig. 2.

6.1. Data filtering and identification of normal operation

The datasets for both case studies followed a consistent structure. Each case study comprised multiple files, with each file representing a specific fault at an assigned severity level over a one-year simulation period. A separate file documented normal system operation. In the first case study (SDAHU), the dataset included 12 faulty conditions and one normal operation all of them simulated over one year using the same weather conditions. Similarly, the second case study (FCU) contained a single file for normal operation and 45 faulty conditions. The original timestep of 1 min was aggregated into 15-min intervals using a moving average to reduce noise and computational complexity. Furthermore, time periods referred to system inactivity, operational mode transitions, or transient behavior were excluded to ensure the analysis focused on stable and representative conditions.

To distinguish between transient and steady-state periods, the gradient of the control signals for system components was analysed over time, following the approach proposed by [60]. This filtering was helpful for minimizing inaccuracies in fault detection and diagnosis caused by the high variability inherent in transient phases [61]. The second step aimed to identify data within each available file representing normal system behavior. Although only one file specifically described normal operation, the remaining ones representing faulty conditions do not necessarily indicate improper system performance at all timestamps. For instance, a cooling coil valve stuck at 0 % is a fault during cooling mode but not during economizing, when no heat exchange is expected at the coil level.

Identifying normal operating patterns is essential for data-driven FDD, as reliable baselines depend on labeled normal data. However, as reported in Section 2, such data are often difficult to obtain, requiring dedicated approaches to address this task. Accurate baseline identification improves FDD robustness by reducing false positives and false negatives. In this work, three approaches were applied, reflecting different levels of prior knowledge about HVAC system operation. These scenarios are detailed as follows:

1. **Limited knowledge of the system (Scenario 1):** Data points were labeled as normal exclusively through the use of standard rules, such as the APAR rule set [62], without any prior knowledge of the original label assigned to the normal operation file. This represents the most challenging scenario, where the system control logic and contextual information were entirely unknown.
2. **Comprehensive knowledge of the system and its control logics (Scenario 2):** Within this scenario the normal operation was identified among the available files without any a-priori knowledge of the original label assigned to the normal file. In this case it was exploited a detailed knowledge of the control logic and setpoints governing all components that were available together with the simulated data [23]. Together with the knowledge of the control logics, all the applicable APAR rules on the considered case study were used to label datapoints as normal. Additionally, some expert rules were derived from the control logic to better complement the results of the APAR rule subset.
3. **Ground truth + complete knowledge of system operation (Scenario 3):** In this approach, normal system operation was identified across all available files by leveraging prior knowledge of the explicitly labeled normal operation file, along with complete information on control signals and actual component positions at each timestep. This allowed for the extraction of batches of normal operation data also from simulated datasets initially classified as faulty. Specifically, data points within faulty files were labeled as normal if the actual component positions closely matched their corresponding control signals within a predefined margin of error. This method provided a precise identification of normal operating conditions, with the resulting labeled data serving as the ground truth for subsequent analysis.

The three approaches yielded distinct normal operation datasets, all used to establish baselines. This enabled assessment of how normal data quality affects the reliability of the data-driven FDD tool.

After labeling normal operation data under Scenario 3, each dataset was split by assigning the first week of each simulated month to the test set (Fig. 3). The remaining data were divided into training and validation sets using an 80–20 split.

For Scenarios 1 and 2, the first week of each month was removed to ensure a common test set across all three scenarios (Fig. 3). However, due to differences in how normal operation was defined in each scenario, the resulting training and validation sets contained varying proportions of normal and faulty labeled data.

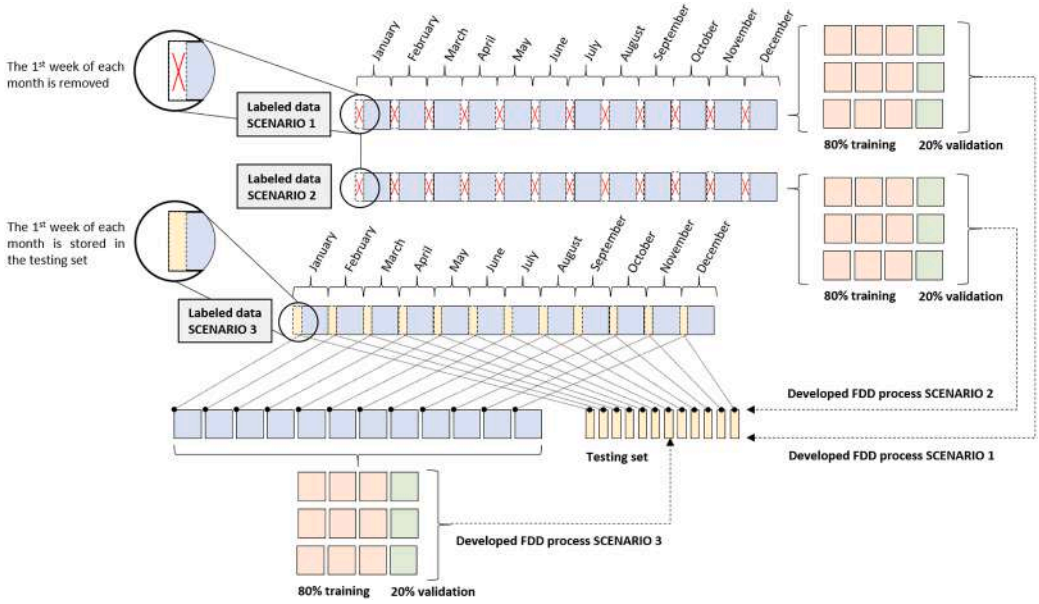


Fig. 3. Selection process of the training, validation and testing sets across the three scenarios.

Moreover In the test set, faulty data were categorized by affected component. For instance, datasets where the outdoor air damper was stuck at 10 %, 25 %, 75 %, or 100 % were grouped under "economizer faults" in the SDAHU. This fault classification enabled the creation of a consistent ground truth for both normal and faulty conditions, supporting systematic fault detection and isolation, which in turn informed the fault diagnosis process.

Importantly, fault labels were used exclusively during testing to evaluate the framework performance and were not employed in training.

6.2. Baseline establishment and evidence assessment

A baseline describing the normal operation of the system was established starting from the labeled normal data included in the training sets (three in total, one for each scenario pertaining the identification of the normal data points). The baseline was established by means of multiple RF models according to the methodology outlined in Piscitelli et al. [63]. Specifically, regression-based RF models were employed to establish a fault-free operational benchmark for the system. Some key operational variables were selected as a target attribute in separate models, while a subset of the remaining system-related variables served as inputs (details were provided in Section 7 for the SDAHU system and in Appendix A for the FCU system). External forcing variables, such as environmental conditions and upstream system parameters, were exclusively used as input features.

This methodology resulted in a comprehensive set of regression models for both case studies ensuring a robust representation of the normal operation.

Additionally, the residuals from the trained Random Forest models were used as virtual evidence in the subsequent BN-based algorithm for fault detection and isolation. As discussed in Peng et al. [64], Mrad et al. [65], virtual evidence represents uncertain conditions and can be derived from the residuals of estimation models. Assuming the residuals follow a Gaussian distribution, deviations between actual and estimated values can be transformed into fault probability values ranging from 0 (normal) to 1 (faulty) [66]. These probabilistic values were then used to update the probability of specific BN nodes, as detailed in the BN architecture.

In addition to virtual evidences, hard evidences and expert rules were also incorporated into the BNs. Differently from virtual evidence, which is based on deviations from estimated values, hard evidence relies on direct observations such as deviations of variables from their setpoints. On the other hand, expert rules introduced multiple conditions that can exploit KPIs, residuals from estimation models and variable deviations from setpoint or expected values at the same time. This combined approach enhanced the robustness of the BN-based fault detection and isolation framework by integrating both data-driven and physics-based reasoning.

6.3. Definition of the Bayesian network structure and parameters

The definition of the BN architecture is ontology-driven, starting from an available semantic metadata model using Brick. The core function of the Brick Schema is to provide a consistent and machine-readable description of building systems [22]. It defines a formal hierarchy of classes representing common elements such as equipment, sensors, and control points, and establishes relationships between them through a set of clearly defined semantic properties (such as *Brick: hasPart, isPartOf, hasPoint*). These relationships express spatial, functional, and logical connections. This semantic schema is usually employed to detail information about the equipments of the HVAC system and the sensors associated with each component [22,67]. Using this information, a preliminary draft of the BN structure can be created, according to the following procedure:

1. From ***Brick Equipment*** to ***System Node:*** The type of the *Equipment* defined in the Brick Schema is mapped to be the system fault node, where fault detection is performed (i.e., the decision node that determine if the observed condition of the entire system is normal or faulty).This node aggregates the posterior probabilities from related component-level fault nodes, which correspond to sub-equipments linked to the system equipment via the *Brick: hasPart* relationship. If the fault probability of the system node exceeds 50 %, the system is classified as operating in faulty conditions.
2. From ***Brick Equipment*** to ***Component Node:*** For each *Equipment* defined in the Brick Schema associated with the one in the system node, a corresponding component fault node is created within the BN. An arc is established from each component fault node to the system fault node, reflecting the *Brick: isPartOf* relationship. In detail, if the system-level node suggests the occurrence of a faulty state of the system, the component-level node with the highest posterior proba-

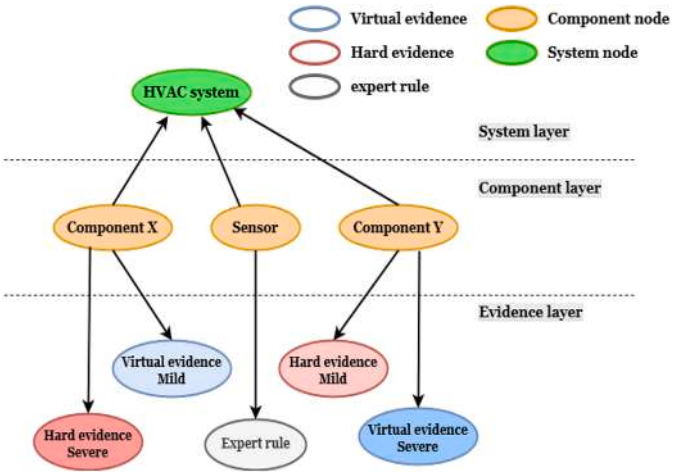


Fig. 4. Example of BN structure.

bility is identified as the primary source of the fault performing the fault isolation.

3. From ***Brick Point*** to ***Evidence Node:*** For each *Point* defined in the Brick Schema, an evidence node is created within the BN. Each component fault node is then linked to its corresponding evidence nodes through arcs, following the *Brick: hasPoints* relationship. These evidence nodes can represent either hard evidence or virtual evidence, capturing the symptoms of the HVAC system under faulty conditions and contributing to fault detection and diagnosis.

The initial draft of BN structure is then refined by the expert through the integration of domain knowledge. This refinement process allows for the elimination of redundant or unrealistic monitored variables, and the addition of extra nodes, which may account for expert rules. Moreover, sensors that measure controlled variables, such as supply air temperature and mixed air temperature sensors, are also treated as fault component nodes. If these sensors have a measurement bias, their effect on system performance can mimic the symptoms of real component failures [68]. By modeling them as fault nodes, the BN can account for sensor-related anomalies, improving the accuracy of fault detection and isolation. An example of a general BN structure that can be determined by following the above introduced approach is reported in Fig. 4. Once the BN structure was established, separate BNs were developed for each operational mode of the HVAC system (e.g., heating, cooling, economizing). This tailored approach ensures that the BN remains both relevant and effective across varying operational scenarios, allowing it to adapt dynamically to changes in system behavior.

Eventually, prior and conditional probabilities were assigned to fault component nodes, the system fault node, and evidence nodes to enable a probabilistic assessment of system operation. To each fault component node, representing individual HVAC components, was assigned an equal prior probability of 0.1 for being faulty and 0.9 for being normal, reflecting an initial assumption that faults are relatively rare. The system fault node, which determines whether the overall HVAC system is in a faulty state, was conditioned on the component fault nodes using a deterministic "OR" rule: if at least one component was faulty, the system was classified as faulty (probability = 1); otherwise, it was classified as normal (probability = 0). Evidence nodes were incorporated into the BN to capture symptoms related to system faults. These nodes were linked to fault component nodes through probabilistic dependencies, representing how observed data influences the likelihood of component failure. The conditional probabilities assigned to evidence nodes varied based on the nature and severity of the observed evidence, as detailed in Table 1.

Table 1 differentiates virtual evidence, derived from regression residuals, and hard evidence, based on rules and physical measurements. Virtual evidence indicating a fault implies a 90 % (mild) or 95 % (severe) fault probability. Hard evidence has stronger weight: mild cases yield a 99 % fault probability, while severe cases determine a state with 100 % fault probability. Conditional probabilities for expert-rule nodes align with those of severe hard evidence.

Table 1
Conditional probabilities for hard and virtual evidence node.

		Fault node	
		Fault	Normal
Virtual evidence - Mild	Fault	0.90	0.10
	Normal	0.10	0.90
Virtual evidence - Severe	Fault	0.95	0.05
	Normal	0.05	0.95
Hard evidence - Mild	Fault	0.99	0.01
	Normal	0.01	0.99
Hard evidence - Severe	Fault	1.00	0.00
	Normal	0.00	1.00

6.4. *Post-processing for fault detection and isolation*

The developed BN continuously analyses system behavior and updates fault predictions at a 15-min time step. However, rather than evaluating the results of the fault detection and isolation at each time step, the analysis is further aggregated and assessed on a daily basis to improve the reliability. Throughout the day, the algorithm accumulates and stores fault probabilities for each component, enabling a more stable and comprehensive assessment over an extended period. At the end of each day, a comprehensive fault detection and isolation analysis is performed, summarizing all detected faults along with their associated probabilities. This post-processing step integrates expert rules to refine the analysis, filtering out transient anomalies that might otherwise lead to false alarms. These rules account for operational mode transitions and hierarchical fault interactions, factors often neglected in standard data-driven FDD approaches. By adopting a daily evaluation strategy that integrates cross-mode consistency and expert-driven refinement, the proposed method ensures that fault detection and isolation remain accurate and actionable, providing users with reliable information while avoiding unnecessary alarms.

6.5. *Knowledge-driven fault diagnosis*

The last step of the methodology focuses on fault diagnosis, complementing the results of the detection and isolation process to determine

the type and severity of the identified malfunction. In the literature, knowledge-driven approaches, such as those based on expert rules [69] or statistical analyses [70], have proven effective in diagnosing faults in HVAC systems and most of all do not require fault labels as required by supervised classification models to perform a diagnosis of a predetermined severity. In this framework, the diagnostic process was conducted at the end of each day, once the faulty components have been identified.

To achieve an accurate diagnosis, the methodology combines statistical analysis with an assessment of deviations from baseline conditions for key operating variables.

By integrating expert-driven validation with statistical analysis, the methodology enables system operators to infer fault diagnoses independently of previous observations of similar fault severity. This approach overcomes a key limitation of purely data-driven fault diagnosis tools, which rely on a-priori knowledge of fault labels and often struggle to diagnose faults that deviate from previously observed patterns.

7. Results: analysis of the single-duct air handling unit

The methodological framework described in Section 6 was tested on the two case studies illustrated in Section 4. The results of the first case study were presented and discussed below. To enhance readability, the results of the second case study were summarized in Section 8, with further details provided in Appendix A.

