RADIATOR FAULT DETECTION IN A MULTI-ENERGY SOURCE BUILDING USING UNSUPERVISED LEARNING TECHNIQUES

Nima Monghasemi[1]*, Amir Vadiee[1], Stavros Vouros[2], Konstantinos Kyprianidis[1]

[1]Mälardalen University, School of Business, Society, and Engineering, Västerås, Sweden

[2]Mälardalen University, School of Innovation, Design, and Technology, Västerås, Sweden

*Corresponding Author: nima.monghasemi@mdu.se

ABSTRACT

As modern district heating networks integrate buildings with multiple energy sources, fault detection has become increasingly relevant and critical. This study investigates the effectiveness of an unsupervised data-driven fault detection approach to identify stuck valve and faulty thermostatic radiator valve scenarios in the baseboard radiators of an office building. A baseline model for a typical Swedish office building was developed, featuring a ground-source heat pump, solar photovoltaic-thermal panel, water-based radiators, and a connection to the district heating system to support its heating demand. Multiple fault scenarios were considered in the model, involving partially stuck valves and thermostatic radiator valves that deviated from their intended setpoints. Synthetic noise was added to generate faulty scenarios. The model performed well in detecting severe stuck valve faults but showed lower performance on less severe faults and faulty thermostatic radiator valves. The insights gained from this research emphasize the importance of fault monitoring in the context of evolving buildings connected to district heating networks.

1 INTRODUCTION

District heating (DH) systems have emerged as a pivotal element in the transition towards a sustainable energy future, particularly in the context of energy systems based on renewable energy sources (Connolly, et al. 2014). One of the key advantages of these systems is their ability to utilize heat that would otherwise