7.1. Data preparation and identification of normal operation datapoints

For the first case study (i.e., SDAHU), the considered variables were as follows:

- **CHWC_VLV_DM**: Control signal for AHU cooling coil valve; ranges from 0 to 1; 0 – valve should be fully closed, 1 – valve should be fully open.
- **OA_DMPR_DM**: Control signal for AHU outdoor air damper; ranges from 0 to 1; 0 – damper should be fully closed, 1 – damper should be fully open.
- **RA_DMPR_DM**: Control signal for AHU return air damper; ranges from 0 to 1; 0 – damper should be fully closed, 1 – damper should be fully open.
- **OA_TEMP**: AHU outdoor air temperature [°C].
- **RA_TEMP**: AHU return air temperature [°C].
- **MA_TEMP**: AHU mixed air temperature [°C].
- **SA_TEMP**: AHU supply air temperature [°C].
- **SA_SP**: AHU supply air duct static pressure [Pa].
- **SF_WAT**: AHU supply air fan power [W].
- **RF_WAT**: AHU return air fan power [W].
- **SF_CS**: Control signal for AHU supply air fan speed; ranges from 0 to 1; 0 - fan speed is 0 %, 1 - fan speed is 100 %.
- **RF_CS**: Control signal for AHU return air fan speed; ranges from 0 to 1; 0 - fan speed is 0 %, 1 - fan speed is 100 %.
- **SA_SPSPT**: AHU supply air duct static pressure setpoint [Pa].
- **SA_TEMPSPT**: AHU supply air temperature setpoint [°C]. During the economizing mode when the cooling coil is not operated this value is used as a target for the mixed air temperature (MA_TEMPSPT).

As explained previously, three different scenarios were considered for identifying normal operation data in the available datasets. Before this, data were aggregated at 15-min and specific periods pertaining transient, operational mode switch and system inactivity were filtered out from the dataset. For the first case study, among the available 13 files, 167,571 15-min timesteps were considered after the data filtering step. The faults included in the datasets were the following:

- CCV stuck 10 %: Cooling coil valve stuck at 10 %
- CCV stuck 25 %: Cooling coil valve stuck at 25 %
- CCV stuck 75 %: Cooling coil valve stuck at 75 %
- CCV stuck 100 %: Cooling coil valve stuck at 100 %
- OAD stuck 10 %: Outdoor air damper stuck at 10 %
- OAD stuck 25 %: Outdoor air damper stuck at 25 %
- OAD stuck 75 %: Outdoor air damper stuck at 75 %
- OAD stuck 100 %: Outdoor air damper stuck at 100 %
- SAT bias −2 °C: Supply air temperature sensor bias of −2 °C
- SAT bias −4 °C: Supply air temperature sensor bias of −4 °C
- SAT bias 2 °C: Supply air temperature sensor bias of 2 °C
- SAT bias 4 °C: Supply air temperature sensor bias of 4 °C

In the Scenario 3 (i.e., **Ground truth + complete knowledge of system operation**) normal operation data were identified by leveraging the explicitly labeled normal operation file along with complete knowledge of control signals and actual component positions at each time step of operation. The resulting labeled data are visualized in Fig. 5.

In Fig. 5, each row represents a file corresponding to a one-year system simulation, with time reported along the x-axis. The green segments indicate portions of the dataset labeled as normal, while red segments denote faulty conditions. Blank areas represent removed or not available data points, which were excluded from the analysis. It is important to note that this labeling scenario represents the most accurate possible identification of normal operation and was therefore considered the Ground Truth for validating the FDD process. The resulting portion of normal data consisted in 22,064 observations (i.e., about the 13 % of the total).

In the Scenario 2 (i.e., **Comprehensive knowledge of the system and its control logics**) normal operation was identified from the available file without prior knowledge of its label, using information on control logics, setpoints, expert-derived rules and APAR rules. The control logics, pertaining to the first case study, are in the following reported:

- **Occupied Mode (Monday–Friday 6:00 AM–10:00 PM, Saturday 6:00 AM–6:00 PM)**
  - **Fan Status:** The supply and return fans operated continuously during occupied periods.
  - **Supply Air Temperature Control:** The cooling coil valve regulated the supply air temperature to maintain a fixed setpoint of 12.8 °C when the outdoor air temperature exceeded 15.6 °C or dropped below 0 °C. For outdoor air temperatures between 0 °C and 15.6 °C, mechanical cooling was activated only if the economizer dampers were fully open. Otherwise, mechanical cooling was disabled, and the cooling coil valve remained closed. When mechanical cooling was locked out, the supply air temperature was not directly controlled and could deviate from 12.8 °C.
  - **Supply Fan Control:** The supply fan speed was modulated via a variable frequency drive to maintain a constant static pressure setpoint of 398.54 Pa.
  - **Return Fan Control:** The return fan speed was adjusted via a variable frequency drive to operate at 90 % of the supply fan speed.
  - **Minimum Outdoor Air Control:** The outdoor air damper remained at a fixed minimum position (10 % open) when the unit was not in economizer mode.
  - **Economizer Mode:** The AHU entered economizer mode when the outdoor air temperature was between 0 °C and 15.6 °C. In this mode, the outdoor air damper modulated in sequence with the return air damper to maintain a mixed air temperature setpoint of 12.8 °C. Once the outdoor air damper exceeded 100 % open, the cooling coil valve was enabled to maintain the supply air temperature setpoint.
  - **Space Temperature Control:** The zone heating and cooling setpoints were 21.1 °C and 23.9 °C, respectively, during the occupied period.
- **Unoccupied Mode**
  - **Fan Status:** The supply and return fans were turned off, while the cooling coil valve and outdoor air damper remained closed. The system cycled on and off to maintain space temperatures within the unoccupied heating and cooling setpoints.

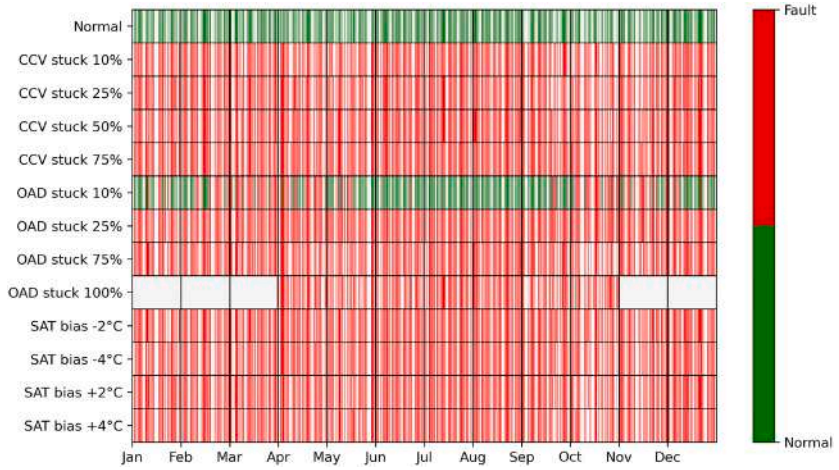


Fig. 5. Identification of Normal operation in Scenario 3 - Ground truth + complete knowledge of system operation.

Fig. 6. Identification of Normal operation in Scenario 2 - Comprehensive knowledge of the system and its control logics.

- **Minimum Outdoor Air Control:** The outdoor air damper remained closed.
- **Space Temperature Control:** The zone heating and cooling setpoints were 15.6 °C and 26.7 °C, respectively, during the unoccupied period.

From the control logics some expert rules were defined to check the existence of deviations from key setpoints, such as the mixing air temperature (only during economizing mode), the zone temperature, the supply air temperature and the supply static pressure. Furthermore also the cooling coil control signal was considered when it reaches a zero value during cooling mode, which indicates potential issues. Eventually for this case study, from the set of 28 APAR rules defined in Schein et al. [62], considering the available components in the analysed SDAHU, it was possible to apply the following rules: 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 25, 26, and 27 (the reader is reminded to Schein et al. [62], Schein [71] for the mathematical formulation of the rules). For all the rules (both APAR and knowledge-based), the same thresholds were employed: $\Delta T_{sf} = \Delta T_{rf} = 0.1\,°C$; $\epsilon_t = 1$; $\epsilon_{cc} = \epsilon_d = 0.1$ [71]. The results obtained for this second scenario are reported in Fig. 6. Fig. 6 depicts a larger portion of data labeled as normal compared to Fig. 5, indicating the presence of false negatives in the dataset intended for establishing baseline models. The resulting portion of normal data consisted in 49,023 observations (i.e., about the 29 % of the total).

In the Scenario 1 (i.e., Limited knowledge of the system), only the above specified APAR rules were applied to identify normal operation data within the files, without leveraging additional knowledge about the AHU's management and control. The results of this scenario are presented in Fig. 7, revealing the highest ratio of false negatives respect to the normal ground truth, resulting in 80,263 records labeled as normal (i.e., about 48 % of the total).

Using Scenario 3 as a reference, detailed knowledge of system behavior clearly improved the identification of normal data compared to relying solely on APAR rules. After labeling, each dataset was split in training, validation and testing set as explained in Fig. 3. Specifically, variations in normal operation identification across scenarios led to different proportions of normal and faulty data in each training/validation set.

7.2. Baseline establishment results

Once normal operation data were identified using the three approaches, baseline models were developed to capture fault-free behavior. Regression-based Random Forest models were used, with selected key variables as targets and a subset of accessible system-related variables as inputs. The selection of the input-output tuples for the set of baseline models was expert-driven using only easily and widely accessible variables. The developed regression models had the following output and input variables:

- **MA_TEMP model**: to estimate the mixed air temperature were used as input variables OA_TEMP, RA_TEMP, OA_DMPR_DM.

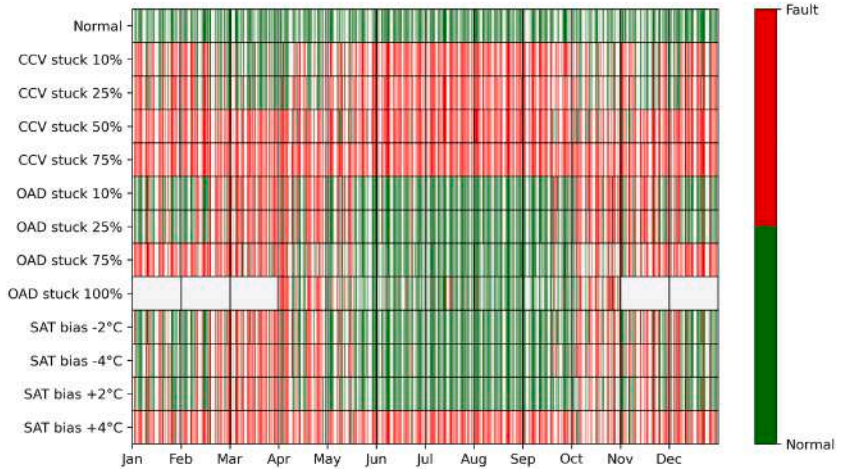


Fig. 7. Identification of Normal operation in Scenario 1 - Limited knowledge of the system.

Table 2
RMSE values of reference baseline models evaluated for the validation set.

Predicted variable	Scenario 1	Scenario 2	Scenario 3
MA_TEMP [°C]	1.31	0.15	0.11
SA_TEMP [°C]	0.14	0.12	0.08
CHWC_VLV_DM [0,1]	0.101	0.093	0.022
OA_DMPR_DM [0,1]	0.020	0.007	0.003
SF_WAT [W]	5.16	3.15	3.07
RF_WAT [W]	2.75	2.69	2.15

- **SA_TEMP model**: to estimate the supply air temperature were used as input variables OA_TEMP, RA_TEMP, MA_TEMP and CHWC_VLV_DM.
- **CHWC_VLV_DM model**: to estimate the control signal for AHU cooling coil valve were used as input variables OA_TEMP, RA_TEMP, MA_TEMP and SA_TEMP.
- **OA_DMPR_DM model**: to estimate the control signal for AHU outdoor air damper were used as input variables OA_TEMP, RA_TEMP and MA_TEMP.
- **SF_WAT model**: to estimate the supply air fan power were used as input variables OA_TEMP, RA_TEMP, MA_TEMP, SA_TEMP, CHWC_VLV_DM, OA_DMPR_DM and SF_CS.
- **RF_WAT model**: to estimate the return air fan power were used as input variables OA_TEMP, RA_TEMP, MA_TEMP, SA_TEMP, CHWC_VLV_DM, OA_DMPR_DM and RA_CS.

For each normal operation dataset, six RF models were trained and validated. Table 2 reports the RMSE values on the validation set. The lowest errors were obtained in Scenario 3, where fault-free data were most accurately identified. Model accuracy decreased progressively in Scenarios 2 and 1 due to less reliable identification of normal operation.

For completeness, Fig. 8 shows a visual comparison between actual and predicted values for each RF model across the three scenarios, using data from a selected normal day in the test set (April 4th), which includes all SDAHU operational modes. The green-shaded area represents three times the RMSE from the baseline models (Table 2), illustrating model accuracy throughout the day. This visualization helps evaluate model accuracy at each timestep during the daily operational period and highlights how the presence of noise (false negatives) in fault-free data influenced pattern learning and the quality of final predictions. In Fig. 8, this effect is especially evident for the supply air temperature model, where Scenarios 1 and 2 show larger deviations between predicted and actual values. In the following step, the baseline models were deployed on the test dataset (that contains out-of-sample data) and the residuals were used as virtual evidence in the subsequent BN-based algorithm for fault detection and isolation.

7.3. Fault detection and isolation results

As described in Section 6, after the definition of the baseline models, the structure of the BN model was preliminary defined. Starting from the Brick Schema, that semantically describe the SDAHU, an initial draft of the BN structure was automatically generated considering the System Layer, Component Layer and Evidence Layer. After that, for each representative operational mode (cooling and economizing mode for the SDAHU) the BN structure was refined in the Component Layer to aggregate, where possible, AHU components together (e.g., outdoor and return dampers in the economizer, supply and return fan in the air duct) and refined in the Evidence Layer to include specific Hard evidences, virtual evidences and expert rules that better fit with the aim of detecting a defective component, or group of components, in a particular operation mode. As a result, two separate BNs were developed, each with a unique configuration of nodes and arcs. These are shown in Fig. 9, representing the mechanical cooling mode with both maximum and minimum outdoor air (i.e., mode 3 and 4), and Fig. 10, corresponding to the economizing mode (i.e., mode 2). From the Figures it is possible to see that the Evidence Layers of the two BNs are different while the Component Layers are exactly the same.

It is worth to note that despite in the analysed use case no faults related to the supply/return fan, filters and mixing air temperature were analysed, the Air Duct and MAT sensor component were included in the Component Layer of the BNs. These components were considered to make the formulation of the detection and isolation problem as much as possible generalizable. In fact while testing the detection capabilities of the developed BNs, fault inference was provided also for these nodes to prove the capability of the model in not generating false positives or faulty pattern misinterpretation involving those components. In particular the BN pertaining the cooling modes of the SDAHU (Fig. 9) used as virtual evidence - mild intensity - the absolute values of the residuals of the SF_WAT model and RF_WAT model, used as virtual evidence - severe intensity - the absolute values of the residuals of the MA_TEMP model and CHWC_VLV_DM model and used as hard evidence - mild intensity- the absolute difference between the air duct static pressure SA_SP and the setpoint SA_SPSPT and the absolute difference between the supply air temperature SA_TEMP and the setpoint SA_TEMPSPT. In addition two expert rules were added as hard evidence with severe intensity for the SAT and MAT node respectively.

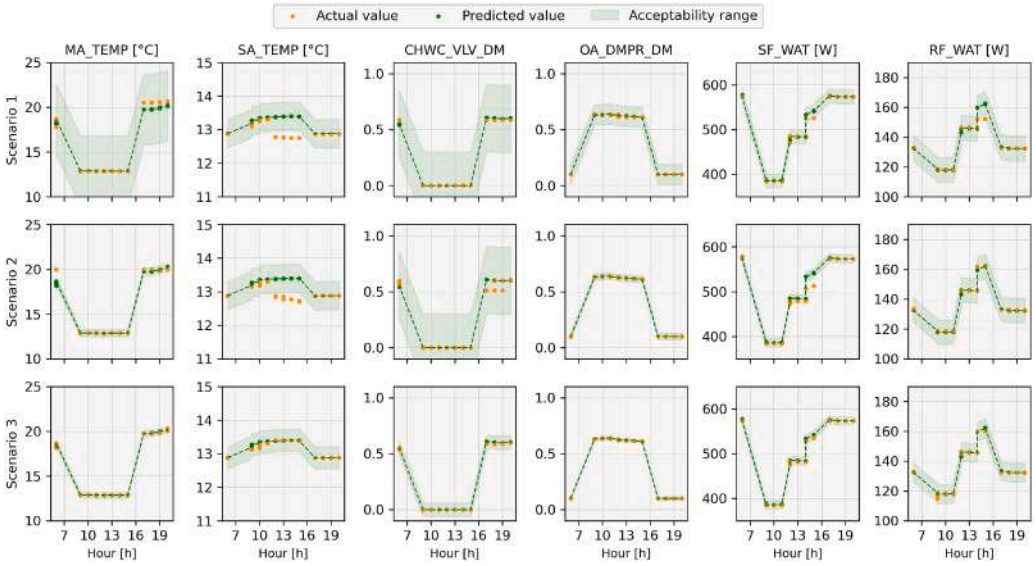


Fig. 8. Random Forest predictions for each variable on normal testing sets (April 4th).

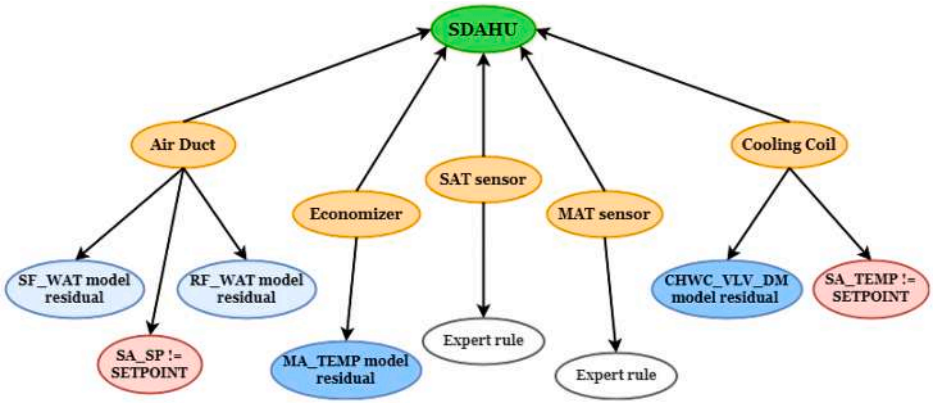


Fig. 9. BN structure for mechanical cooling modes (modes 3 and 4).

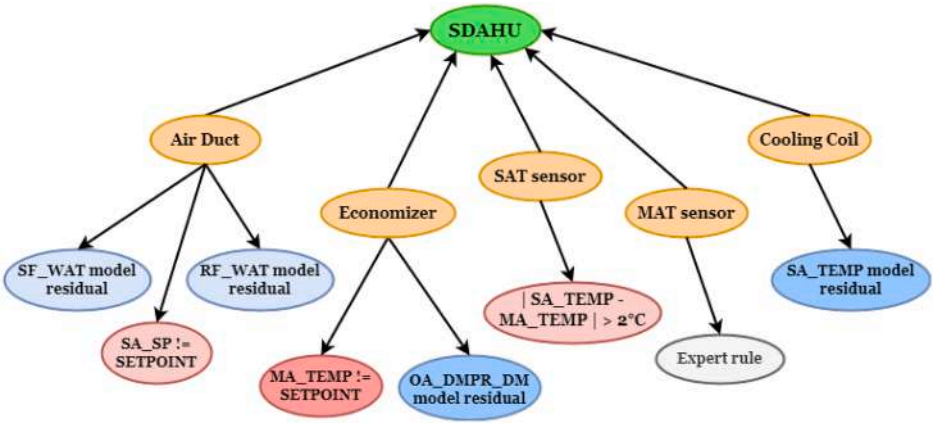


Fig. 10. BN structure for economizing mode (mode 2).

- **Expert rule SAT Sensor**:

$$\left(|\text{SA\_TEMP} - \text{SA\_TEMPSPT}| \leq 1°\text{C}\right)$$
$$\text{and} \quad \left(|\text{CHWC\_VLV\_DM}_{\text{residual}}| > 3 \cdot \text{RMSE}_{\text{validation}}\right)$$
$$\Rightarrow P(\text{FAULT}) = 100\,\%$$

- **Expert rule MAT Sensor**:

$$\left(|\text{MA\_TEMP}_{\text{residual}}| \leq 2°\text{C}\right)$$
$$\text{and} \quad \left(|\text{CHWC\_VLV\_DM}_{\text{residual}}| > 3 \cdot \text{RMSE}_{\text{validation}}\right)$$
$$\Rightarrow P(\text{FAULT}) = 100\,\%$$

On the other hand, the BN pertaining the economizing mode of the SDAHU (Fig. 10) used as virtual evidence - mild intensity - the abso-

lute values of the residuals of the SF_WAT model and RF_WAT model, used as virtual evidence - severe intensity - the absolute values of the residuals of the SA_TEMP model and OA_DMPR_DM model, used as hard evidence - mild intensity - the absolute difference between the air duct static pressure SA_SP and the setpoint SA_SPSPT and used as hard evidence - severe intensity - the absolute difference between the mixed air temperature MA_TEMP and the setpoint MA_TEMPSPT (in the economizing mode the SA_TEMPSPT is used as a target of the mixed air temperature). In addition one expert rule was added as hard evidence with severe intensity for the MAT node respectively.

- **Expert rule MAT Sensor**:

$$\begin{aligned}&\left(|\text{MA\_TEMP} - \text{MA\_TEMPSPT}| \leq 1°\text{C}\right)\\ &\quad \text{and} \quad \left(|\text{OA\_DMPR\_DM}_{\text{residual}}| > 3 \cdot \text{RMSE}_{\text{validation}}\right)\\ &\quad \Rightarrow P(\text{FAULT}) = 100\,\%\end{aligned}$$

To evaluate the performance of the developed BNs in detecting and isolating faults at the component or subsystem level, the test set was used. Fault-free data in the test set kept their label as “Normal”, while faulty data were grouped as described below. In detail, all the faulty data points related to the operation of the cooling coil (i.e., CCV stuck at 10 %, 25 %, 50 % and 75 %) were grouped together with the label ***Cooling coil fault***. All the faulty data points related to the operation of the outdoor air damper (i.e., OAD stuck at 10 %, 25 %, 75 % and 100 %) were grouped together with the label ***Economizer fault***. Eventually, all the faulty data points related to the bias of the supply air temperature sensor were grouped together with the label ***SAT sensor fault***.

These labels served as ground truth to assess the BNs’ ability to detect fault conditions and isolate the most likely affected component. Since the baseline models were trained and validated using three different normal operation scenarios, BN performance was evaluated across all three to examine how noise in the normal dataset impacts detection and isolation results. Additionally, the BNs were free to evaluate fault probabilities even for component nodes that were not affected by the considered faulty conditions. This allowed for evaluating potential misinterpretations or false alarms during the isolation phase.

Once the test dataset was prepared, the two BNs, one for economizing mode and one for cooling modes, provided inference every 15 min. Applying the appropriate BN at the end or during each operational mode period it was possible to identify (if any) the component with the highest fault probability. Fig. 11 shows the 15-min posterior fault probabilities for each component across test days covering all faulty conditions. These probabilities were derived using Scenario 3 (i.e., ground truth + complete system knowledge) for baseline establishment. Each plot highlights periods detected as faulty, (i.e., the fault probability at the system node exceeds 50 %) using red (fault) and green (normal) shading. Additionally, the component with the highest fault probability is shown at each timestep, enabling detailed assessment of the BN’s isolation capabilities. The two BNs operate independently, and information inferred in one operational mode is not transferred to the other. This separation can result in weaker detection or uncertain isolation. For instance, a component may show a high fault probability during cooling mode, while appearing normal in economizing mode on the same day. ChatGPT ha detto:

All days shown in the first column of Fig. 11 correspond to cooling coil faults. In most timesteps, the BNs correctly identified the cooling coil as the most likely faulty component. However, on the day when the coil was stuck at 75 % for several hours, the SAT sensor briefly appeared as the top fault candidate. This occurred only during the economizing mode, while in adjacent cooling mode periods, the cooling coil remained correctly prioritized.

To address this issue and increase the reliability of the results, a post-processing inference process was also proposed and shown in Fig. 12. This method analyses the fault probabilities across different operational modes over a daily timescale aiming to refine fault detection and isolation by sharing information inferred within each mode throughout the day.

Table 3
Fault detection accuracy across the three scenarios.

	Scenario 1	Scenario 2	Scenario 3
Inference at 15-min	68 %	74 %	95 %

The post-processing inference process follows a structured path. If the system-level fault probability exceeds 50 %, the component with the highest fault probability is identified. If the SDAHU operated in only one mode throughout the day and a component fault occurrence exceeds 80 %, that component is flagged as faulty. If the AHU operated in more than one mode and the same component consistently shows the highest occurrence across all modes, it is also marked as faulty (for reference see the plots in the third column of Fig. 11 pertain to the occurrence of SAT sensor faults and for all the timesteps the BNs correctly suggested the SAT sensor as the most probable faulty component).

If neither of these conditions is met, the process checks whether the economizer is estimated to be in faulty condition in mode 2, mode 3 or mode 4. If so, the economizer is flagged. Otherwise, the system checks whether the cooling coil is predicted faulty in mode 3 or 4. If this is the case, it then verifies whether the SAT sensor is predicted faulty in mode 2 to confirm a fault in the cooling coil. If not, the system checks whether the SAT sensor is predicted faulty in mode 3 or 4. At this point, the cooling coil or the MAT sensor is assessed during mode 2 to isolate potential faults in the SAT or MAT sensors.

In order to assess the performance of the fault detection and isolation process (the latter with and without the post-processing inference process) the entire test set, was used considering all the three investigated scenarios for baseline establishment. The test set was configured as follows:

- 5235 datapoints referred to the normal operation of the system
- 34,224 datapoints referred to the faulty operation of the system
  - 12,618 datapoints referred to cooling coil faults (i.e., stuck of the cooling coil valve at 10 %, 25 %, 75 % and 100 %)
  - 8998 datapoints referred to economizer faults (i.e., stuck of the outdoor air damper at 10 %, 25 %, 75 % and 100 %)
  - 12,608 datapoints referred to supply air temperature sensor fault (i.e., sensor bias of −2, +2, −4 and +4 °C)
  - 0 datapoints referred to mixed air temperature sensor fault (in the test the BNs are free to infer fault probability on this component even if there are no faulty data pertaining to that component)
  - 0 datapoints referred to fault of air duct components - fans and filters - (in the test the BNs are free to infer fault probability on this component even if there are no faulty data pertaining to that component)

In Table 3 the fault detection accuracies, defined as the percentage of time the BNs correctly estimated the presence or absence of faults, are reported for the three baseline establishment scenarios. The Table shows that a noisy normal operation dataset significantly affected the detection capabilities of the BNs, with a loss of about 21 % accuracy in Scenario 2 and 27 % in Scenario 1, compared to Scenario 3, which serves as the reference ground truth.

On the other hand, in Table 4 are reported the fault isolation accuracies (percentage of time when the BNs estimated the presence or not of a fault and correctly identified the affected component) obtained in the three baseline establishment scenarios and considering or not the implementation of the post-processing daily inference process.

Firstly, post-processing led to an accuracy improvement in the range of 2 % to 4 %. However, the most notable result concerns the impact that the noisy normal operation datasets had on fault isolation accuracy. In this case, the accuracy drop was even more significant than that observed in fault detection (Table 3), with a decrease of about 22 % in

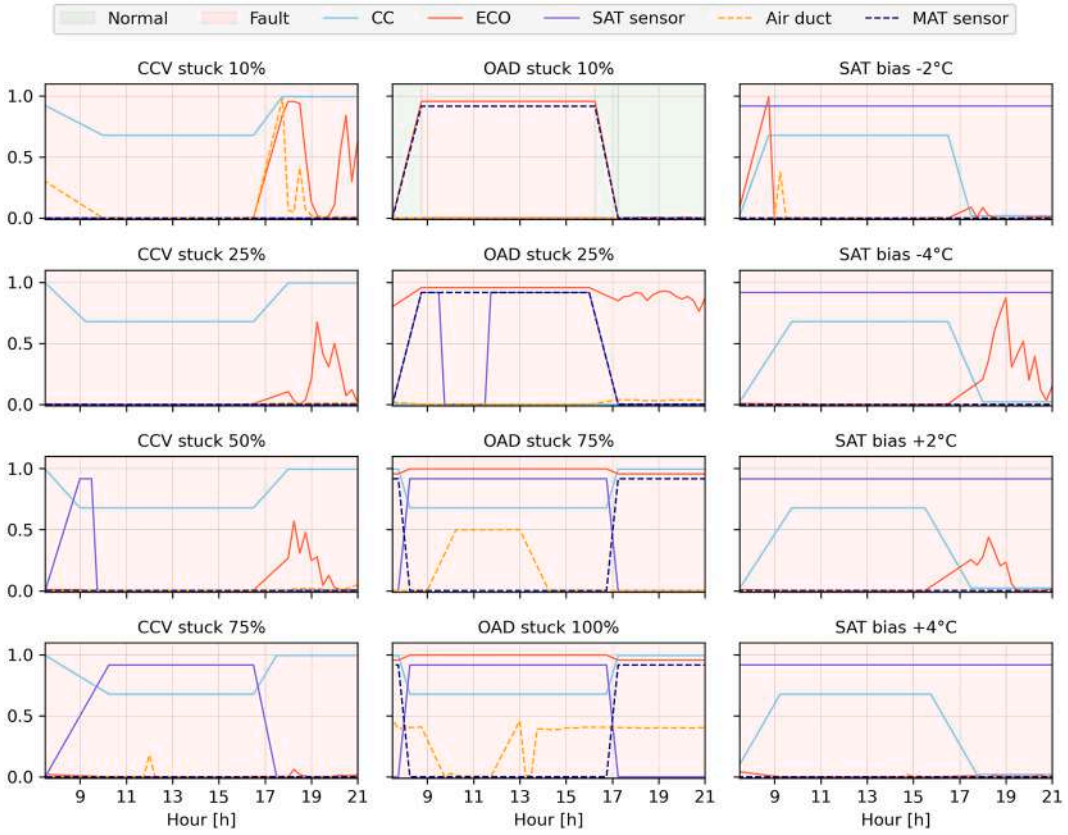


Fig. 11. Posterior probabilities of each components (Scenario 3).

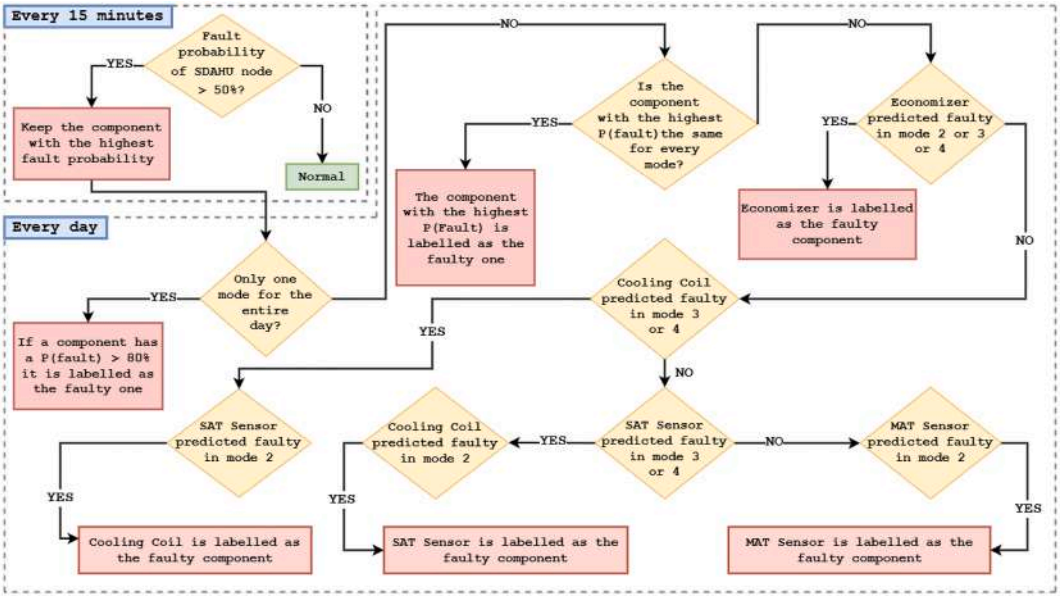


Fig. 12. Post-processing expert rules for fault detection and isolation.

Scenario 2 and 35 % in Scenario 1, compared to Scenario 3. In the latter, the BNs proved to be highly accurate, reaching an overall accuracy of 91 %. For the sake of completeness, in the following are reported the confusion matrices (Tables 5–7) pertaining the fault detection and isolation results obtained through the daily post-processing for each baseline establishment scenario analysed. In the confusion matrices the rows represent the actual classes (Ground Truth) and the columns represent the output of the fault detection and isolation provided by BNs.The Tables also report values for precision, recall and accuracy, further demonstrating the significant impact that the identified reference normal dataset had on the robustness of the baseline regression models and on the overall inference quality of the BNs.

Finally, focusing only on the results of Scenario 3, Fig. 13 presents a temporal visualization of actual fault and normal labels distributed across the test dataset, while Fig. 14 shows the estimations provided by the proposed fault detection and isolation process. The x-axis spans

Table 4
Fault detection and isolation accuracy across the three scenarios and BN inference approach.

	Scenario 1	Scenario 2	Scenario 3
Inference at 15-min	53 %	67 %	87 %
Daily post-processing	56 %	69 %	91 %

Table 5
Confusion matrix of the fault detection and isolation results in Scenario 1 (Limited knowledge of the system).

	CC	ECO	SAT	MAT	Air duct	Normal	Recall
CC	**10078**	4	513	0	2	2021	80 %
ECO	0	**1596**	273	3747	1	3381	18 %
SAT	6	33	**5166**	162	3	7238	41 %
MAT	–	–	–	–	–	–	–
Air Duct	–	–	–	–	–	–	–
Normal	0	29	17	39	1	**5149**	98 %
Precision	100 %	96 %	87 %	–	–	29 %	**Acc. = 56 %**

Table 6
Confusion matrix of the fault detection and isolation results in Scenario 2 (Comprehensive knowledge of the system and its control logics).

	CC	ECO	SAT	MAT	Air duct	Normal	Recall
CC	**10490**	6	564	0	34	1524	83 %
ECO	0	**6764**	9	2	94	2129	75 %
SAT	532	427	**4736**	0	10	6903	38 %
MAT	–	–	–	–	–	–	–
Air Duct	–	–	–	–	–	–	–
Normal	1	75	0	0	3	**5156**	98 %
Precision	95 %	93 %	89 %	–	–	33 %	**Acc. = 69 %**

Table 7
Confusion matrix of the fault detection and isolation results in Scenario 3 (ground truth + complete knowledge of system).

	CC	ECO	SAT	MAT	Air duct	Normal	Recall
CC	**12130**	0	487	0	1	0	96 %
ECO	0	**7274**	145	0	6	1573	81 %
SAT	587	506	**11515**	0	0	0	91 %
MAT	–	–	–	–	–	–	–
Air Duct	–	–	–	–	–	–	–
Normal	19	45	89	0	1	**5081**	97 %
Precision	95 %	93 %	94 %	–	–	76 %	**Acc. = 91 %**

from January to December, while the y-axis lists the considered 13 system conditions across the analysed files, including normal operation and the various faults. Each row represents a distinct normal or faulty condition, with the color indicating the affected component by each fault. The color legend on the right identifies four label categories: green for normal conditions, light blue for cooling coil faults (Fault CC), red for economizer faults (Fault ECO), and purple for SAT sensor faults. Rows contain gaps indicating periods out of the test dataset. Comparing the two figures, estimations of normal conditions appear largely accurate. Cooling coil faults are also captured with good temporal consistency across the different severity levels. Economizer faults show broader differences, specially referred to OAD stuck at 25 %. In that case the model occasionally mislabeled OAD faults as normal.

SAT sensor faults were generally well identified. However, in some cases for negative bias the model occasionally mislabeled SAT sensor faults as Cooling coil or Economizer faults. Notably, Fig. 14 visualizes also estimations for additional fault components (i.e., MAT sensor and Air Duct components) which are not included in the actual labels of Fig. 13. These estimations are sparse and scattered across different faults and time periods, without a clear association to specific fault conditions or consistent temporal patterns.

7.4. Fault diagnosis results

A key strength of the proposed FDD method is the clear separation between fault detection, isolation, and diagnosis. Differently from typical data-driven approaches that treat the problem as multi-class classification with fixed fault labels, this structure allows greater flexibility and interpretability. Instead, the proposed framework offers a graphical representation of the abnormal behavior of the identified faulty component, using statistical analyses or simple KPIs. This supports the user in determining the most likely diagnosis through clear, interpretable visualizations exploiting the results produced by the BNs.

To demonstrate the framework effectiveness, selected days from the Scenario 3 in the test dataset were analysed, enabling an evaluation of system performance under different operational modes and varying external weather conditions.

As a reference, for days in which the BNs identified the cooling coil as the faulty component, the analysis shown in Fig. 15 proved useful for diagnosing the fault and assessing its severity, particularly in relation to operational modes 3 and 4 (mechanical cooling modes). Fig. 15 illustrates the relationship between the air temperature drop across the cooling coil (calculated as the difference between SA_TEMP and MA_TEMP) and the control signal of the cooling coil valve (CCV_VLV_DM).

The green line represents the logistic regression fit for the cooling coil, as reported in [70], and was obtained using the normal datapoints (grey points) identified in the training set under Scenario 3. The shaded green band around this line defines the acceptable range of variation under normal operating conditions, calculated as three times the standard deviation. The green points indicate actual normal data in the test set, classified as such by the BN associated with the cooling operational modes. In contrast, the colored points correspond to fault scenarios in which the cooling coil valve was stuck at various positions: 10 %, 25 %, 50 %, and 75 %. These points had already been detected and isolated by the BN as potential faults affecting the cooling coil. Their clear deviation from the expected trend reveals that the air temperature drop across the coil does not align with the control signal as it should.

Specifically, when the control signal (CCV_VLV_DM) consistently exceeds the upper bound of the normal range, it suggests that the system is attempting to compensate for a valve stuck in a more closed position than required. Conversely, if the control signal remains below the expected range, it indicates a response to a valve stuck in a more open position. This analysis helps the user understand whether the valve is more or less open than it should be, and supports a qualitative assessment of the severity of the fault based on its effect on the supply air temperature (SA_TEMP).

It is important to note that this analysis is meant to complement the BN results by extracting meaningful qualitative patterns for diagnostic support. The statistical model alone is not sufficiently accurate to serve as the sole method for detection, isolation, and diagnosis. In fact, some normal datapoints identified by the BN may fall outside the acceptable range, while some faulty datapoints may appear close to the normal trend.

For a single day characterized by multiple operational modes, it is also possible to leverage the information extracted from the analysis presented in Fig. 16. The figure illustrates the effect of a cooling coil valve stuck condition on the supply air temperature across the two HVAC operational modes observed during the analysed days in the test set.

The top plot focuses on economizing mode (mode 2), when the coil is not expected to operate. It shows the residual between the actual supply air temperature (SA_TEMP) and its baseline model estimate. Green dots indicate normal data, and the green shaded area marks the acceptable residual range ($\pm 3 \times$ RMSE). As the valve becomes increasingly stuck open, the residual shifts negatively, indicating that the system is delivering cooling energy even when it is not required.

The bottom plot refers to mechanical cooling modes (modes 3 and 4), and displays the deviation between the actual SA_TEMP and its set-

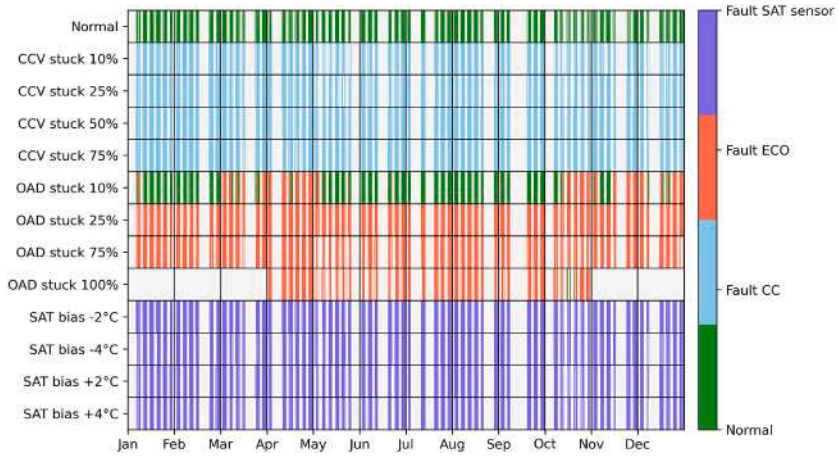


Fig. 13. Actual fault and normal labels (Scenario 3).

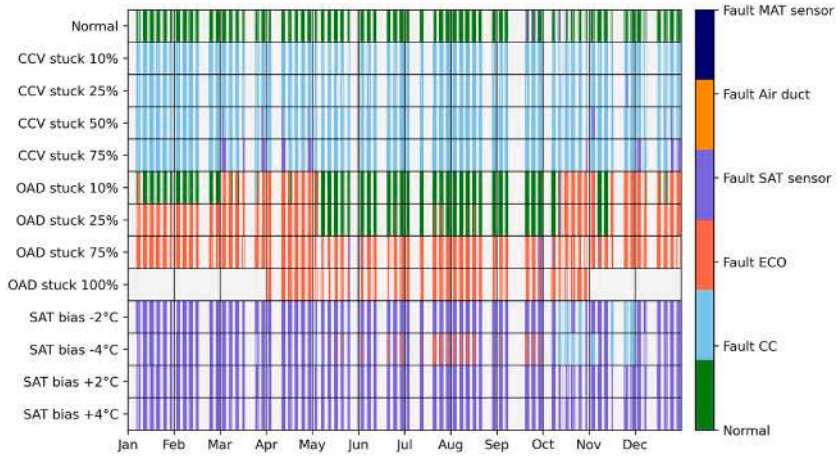


Fig. 14. Predicted fault and normal labels (Scenario 3).

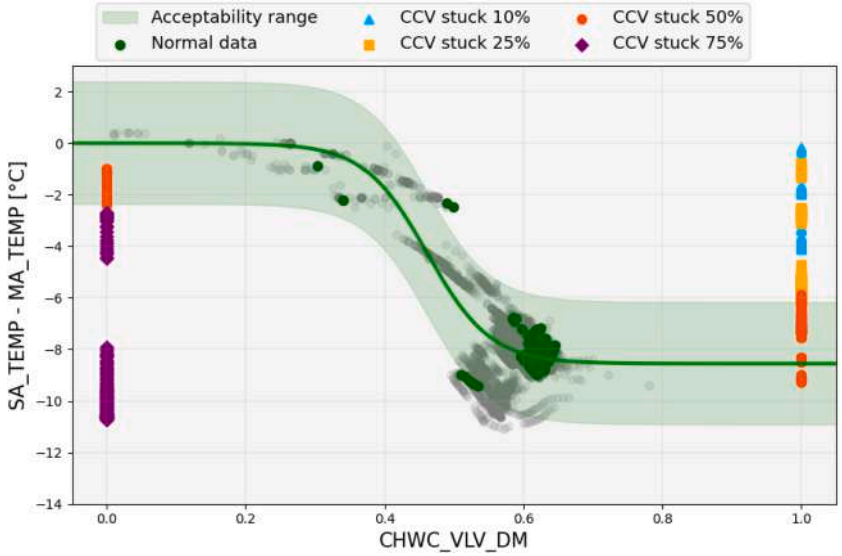


Fig. 15. Statistical analysis describing the relation between the cooling valve control signal and the drop of the air temperature across the coil for qualitatively diagnosing cooling valve faults. The normal ranges and the fitted curves were assessed considering the training set following the Scenario 3.

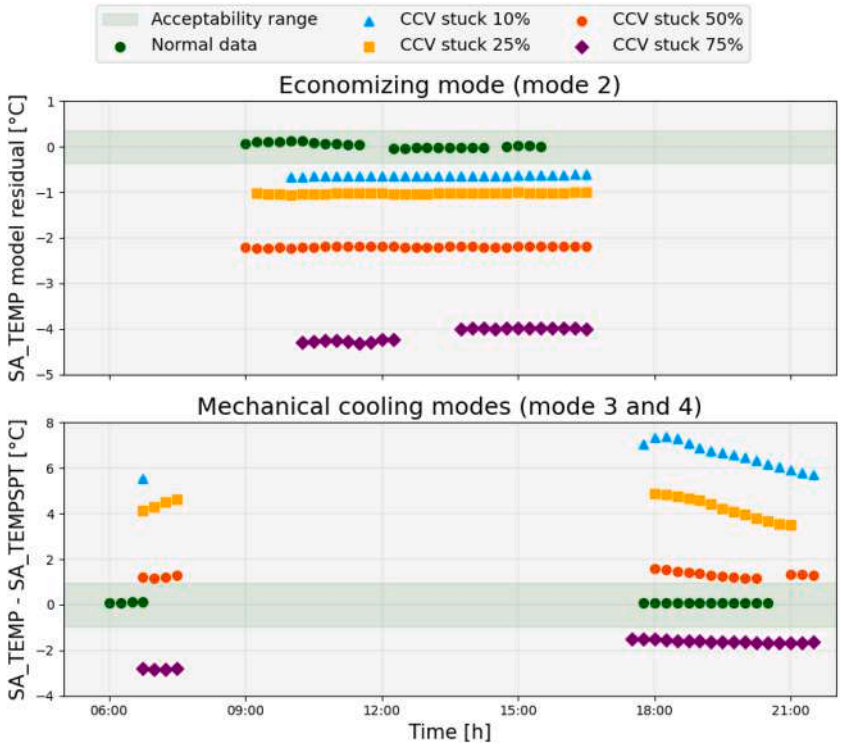


Fig. 16. Daily analysis across multiple operational mode for qualitatively diagnosing cooling coil valve faults (Scenario 3).

point SA_TEMPSPT. Normal data remain within the acceptability band (±1,°C), while faulty conditions result in either negative or positive deviations. This indicates that when the coil is operated, the system may overcool or undercool the airflow due to an unexpected heat exchange at the coil level.

Performing this analysis after fault isolation ensures that faults in other components have already been considered and ruled out. For example, the behavior in economizing mode observed in Fig. 16 could also result from a biased SA_TEMP sensor, as both cases cause systematic deviations. However, since the BNs did not flag the sensor as faulty, are more plausibly attributed to an issue at the coil level, such as a stuck valve.

Similarly to the analysis conducted for the cooling coil, for days in which the BNs identified the economizer as the faulty component, the analysis presented in Fig. 15 can support fault diagnosis and help assess the severity of the fault across all operational modes. Fig. 17 shows the relationship between the Outdoor Air Fraction (OAF) and the outdoor air temperature (OA_TEMP), where the OAF was evaluated as follows [70]:

$$OAF = \frac{MA\_TEMP - RA\_TEMP}{OA\_TEMP - RA\_TEMP} \quad (4)$$

The green line represents the fit of the outdoor air fraction obtained from the normal datapoints (grey points) identified in the training set according to Scenario 3. The green shaded band around this line defines the acceptable range of variation under normal operating conditions. The green points correspond to actual normal data in the test set, detected as normal operation by the BNs. In contrast, the colored points indicate fault scenarios in which the outdoor air damper is stuck at specific positions (10 %, 25 %, 75 %, 100 %), leading to a constant outdoor air fraction regardless of changing outdoor conditions or damper control signals.

When the actual OAF remains consistently above the normal range, it suggests that the damper is stuck in a more open position than required. Conversely, values that fall consistently below the expected range point to a damper stuck in a more closed position. This allows the user to determine whether the economizer is introducing more or less outdoor air than intended and to qualitatively assess the severity of the fault.

For a single day characterized by the occurrence of multiple operational modes, it is also possible to combine this information with the insights derived from the analysis shown in Fig. 18, enhancing the accuracy and interpretability of the diagnosis.

The figure illustrates the impact of a stuck outdoor air damper on the MA_TEMP across the two SDAHU operational modes observed during the analysed days in the test set. The top plot focuses on the economizing mode (mode 2), during which the damper is expected to operate in order to maintain MA_TEMP equal to MA_TEMPSTP, which in this specific mode is set equal to SA_TEMPSPT. Green dots indicate data points classified as normal, while the green shaded area denotes the corresponding acceptability range. As the damper becomes increasingly stuck in the open position, MA_TEMP consistently falls below MA_TEMPSTP, indicating a deviation from expected behavior.

The bottom plot shows mechanical cooling modes (modes 3 and 4), and it visualizes the difference between the actual MA_TEMP and its estimation provided by the baseline model. During cooling modes, MA_TEMP is not regulated according to a setpoint as it is in the economizing mode. Therefore, the estimation from the baseline model was considered as a reference. Since the baseline model uses the control signal OA_DMPR_DM as an input variable, it is possible to distinguish a damper stuck in a more closed position from one stuck in a more open position by evaluating whether the model residual is negative or positive, depending on the weather conditions. In the plot, the blue dotted line represents the outdoor air temperature, and the black dashed line indicates the lower limit of the economizing region.

As a reference, between 8:00 and 17:00, the outdoor air temperature was within a range suitable for economizing. However, a damper stuck in a more closed position resulted in a lower-than-expected outdoor air fraction, causing a higher MA_TEMP than predicted determining the activation of the cooling coil to compensate. Even in this case, the analysis became useful for diagnosis purposes, as the BN had already excluded, for the days under analysis, the presence of a fault at the MA_TEMP sen-

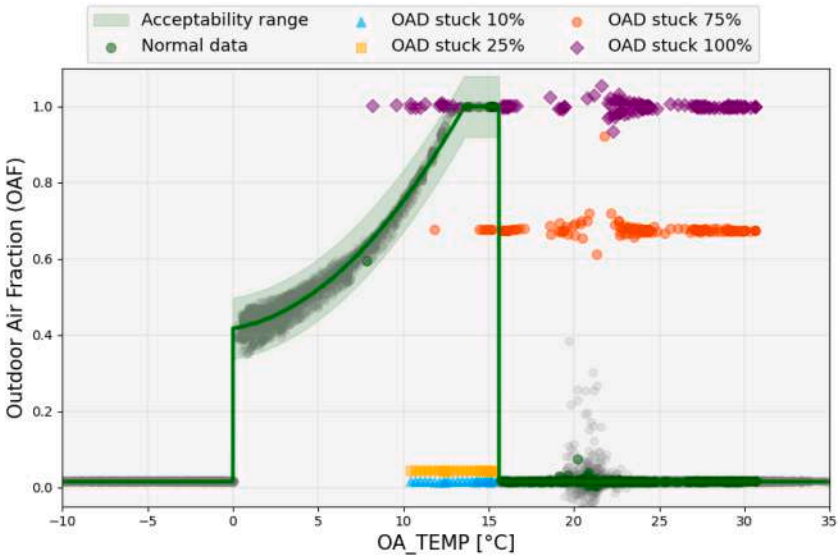


Fig. 17. Statistical analysis describing the relation between the outdoor air temperature and the outdoor air fraction across multiple operational mode for qualitatively diagnosing outdoor air damper faults. The normal ranges and the fitted curves were assessed considering the training set following the Scenario 3.

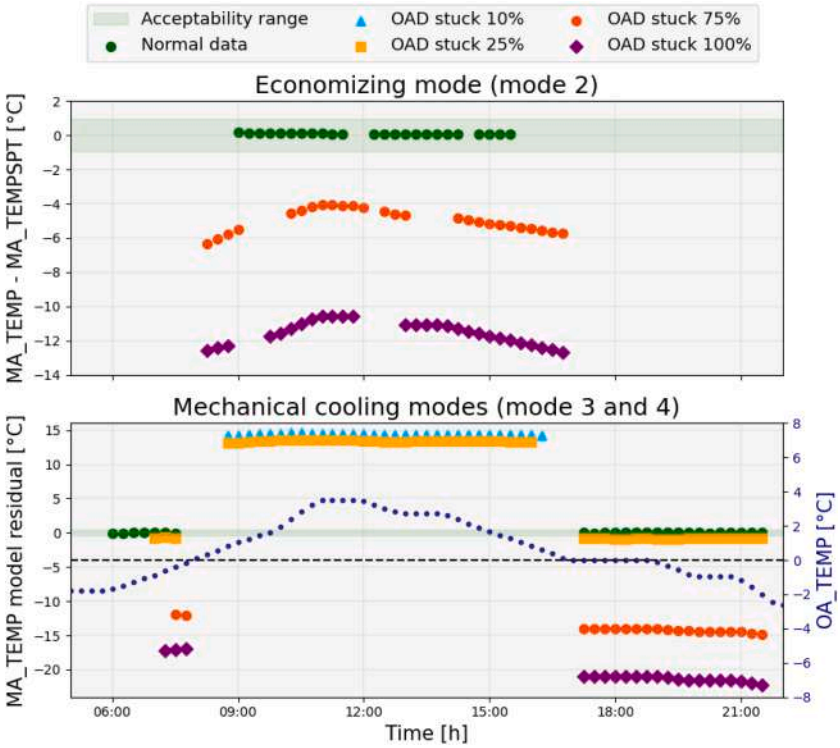


Fig. 18. Daily analysis across multiple operational mode for qualitatively diagnosing outdoor air damper faults (Scenario 3).

sor level, which would have caused a similar effect to the one observed in the Fig. 18.

Finally, for days in which the BNs detected the SA_TEMP sensor as the faulty component, the analysis shown in Fig. 19 can support fault diagnosis and help assess the severity of the fault across all operational modes.

The top plot focuses on the economizing mode (mode 2), when the damper is expected to operate in order to achieve a MA_TEMP equal to MA_TEMPSTP, which in this mode is set equal to SA_TEMPSPT. Given that the outdoor air damper, the MA_TEMP sensor, and the cooling coil were not identified as faulty, the most probable cause of persistent positive or negative differences between SA_TEMP and MA_TEMP is a bias in the SA_TEMP sensor.

During mechanical cooling modes (modes 3 and 4), the bottom plot shows the difference between the actual CHWC_VLV_DM and its estimation provided by the baseline model. The model used as input variables: OA_TEMP, RA_TEMP, MA_TEMP, and SA_TEMP. In this configuration, if the BNs have already excluded faults at the economizer, MA_TEMP sensor, and cooling coil levels, it becomes possible to infer the direction of the SA_TEMP sensor bias. A consistently higher control signal than estimated suggests a positive bias, while a lower control signal indicates a negative bias.

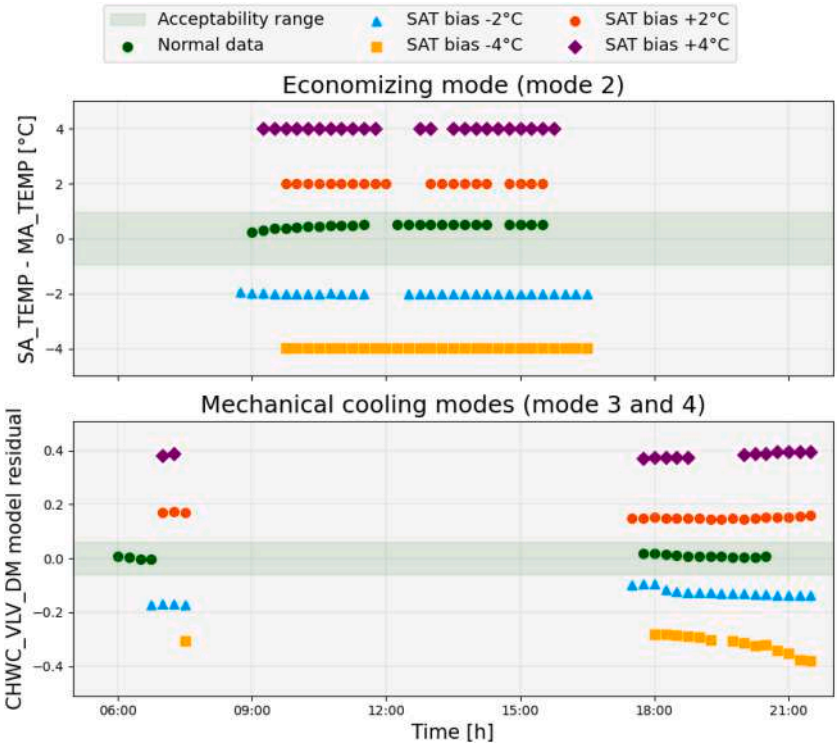


Fig. 19. Daily analysis across multiple operational mode for qualitatively diagnosing supply air temperature sensor fault (Scenario 3).

8. Results: analysis of the fan coil unit

The proposed FDD framework was also tested on a Fan Coil Unit [23] to demonstrate replicability and generalizability behind the developed process. For the sake of conciseness, in this section were reported only the most relevant findings. Further information pertaining the case study and the obtained results is reported in the Appendix A. Also in this case the process started by removing transient periods from the dataset that consisted in 46 files of which 1 represents the fault free operation. Faulty conditions are reported in Table 9. The following step was identify the normal operation data. For this case study the Scenario 3 (i.e., ground truth + complete knowledge of system operation) was considered for the results reported in Appendix A. Specifically, Fig. 20 illustrates the identification of normal operation for each dataset.

In this case study, six distinct operational modes for the FCU were identified [23], leading to the definition of six different BN structures:

- Cooling - occupied hours
- Heating - occupied hours
- Ventilation - occupied hours
- Cooling - setback mode
- Heating - setback mode
- Ventilation - setback mode

Once the normal operation data have been identified the next step was to establish baseline models that capture all relevant patterns characterizing fault-free conditions. Specifically, 8 regression-based RF models were employed to establish a fault-free operational benchmark for the system. The developed regression models had the output and input variables as reported in Appendix A . The structure of the BNs was diversified for considering the different operational patterns among cooling, heating and ventilation modes. The only difference between the BNs pertaining the occupied hours and setback mode consist in the component node of the economizer. For this reason all the BNs were reported in Appendix A with specification of the rules or KPIs included in the expert based evidence nodes. For what is concerned virtual evidences, hard evidences and expert rules for this case study were used the same conditional probabilities as previously defined in Table 1. In order to assess the performance of the developed BNs in detecting and isolating faults at component/group level, the available normal and fault data in the test set were used. Specifically, fault-free data kept the label “Normal”, while the faulty data were grouped under the following categories: air duct faults, cooling coil faults, heating coil faults, economizer faults, water pipe faults, and room sensor faults (Table 9). As for the previous case study, the BNs provided inference on fault probabilities every 15-mins. However in order to increase the reliability of the results, a post-processing inference process was also proposed to refine the fault detection and isolation produced by the BNs in all the periods during the day with the same operational mode by sharing information inferred across them (as explained in Section 7). In particular, if the system-level fault probability exceeds 50 %, the component with the highest fault probability is identified. If the FCU operated in only one mode throughout the day and a component fault occurrence exceeds 80 %, that component is flagged as faulty. If the FCU operated in more than one mode and the same component consistently shows the highest occurrence across all modes, it is also marked as faulty . However, if different modes have different dominant labels, then the final label for the day is determined using a predefined priority: Economizer, followed by Air Duct system, then Piping system, Heating/Cooling coils, and finally Room sensor. In the following is reported the confusion matrix pertaining the fault detection and isolation results obtained through the daily post-processing considering the Scenario 3 for the normal data identification. In the confusion matrix the rows represent the actual classes (ground truth) and the columns represent the output of the fault detection and isolation provided by BNs.The Table also reports values for precision, recall and accuracy.

Table 8 shows that the proposed framework achieved an overall accuracy of 87.2 % in detecting and isolating faults, confirming its robustness and generalizability. Misclassifications mainly involved piping-related faults (e.g., coil fouling), likely due to their limited impact on FCU operation. Such faults could have a relatively small and gradual

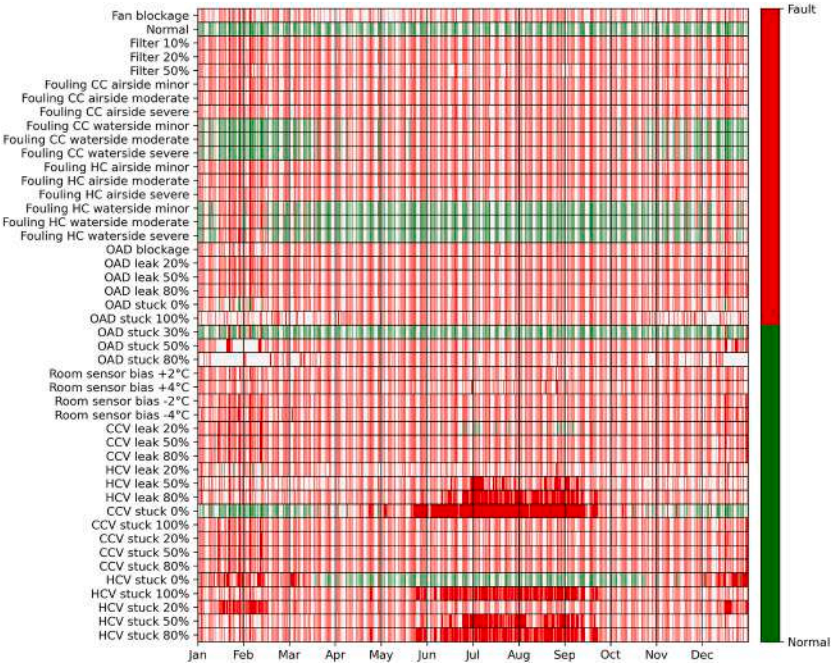


Fig. 20. Identification of Normal operation for FCU (Scenario 3).

Table 8
Confusion matrix of the fault detection and isolation results for FCU in Scenario 3 (ground truth + complete knowledge of system).

	Air duct	CC	HC	ECO	Piping	Room sensor	Normal	Recall
Air duct	**15318**	28	2	186	72	0	4355	77 %
CC	11	**18505**	0	144	131	0	83	98 %
HC	0	0	**18369**	0	689	0	3	96 %
ECO	291	163	323	**14536**	0	0	409	92 %
Piping	33	0	93	14	**341**	0	5028	1 %
Room sensor	36	0	0	191	0	**7475**	381	92 %
Normal	36	27	11	102	204	0	**14258**	97 %
Precision	97 %	99 %	98 %	96 %	24 %	100 %	58 %	**Acc. = 87 %**

impact on FCU behavior under typical control strategies. In practice, coil fouling increases thermal resistance, but the control system often compensates by modulating valve positions, making the fault less evident from commonly available measurements such as room temperature, flow rate, or control signals. To support diagnosis in such cases, statistical analyses could focus on tracking long-term deviations in heat transfer effectiveness. For instance, comparing expected and actual temperature differences across the coil (as shown in Fig. 15) given known flow rates and control signals over multiple days can reveal subtle degradation patterns. In contrast, average precision and recall for other classes exceeded 90 %. A full comparison of actual and predicted labels is provided in Appendix A (Figs. 25 and 26).

9. Discussion

The key findings of this study are outlined below to highlight the strengths of the proposed methodology, as well as the current challenges and areas of ongoing effort in applying hybrid FDD approaches to building energy systems.

9.1. Can FDD succeed without labels and with imperfect knowledge of normal operation?

Labeled fault data are often difficult to obtain in real-world building HVAC systems, and even labeled data representing normal operation are frequently limited. Yet, accurately characterizing normal system behavior is critical for effective fault detection and diagnosis. To address these limitations, this study introduced an FDD framework designed to operate independently of labeled data. Furthermore, the framework enables the evaluation of how the quality of normal operation data influences the overall fault detection and isolation process. The results demonstrated that the proposed approach can achieve approximately 90 % accuracy in detecting and isolating faulty components in SDAHU and FCU systems -without relying on fault labels. However, a clear decline in performance was observed as the quality of normal operation data decreases. These findings emphasize the importance of a reliable representation of the system normal behavior and suggest that, in its absence, supervised data-driven methods tend to under-perform in fault detection tasks.

9.2. Can we detect and diagnose faults without monitoring everything?

Although the monitoring infrastructures in building HVAC systems have increased in recent years, many crucial operational variables remain unmonitored in practice. For example, the actual positions of valves and dampers are often not recorded. If such data were available, fault detection and diagnosis could be performed using straightforward and effective if-then rules (e.g., if the component is in a position different from the one imposed by the control signal then a faulty condition is occurring). However, in most of the cases, practitioners must often rely on indirect indicators - such as control signals or measured air or water

flow rates - which offer only a partial view of system behavior. In this sense, this study argues that an effective FDD solution for real-world deployment should prioritize the use of commonly available variables. In this way, it becomes possible to minimize sensor requirements while still enabling meaningful diagnostics. While this approach may not reach the performance of methods relying on fully monitored systems, it proves that reliable fault detection and diagnosis are achievable using a limited but accessible set of inputs.

9.3. Can data-driven and expert knowledge approach work together for FDD?

The proposed hybrid FDD approach integrates the strengths of data-driven models with the interpretability and robustness offered by knowledge-based methods. In this framework, RF models were employed to establish baseline behavior and generate residuals, which were then transformed into virtual evidence to guide fault isolation within a BN. This process relies solely on identified normal operation data, eliminating the need for labeled fault instances. In parallel, hard evidence was derived from expert-defined rules, key performance indicators, or significant deviations from sensor readings and setpoint values. These expert-driven inputs help overcome the limitations of purely data-driven methods, especially in handling varying control strategies and operating modes. By incorporating system understanding into the BN structure, the hybrid model was capable of embedding case-specific knowledge.

9.4. How does decoupling diagnosis from detection and isolation improve FDD?

Unlike traditional BNs, which are often overly detailed and complex, the networks developed in this study were intentionally kept simple and interpretable. This was achieved by modeling at the system level rather than representing all possible fault classes, enhancing transparency, computational efficiency, and real-world applicability. Crucially, the proposed approach did not rely on predefined fault categories. This was a deliberate design choice: in real applications, a component may exhibit behavior (such as a “stuck” fault) that differs from historical patterns or known fault types. In such cases, forcing a predefined label could lead to misclassifications or missed detections.

Instead, the strategy focused on accurately identifying the faulty component, with the specific diagnosis addressed in a separate analysis stage. This decoupling of detection and isolation from diagnosis reduces dependence on labeled data and enhances flexibility. Once a component was isolated, a targeted statistical analysis was applied retrospectively over a one day window to infer the fault type. This approach provided clearer insight into anomalies, particularly under varying operating conditions.

The hybrid model, consisting of multiple BNs tailored to different operational scenarios, continuously monitored the system, while the diagnostic layer incorporated temporal trends and hierarchical relationships among components. Graphical outputs were used to present the results, providing users with an intuitive and informative view of the most probable root causes. This approach preserved flexibility in the diagnosis process, allowing practitioners to identify new or unexpected fault conditions without being limited by rigid predefined fault categories.

9.5. Is semantic modeling the missing link for scalable FDD processes?

In this study, system-level BN structures were built using semantic metadata modeling. The Brick Schema representation of the HVAC system enabled the automatic generation of an initial BN structure, which was then refined with expert input. A key advantage of using semantic metadata is its ability to systematically identify components and associated sensors, supporting the definition of meaningful relationships within the BN and ensuring alignment with the physical system layout. This automation streamlines the modeling process and improves consistency. However, expert input remained essential, as the semantic layer does not fully capture control logic, operational sequences, and context-specific behaviors. These elements are critical for system performance and FDD accuracy, and were manually integrated to ensure a robust and context-aware BN model.

9.6. Can simulated data truly reflect real-world system behavior?

The ultimate goal of any EMIS is successful deployment in real-world settings. Before field implementation, however, it is essential to test and validate FDD tools under controlled, static conditions using available datasets. A key distinction must be made between data from real systems and that generated through simulations or experiments. Although simulated environments provide controlled settings for initial testing, real-world conditions often introduce significant variability and noise that can compromise model performance. Consequently, models that perform well in simulation may not retain the same level of accuracy when deployed in actual systems, even when the components and control logic are nominally similar. In this study, both case studies were based on simulated data, which, while useful for demonstrating the potential of the proposed approach, limits the assessment of its real-world robustness. Further validation on operational datasets is therefore essential to fully evaluate the reliability, adaptability, and practical applicability of the proposed FDD strategy.

9.7. Can FDD models be portable, transferable, and still interpretable?

The FDD framework proposed in this study was applied to two substantially different case studies, each with its own set of components, control strategies, and sensor configurations, to evaluate its generalizability. The results were encouraging, demonstrating that the approach was both scalable and adaptable across diverse building systems. This adaptability was further supported by the use of semantic metadata schemas, which streamlined the configuration of analytics tools and promoted consistency in system representation.

However, while portability was partially addressed, the transferability of extracted knowledge across different systems remains an open question. Future research could explore this objective through the development and validation of transfer learning techniques aimed at reusing trained models or insights across different HVAC configurations or operational contexts.

Interpretability was another key focus of the proposed methodology. It was achieved by integrating expert knowledge and decoupling the detection of the faulty component from the subsequent diagnostic phase. This separation allows for greater flexibility and clarity in fault analysis. Looking ahead, the application of Large Language Models (LLMs) offers strong potential to enhance interpretability by translating statistical outputs and diagnostic analyses into natural language explanations, making the results more accessible and actionable for non-expert end users.

10. Conclusions

This paper introduced a hybrid FDD methodology for building HVAC systems that addresses two key challenges: conducting fault detection and diagnosis using a limited set of commonly measured operational variables, and reducing dependence on labeled faulty data.

The proposed approach integrated data-driven modeling (Random Forest) with expert knowledge through BNs. The Random Forest models generated predictive baselines of normal operation, and deviations from these baselines were incorporated as virtual evidence into the BNs. This combination enhanced fault isolation by linking statistical anomalies with domain-specific knowledge, improving both accuracy and interpretability.

To examine the impact of baseline quality on diagnostic performance, the study compared three scenarios reflecting increasing levels of system understanding. The results showed that more accurate representations of normal operation lead to better fault detection and isolation, addressing a gap in the literature on evaluating the significance of baseline reliability.

A semantic metadata schema was employed to describe the HVAC system and support the automatic generation of a system-level BN, providing a simpler and more scalable alternative to the complex Diagnostic Bayesian Networks often used in literature. Fault isolation was carried out at the component level thereby avoiding over-specification of the BN and reducing sensitivity to variations in fault type and severity. This design preserves diagnostic precision while enhancing a better framework applicability.

Additionally, the framework introduces temporally aggregated inference and expert-guided post-processing, accumulating fault probabilities throughout the day. This allows the system to account for operational mode transitions and component interdependencies, reducing false positives and enhancing reliability in dynamic conditions. Validation was carried out on two HVAC system types (SDAHU and FCU) with different layouts and control strategies. The framework achieved fault detection and isolation accuracies of approximately 91 % and 87 %, respectively, demonstrating its robustness and generalizability across diverse contexts.

Looking ahead, several directions for future work emerged. Removing normal-operation labels from the FDD process would support broader applications in cases where extensive historical data existed without annotation. Applying the framework to real-world systems, rather than simulated or experimental ones, would facilitate field deployment. Transfer learning approaches also appeared promising, enabling knowledge transfer between case studies, especially when supported by metadata that describe the relationships between target and source systems. Eventually, LLMs to automate elements of the hybrid strategy could significantly enhance usability and create a more intuitive interface for end users.

CRediT authorship contribution statement

Marco Paolini: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization; **Marco Savino Piscitelli:** Writing – original draft, Validation, Supervision, Methodology, Investigation, Conceptualization; **Alfonso Capozzoli:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization.

Data availability

Data are open.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: co-author Marco Savino Piscitelli is an editorial board member in the Journal. Given his role, Marco Savino Piscitelli had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. Other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study has been carried out in the framework of the PRIN 2022 research project titled “UTMOST FDD: an aUToMated, Open, Scalable and Transparent Fault Detection and Diagnosis process for air-handling units based on a hybrid expert and artificial intelligence approach. From experimental open data to transfer model learning for the enhancement of energy management and indoor environmental quality in buildings” (CUP: E53D23002840006) funded by the Italian Ministry of University and Research (MUR). The work of Marco Savino Piscitelli was carried out within the Ministerial Decree no. 1062/2021 and received funding from the FSE REACT-EU-PON Ricerca e Innovazione 2014–2020. This manuscript reflects only the authors’ views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Appendix A.

The second case study is referred to a FCU which layout is reported in Fig. 21.

The considered variables were the following:

- **FCU_CVLV_DM**: Control signal for FCU cooling coil valve; ranges from 0 to 1; 0 – valve should be fully closed, 1 – valve should be fully open.
- **FCU_HVLV_DM**: Control signal for FCU heating coil valve; ranges from 0 to 1; 0 – valve should be fully closed, 1 – valve should be fully open.
- **FCU_DMPR_DM**: Control signal for FCU outdoor air damper; ranges from 0 to 1; 0 – damper should be fully closed, 1 – damper should be fully open.
- **RM_TEMP**: FCU room temperature [°C].
- **RMCLGSPT**: FCU cooling setpoint temperature [°C].
- **RMHTGSPT**: FCU heating setpoint temperature [°C].
- **OA_TEMP**: FCU outdoor air temperature [°C].
- **RA_TEMP**: FCU return air temperature [°C].
- **MA_TEMP**: FCU mixed air temperature [°C].
- **SA_TEMP**: FCU supply air temperature [°C].
- **FCU_DA_CFM**: FCU discharge air flow rate [m^3/h].
- **FCU_OA_CFM**: FCU outdoor air flow rate [m^3/h].
- **FCU_CLG_EWT**: FCU cooling coil entering water temperature [°C].
- **FCU_HTG_EWT**: FCU heating coil entering water temperature [°C].
- **FCU_CLG_RWT**: FCU cooling coil return water temperature [°C].
- **FCU_HTG_RWT**: FCU heating coil return water temperature [°C].
- **FCU_CLG_GPM**: FCU cooling coil water flow rate [kg/s].
- **FCU_HTG_GPM**: FCU heating coil water flow rate [kg/s].
- **P_COOLING**: thermal power exchanged by the cooling coil [W] (this variable is not included in the original dataset and was calculated starting from FCU_CLG_GPM, FCU_CLG_EWT, FCU_CLG_RWT)
- **P_HEATING**: thermal power exchanged by the heating coil [W] (this variable is not included in the original dataset and was calculated starting from FCU_HTG_GPM, FCU_HTG_EWT, FCU_HTG_RWT)
- **FCU_WAT**: FCU air fan power [W].
- **FCU_SPD**: Control signal for FCU air fan speed; ranges from 0 to 1; 0 - fan speed is 0 %, 1 - fan speed is 100 %.

In addition the considered fault types together with their severity are listed in Table 9. As in the SDAHU case study, once the normal operation data were identified, the baseline models were established. The developed 8 regression RF models had the following output and input variables:

- **MA_TEMP model**: to estimate the mixed air temperature were used as input variables OA_TEMP, RA_TEMP, SA_TEMP, FCU_DMPR_DM, FCU_OA_CFM, FCU_DA_CFM, FCU_WAT and RM_TEMP.
- **SA_TEMP model**: to estimate the supply air temperature were used as input variables OA_TEMP, RA_TEMP, MA_TEMP, FCU_DMPR_DM, FCU_OA_CFM, FCU_DA_CFM, FCU_CVLV_DM, FCU_HVLV_DM, P_COOLING, P_HEATING, FCU_WAT and RM_TEMP.
- **RA_TEMP model**: to estimate the return air temperature were used as input variables OA_TEMP, MA_TEMP, SA_TEMP, FCU_DMPR_DM, FCU_OA_CFM, FCU_DA_CFM and FCU_WAT.

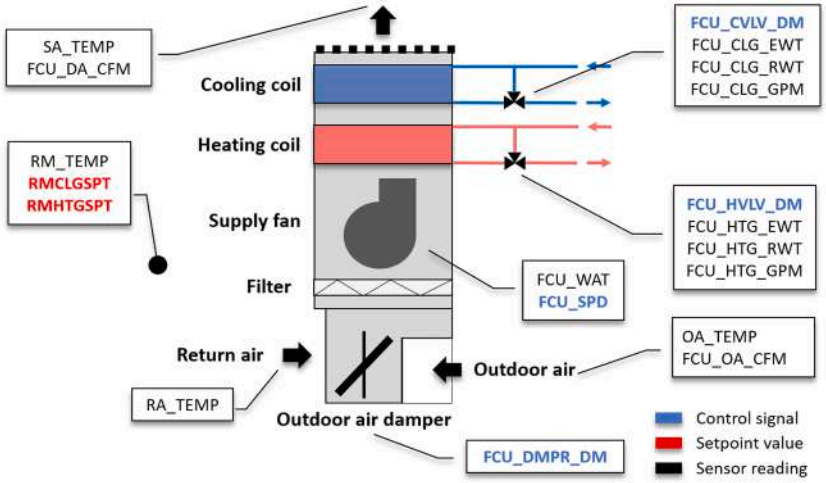


Fig. 21. Fan coil unit layout.

Table 9
Fault categories, fault types and fault intensities for FCU case study.

Faulted component	Fault type	Fault severity
Economizer	Outdoor air damper stuck	(1) 0 %, (2) 20 %, (3) 50 %, (4) 80 % and (5) 100 % of the position
	Outdoor air damper leaking	(1) +20 %, (2) +50 % and (3) +80 % of the damper face area
	Outdoor air damper blockage	(1) -80 % of the damper face area
Air duct	Fan outlet blockage	(1) +2400 % of the air flow pressure resistance
	Filter restriction	(1) +10 %, (2) +20 % and (3) +50 % of air flow pressure resistance
	Heating coil fouling air-side	(1) Severe, (2) Moderate and (3) Minor severity
	Cooling coil fouling air-side	(1) Severe, (2) Moderate and (3) Minor severity
Piping system	Heating coil fouling water-side	(1) Severe, (2) Moderate and (3) Minor severity
	Cooling coil fouling water-side	(1) Severe, (2) Moderate and (3) Minor severity
Heating coil	Heating coil valve stuck	(1) 0 %, (2) 20 %, (3) 50 %, (4) 80 % and (5) 100 % of the position
	Heating coil valve leaking	(1) 20 %, (2) 50 % and (3) 80 % of the nominal flow rate
Cooling coil	Cooling coil valve stuck	(1) 0 %, (2) 20 %, (3) 50 %, (4) 80 % and (5) 100 % of the position
	Cooling coil valve leaking	(1) 20 %, (2) 50 % and (3) 80 % of the nominal flow rate
Room sensor	Room temperature sensor bias	(1) −4°C, (2) −2°C, (3) +2°C and (4) +4°C

- **FCU_WATT model**: to estimate the air fan power were used as input variables OA_TEMP, MA_TEMP, SA_TEMP, RA_TEMP, FCU_DMPR_DM, FCU_OA_CFM, FCU_DA_CFM and RM_TEMP.
- **P_COOLING model**: to estimate the cooling coil power were used as input variables OA_TEMP, MA_TEMP, SA_TEMP, RA_TEMP, FCU_DMPR_DM, FCU_CVLV_DM, FCU_HVLV_DM, FCU_WATT and RM_TEMP.
- **P_HEATING model**: to estimate the heating coil power were used as input variables OA_TEMP, MA_TEMP, SA_TEMP, RA_TEMP, FCU_DMPR_DM, FCU_CVLV_DM, FCU_HVLV_DM, FCU_WATT and RM_TEMP.
- **FCU_CLG_GPM model**: to estimate the cooling coil water flow rate was used as input variable FCU_CVLV_DM.
- **FCU_HTG_GPM model**: to estimate the heating coil water flow rate was used as input variable FCU_HVLV_DM.

Considering all the possible operational modes of the FCU, six separate BNs were developed, each with a unique configuration of nodes and arcs. These are shown in Figs. 22, 23, and 24 corresponding to the heating, cooling and ventilation modes in both occupied hours and setback mode configurations.

As in the previous case study, virtual and hard evidences (e.g., baseline model residuals, differences from setpoint values etc.) were defined with different severity. In addition three expert rules were added as hard evidence with severe intensity to the Air duct, Economizer and Room sensor component. The expert rules are in the following described:

- **Expert rule Air Duct**: This rule is based on the calculation of a KPI and its comparison with reference values of that KPI extracted from the normal operation data. Specifically, the fan of the FCU has three different speed commands that determine the three reference values

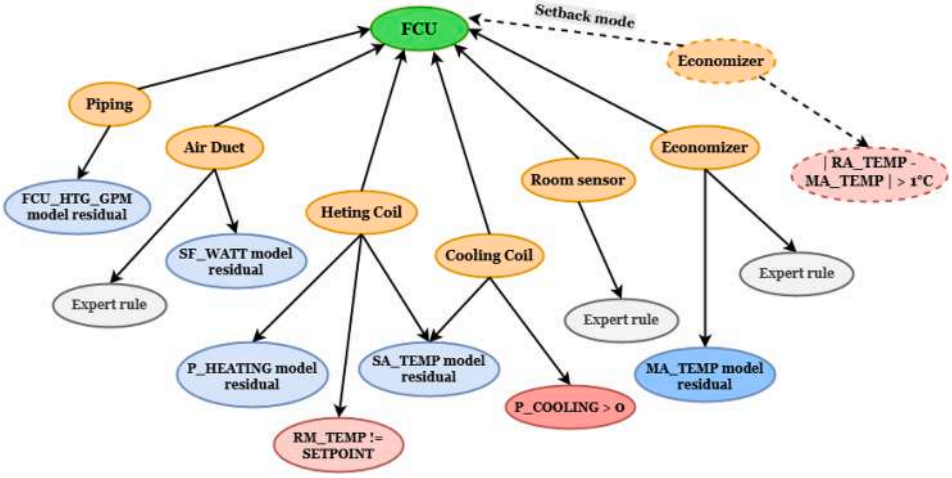


Fig. 22. BN structure for heating mode in both occupied hours and setback mode.

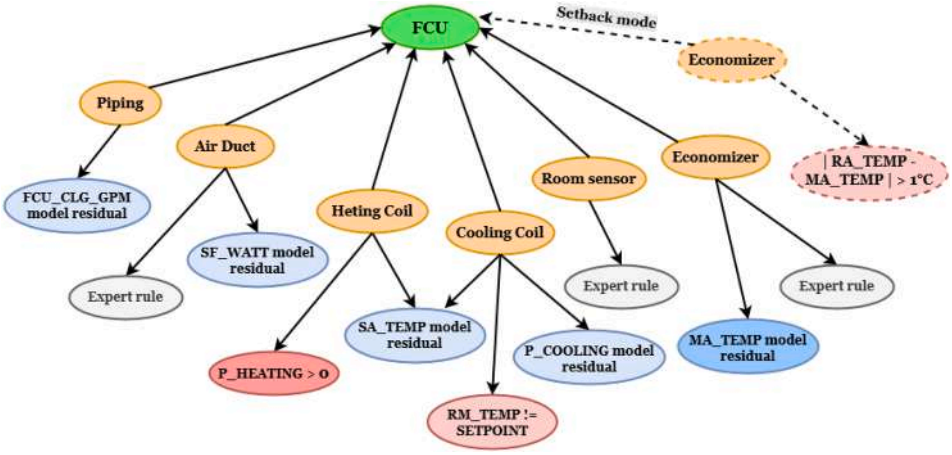


Fig. 23. BN structure for cooling mode in both occupied hours and setback mode.

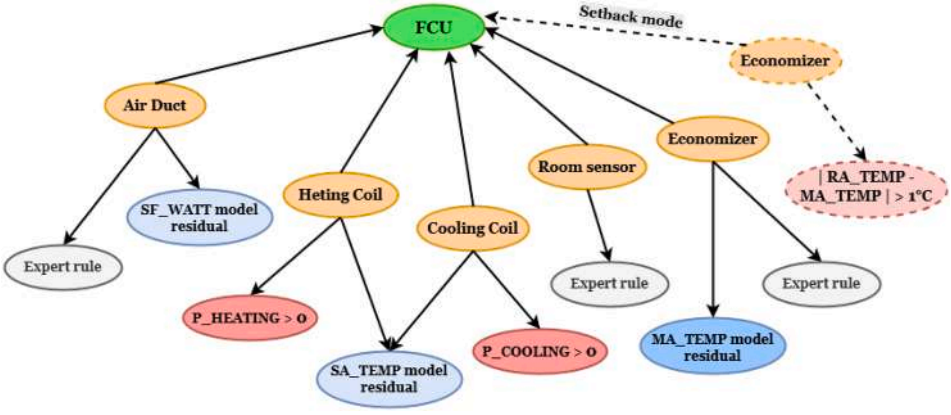


Fig. 24. BN structure for ventilation mode in both occupied hours and setback mode.

of the KPI.

$$\left(\frac{\text{FCU\_DA\_CFM}}{\text{SF\_WATT}}\right) < 0.8 \times \left(\frac{\text{FCU\_DA\_CFM}}{\text{SF\_WATT}}\right)_{\text{normal}}$$

$$\Rightarrow \quad P(\text{FAULT}) = 100\,\%$$

- **Expert rule Economizer**: This rule is based on the calculation of a KPI and its comparison with reference values of that KPI extracted from the normal operation data. Specifically, the outdoor air damper of the FCU has fixed positions according to the operational modes while the fan of the FCU has three different speed commands that determine the three reference values of the KPI.

$$\left(\frac{\text{FCU\_OA\_CFM}}{\text{FCU\_DA\_CFM}}\right) > 1.2 \times \left(\frac{\text{FCU\_OA\_CFM}}{\text{FCU\_DA\_CFM}}\right)_{\text{normal}}$$

$$\Rightarrow \quad P(\text{FAULT}) = 100\,\%$$

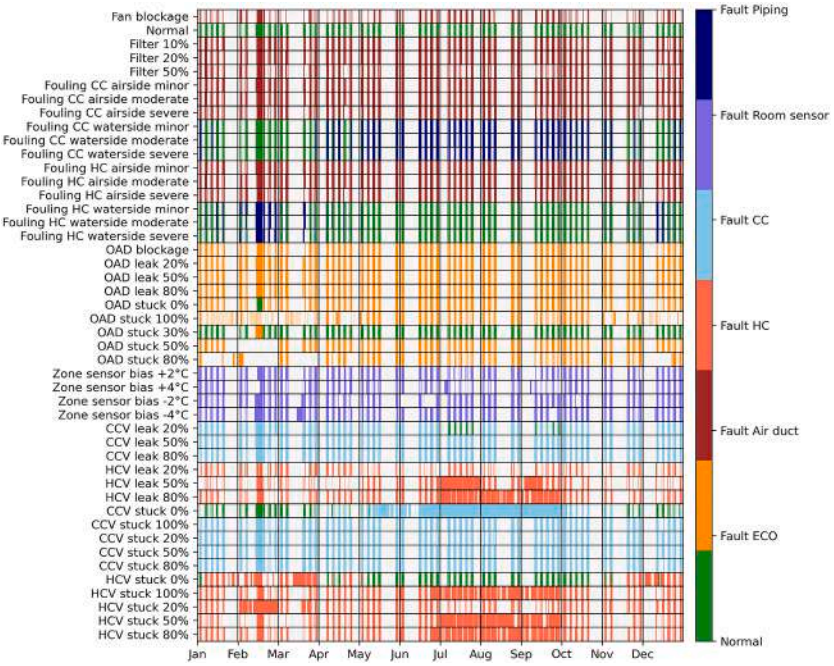


Fig. 25. Actual fault and normal labels for the FCU (Scenario 3).

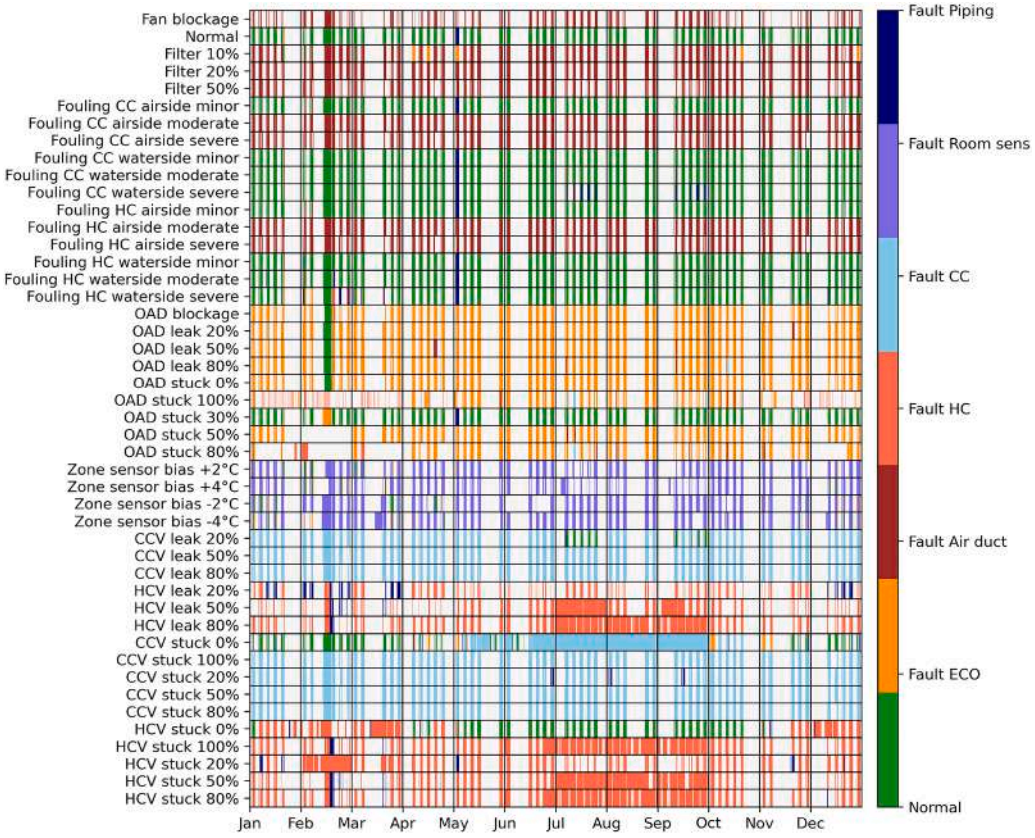


Fig. 26. Predicted fault and normal labels for the FCU (Scenario 3).

- **Expert rule Room Sensor**:

$$\left(|\text{RM\_TEMP} - \text{RA\_TEMP}| \leq 2^{\circ}\text{C}\right) \text{ and } \left(|\text{RA\_TEMP}_{\text{residual}}| > 3 \cdot \text{RMSE}_{\text{validation}}\right) \Rightarrow P(\text{FAULT}) = 100\,\%$$

From Table 8 it is possible to infer that the developed methodological framework was able to achieve an overall accuracy in detecting and isolating faulty components equal to 87.2 %. Specifically, a comparison between the actual and predicted labels over the entire testing set is reported in Figs. 25 and 26.

References

[1] 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) 110492. https://doi.org/10.1016/j.enbuild.2020.110492

[2] T. Daixin, X. Hongwei, Y. Huijuan, Y. Hao, H. Wen, Optimization of group control strategy and analysis of energy saving in refrigeration plant, Energy Built Environ. 3 (4) (2022) 525–535. https://doi.org/10.1016/j.enbenv.2021.05.006

[3] M.A. Piette, S.K. Kinney, P. Haves, Analysis of an information monitoring and diagnostic system to improve building operations, Energy Build. 33 (8) (2001) 783–791. https://doi.org/10.1016/S0378-7788(01)00068-8

[4] S. Katipamula, M. Brambley, Methods for fault detection, diagnostics and prognostics for building systems—A review part i, HVAC&R Res. 11 (2005). https://doi.org/10.1080/10789669.2005.10391133

[5] S. Katipamula, M. Brambley, Review article: methods for fault detection, diagnostics, and prognostics for building systems-a review, part II, HVAC&R Res. 11 (2005) 3–25. https://doi.org/10.1080/10789669.2005.10391123

[6] A. Costa, M.M. Keane, J.I. Torrens, E. Corry, Building operation and energy performance: monitoring, analysis and optimisation toolkit, Appl. Energy 101 (2013) 310–316. Sustainable Development of Energy, Water and Environment Systems, https://doi.org/10.1016/j.apenergy.2011.10.037

[7] 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, Renew. Sustain. Energy Rev. 161 (2022) 112395. https://doi.org/10.1016/j.rser.2022.112395

[8] J. Bi, H. Wang, E. Yan, C. Wang, K. Yan, L. Jiang, B. Yang, AI in HVAC fault detection and diagnosis: a systematic review, Energy Rev. 3 (2) (2024) 100071. https://doi.org/10.1016/j.enrev.2024.100071

[9] Y. Himeur, M. Elnour, F. Fadli, et al., AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives, Artif. Intell. Rev. 56 (2023) 4929–5021. https://doi.org/10.1007/s10462-022-10286-2

[10] C. Fan, Y. Lin, M.S. Piscitelli, R. Chiosa, H. Wang, A. Capozzoli, Y. Ma, Leveraging graph convolutional networks for semi-supervised fault diagnosis of HVAC systems in data-scarce contexts, Build. Simul. 16 (8) (2023) 1499–1517. https://doi.org/10.1007/s12273-023-1041-1

[11] F.J. Montáns, F. Chinesta, R. Gómez-Bombarelli, J.N. Kutz, Data-driven modeling and learning in science and engineering, C. R. Méc. 347 (11) (2019) 845–855. Data-Based Engineering Science and Technology, https://doi.org/10.1016/j.crme.2019.11.009

[12] C. Fan, W. He, Y. Liu, P. Xue, Y. Zhao, A novel image-based transfer learning framework for cross-domain HVAC fault diagnosis: from multi-source data integration to knowledge sharing strategies, Energy Build. 262 (2022) 111995. https://doi.org/10.1016/j.enbuild.2022.111995

[13] Z.-C. Wang, D. Li, Z.-W. Cao, F. Gao, M.-J. Li, A modified transformer and adapter-based transfer learning for fault detection and diagnosis in HVAC systems, Energy Storage Sav. 3 (2) (2024) 96–105. https://doi.org/10.1016/j.enss.2024.02.004

[14] J. Zhang, C. Zhang, J. Lu, Y. Zhao, Domain-specific large language models for fault diagnosis of heating, ventilation, and air conditioning systems by labeled-data-supervised fine-tuning, Appl. Energy 377 (2025) 124378. https://doi.org/10.1016/j.apenergy.2024.124378

[15] V. Singh, J. Mathur, A. Bhatia, A comprehensive review: fault detection, diagnostics, prognostics, and fault modeling in HVAC systems, Int. J. Refrig. 144 (2022) 283–295. https://doi.org/10.1016/j.ijrefrig.2022.08.017

[16] 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) 121030. https://doi.org/10.1016/j.apenergy.2023.121030

[17] M.S. Piscitelli, A. Hooman, A. Rosato, A. Capozzoli, Overview on Fault Detection and Diagnosis Methods in Building HVAC Systems: Toward a Hybrid Approach, Springer Nature Singapore, Singapore, 2024, pp. 709–719. https://doi.org/10.1007/978-981-99-8501-2_61

[18] M. Paolini, M.S. Piscitelli, A. Hooman, A. Rosato, A. Capozzoli, Experimental performance evaluation of cost-sensitive Bayesian networks for fault detection and diagnosis in HVAC systems, in: J.R. Littlewood, R.J. Howlett, L.C. Jain (Eds.), Sustainability in Energy and Buildings 2024, Springer Nature Singapore, Singapore, 2026, pp. 461–472.

[19] W. Li, H. Li, S. Gu, T. Chen, Process fault diagnosis with model- and knowledge-based approaches: advances and opportunities, Control Eng. Pract. 105 (2020) 104637. https://doi.org/10.1016/j.conengprac.2020.104637

[20] G. Lin, J. House, Y. Chen, J. Granderson, W. Zhang, Active multi-mode data analysis to improve fault diagnosis in AHUs, Energy Build. 337 (2025) 115621. https://doi.org/10.1016/j.enbuild.2025.115621

[21] M. Lampis, J.D. Andrews, Bayesian belief networks for system fault diagnostics, Qual. Reliab. Eng. Int. 25 (4) (2009) 409–426. https://doi.org/10.1002/qre.978

[22] B. Balaji, A. Bhattacharya, G. Fierro, J. Gao, J. Gluck, D. Hong, A. Johansen, J. Koh, J. Ploennigs, Y. Agarwal, M. Bergés, D. Culler, R.K. Gupta, M.B. Kjærgaard, M. Srivastava, K. Whitehouse, Brick: metadata schema for portable smart building applications, Appl. Energy 226 (2018) 1273–1292. https://doi.org/10.1016/j.apenergy.2018.02.091

[23] J. Granderson, G. Lin, Y. Chen, A. Casillas, P. Im, S. Jung, K. Benne, J. Ling, R. Gorthala, J. Wen, Z. Chen, S. Huang, D. Vrabie, LBNL Fault Detection and Diagnostics Datasets, 2022, https://doi.org/10.25984/1881324

[24] B. Cai, L. Huang, M. Xie, Bayesian networks in fault diagnosis, IEEE Trans. Ind. Inf. 13 (5) (2017) 2227–2240. https://doi.org/10.1109/TII.2017.2695583

[25] J. Huang, N. Ghalamsiah, A. Patharkar, O. Pradhan, M. Chu, T. Wu, J. Wen, Z. O'Neill, K. Selcuk Candan, An entropy-based causality framework for cross-level faults diagnosis and isolation in building HVAC systems, Energy Build. 317 (2024) 114378. https://doi.org/10.1016/j.enbuild.2024.114378

[26] Z. Wang, B. Liang, J. Guo, L. Wang, Y. Tan, X. Li, Fault diagnosis based on residual–knowledge–data jointly driven method for chillers, Eng. Appl. Artif. Intell. 125 (2023) 106768. https://doi.org/10.1016/j.engappai.2023.106768

[27] D. Wu, H. Yang, K. Xu, X. Meng, S. Yin, C. Zhu, X. Jin, Data and knowledge fusion-driven Bayesian networks for interpretable fault diagnosis of HVAC systems, Int. J. Refrig. 161 (2024) 101–112. https://doi.org/10.1016/j.ijrefrig.2024.02.019

[28] A. Behravan, B. Kiamanesh, R. Obermaisser, Fault diagnosis of DCV and heating systems based on causal relation in fuzzy Bayesian belief networks using relation direction probabilities, Energies 14 (20) (2021). https://doi.org/10.3390/en14206607

[29] K. Verbert, R. Babuška, B. De Schutter, Combining knowledge and historical data for system-level fault diagnosis of HVAC systems, Eng. Appl. Artif. Intell. 59 (2017) 260–273. https://doi.org/10.1016/j.engappai.2016.12.021

[30] O. Pradhan, J. Wen, Y. Chen, X. Lu, M. Chu, Y. Fu, Z. O'Neill, T. Wu, K.S. Candan, Dynamic bayesian network-based fault diagnosis for ASHRAE guideline 36: high performance sequence of operation for HVAC systems, in: Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, BuildSys '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 365–368. https://doi.org/10.1145/3486611.3491124

[31] F. Xiao, Y. Zhao, J. Wen, S. Wang, Bayesian network based FDD strategy for variable air volume terminals, Autom. Constr. 41 (2014) 106–118. https://doi.org/10.1016/j.autcon.2013.10.019

[32] K.H. Ng, F.W.H. Yik, P. Lee, K.K.Y. Lee, D.C.H. Chan, Bayesian method for HVAC plant sensor fault detection and diagnosis, Energy Build. 228 (2020) 110476. https://doi.org/10.1016/j.enbuild.2020.110476

[33] S. He, Z. Wang, Z. Wang, X. Gu, Z. Yan, Fault detection and diagnosis of chiller using Bayesian network classifier with probabilistic boundary, Appl. Therm. Eng. 107 (2016) 37–47. https://doi.org/10.1016/j.applthermaleng.2016.06.153

[34] J. Xu, Q. Wang, J. Zhou, H. Zhou, J. Chen, Improved Bayesian network-based for fault diagnosis of air conditioner system, Int. J. Metrol. Qual. Eng. 14 (2023). https://doi.org/10.1051/ijmqe/2023009

[35] Z. Wang, Z. Wang, S. He, X. Gu, Z.F. Yan, Fault detection and diagnosis of chillers using Bayesian network merged distance rejection and multi-source non-sensor information, Appl. Energy 188 (2017) 200–214. https://doi.org/10.1016/j.apenergy.2016.11.130

[36] Z. Wang, L. Wang, K. Liang,y Y. Tan, Enhanced chiller fault detection using Bayesian network and principal component analysis, Appl. Therm. Eng. 141 (2018) 898–905. https://doi.org/10.1016/j.applthermaleng.2018.06.037

[37] Y. Wang, Z. Wang, S. He, Z. Wang, A practical chiller fault diagnosis method based on discrete Bayesian network, Int. J. Refrig. 102 (2019) 159–167. https://doi.org/10.1016/j.ijrefrig.2019.03.008

[38] Y. Chen, J. Wen, O. Pradhan, L.J. Lo, T. Wu, Using discrete Bayesian networks for diagnosing and isolating cross-level faults in HVAC systems, Appl. Energy 327 (2022) 120050. https://doi.org/10.1016/j.apenergy.2022.120050

[39] Y. Zhao, J. Wen, S. Wang, Diagnostic Bayesian networks for diagnosing air handling units faults—Part II: faults in coils and sensors, Appl. Therm. Eng. 90 (2015) 145–157. https://doi.org/10.1016/j.applthermaleng.2015.07.001

[40] Y. Zhao, J. Wen, F. Xiao, X. Yang, S. Wang, Diagnostic Bayesian networks for diagnosing air handling units faults—Part I: faults in dampers, fans, filters and sensors, Appl. Therm. Eng. 111 (2017) 1272–1286. https://doi.org/10.1016/j.applthermaleng.2015.09.121

[41] 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

[42] T. Gao, S. Marié, P. Béguery, S. Thebault, S. Lecoeuche, Integrated building fault detection and diagnosis using data modeling and Bayesian networks, Energy Build. 306 (2024) 113889. https://doi.org/10.1016/j.enbuild.2024.113889

[43] S. Fenz, An ontology-based approach for constructing Bayesian networks, Data Knowl. Eng. 73 (2012) 73–88. https://doi.org/10.1016/j.datak.2011.12.001

[44] R. Chen, Y. Lu, P. Witherell, T.W. Simpson, S. Kumara, H. Yang, Ontology-driven learning of Bayesian network for causal inference and quality assurance in additive manufacturing, IEEE Robot. Autom. Lett. 6 (3) (2021) 6032–6038. https://doi.org/10.1109/LRA.2021.3090020

[45] M.S. Sayed, N. Lohse, Ontology-driven generation of Bayesian diagnostic models for assembly systems, Int. J. Adv. Manuf. Technol. 74 (5) (2014) 1033–1052. https://doi.org/10.1007/s00170-014-5918-0

[46] M. Munch, P. Buche, S. Dervaux, J. Dibie, L. Ibanescu, C. Manfredotti, P.-H. Wuillemin, H. Angellier-Coussy, Combining ontology and probabilistic models for the design of bio-based product transformation processes, Expert Syst. Appl. 203 (2022) 117406. https://doi.org/10.1016/j.eswa.2022.117406

[47] M.M.-L. Pfaff-Kastner, K. Wenzel, S. Ihlenfeldt, Concept paper for a digital expert: systematic derivation of (causal) Bayesian networks based on ontologies for knowledge-Based production steps, Mach. Learn. Knowl. Extr. 6 (2) (2024) 898–916. https://doi.org/10.3390/make6020042

[48] G. Biau, E. Scornet, A random forest guided tour, TEST 25 (2) (2016) 197–227. https://doi.org/10.1007/s11749-016-0481-7

[49] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32. https://doi.org/10.1023/A:1010933404324

[50] A. Liaw, Classification and Regression by RandomForest, R News (2002).

[51] J. Pearl, Fusion, propagation, and structuring in belief networks, Artif. Intell. 29 (3) (1986) 241–288. https://doi.org/10.1016/0004-3702(86)90072-X

[52] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, San Francisco, CA, USA, 1988.
[53] J. Joyce, Bayes' theorem, in: E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, Metaphysics Research Lab, Stanford University, Fall 2021 edition, 2021.
[54] D. Barber, Machine Learning Concepts, Cambridge University Press, 2012, p. 305–321.
[55] G.E.P. Box, G.C. Tiao, Bayesian Inference in Statistical Analysis, John Wiley & Sons, 2011.
[56] R. van de Schoot, S. Depaoli, R. King, B. Kramer, K. Märtens, M.G. Tadesse, M. Vannucci, A. Gelman, D. Veen, J. Willemsen, C. Yau, Bayesian statistics and modelling, Nat. Rev. Methods Prim. 1 (1) (2021) 1. https://doi.org/10.1038/s43586-020-00001-2
[57] Z. Ji, Q. Xia, G. Meng, A review of parameter learning methods in Bayesian network, in: Advanced Intelligent Computing Theories and Applications: 11th International Conference, ICIC 2015, Fuzhou, China, August 20–23, 2015. Proceedings, Part III 11, Springer, 2015, pp. 3–12.
[58] F.V. Jensen, T.D. Nielsen, Bayesian Networks and Decision Graphs, Information Science and Statistics, Springer New York, second ed., 2007. https://doi.org/10.1007/978-0-387-68282-2
[59] N. Friedman, D. Koller, Being Bayesian about network structure. a Bayesian approach to structure discovery in Bayesian networks, Mach. Learn. 50 (1) (2003) 95–125. https://doi.org/10.1023/A:1020249912095
[60] 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 (2) (2004) 153–170. https://doi.org/10.1016/S0306-2619(03)00107-7
[61] T. Li, Y. Zhao, C. Zhang, J. Luo, X. Zhang, A knowledge-guided and data-driven method for building HVAC systems fault diagnosis, Build. Environ. 198 (2021) 107850. https://doi.org/10.1016/j.buildenv.2021.107850
[62] J. Schein, S.T. Bushby, N.S. Castro, J.M. House, A rule-based fault detection method for air handling units, Energy Build. 38 (12) (2006) 1485–1492. https://doi.org/10.1016/j.enbuild.2006.04.014
[63] 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) 110369. https://doi.org/10.1016/j.enbuild.2020.110369
[64] Y. Peng, S. Zhang, R. Pan, Bayesian network reasoning with uncertain evidences, Int. J. Uncertain., Fuzziness Knowledge-Based Syst. 18 (2010) 539–564. https://doi.org/10.1142/S0218488510006696
[65] A. Mrad, V. Delcroix, S. Piechowiak, P. Leicester, M. Abid, An explication of uncertain evidence in Bayesian networks: likelihood evidence and probabilistic evidence, Appl. Intell. (2015) 1–23. https://doi.org/10.1007/s10489-015-0678-6
[66] T. Gao, Integrated building fault detection and diagnosis using data modeling and Bayesian networks, Theses, Ecole nationale supérieure Mines-Télécom Lille Douai, 2020. https://theses.hal.science/tel-03151257.
[67] C.D. Roa, P. Raftery, R. Sun, L. Paul, A.K. Prakash, M. Pritoni, G. Fierro, T. Peffer, Towards a Stronger Foundation: Digitizing Commercial Buildings with Brick to Enable Portable Advanced Applications, Technical Report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2022.
[68] L. Zhang, M. Leach, Y. Bae, B. Cui, S. Bhattacharya, S. Lee, P. Im, V. Adetola, D. Vrabie, T. Kuruganti, Sensor impact evaluation and verification for fault detection and diagnostics in building energy systems: a review, Adv. Appl. Energy 3 (2021) 100055. https://doi.org/10.1016/j.adapen.2021.100055
[69] H. Yang, S. Cho, C.-S. Tae, M. Zaheeruddin, Sequential rule based algorithms for temperature sensor fault detection in air handling units, Energy Convers. Manag. 49 (8) (2008) 2291–2306. https://doi.org/10.1016/j.enconman.2008.01.029
[70] N. Torabi, H.B. Gunay, W. O'brien, R. Moromisato, A holistic sequential fault detection and diagnostics framework for multiple zone variable air volume air handling unit systems, Build. Serv. Eng. Res. Technol. 43 (2022) 605–625. https://api.semanticscholar.org/CorpusID:249478931.
[71] J. Schein, Results from field testing of embedded air handling unit and variable air volume box fault detection tools:, 2006, https://doi.org/10.6028/NIST.IR.7365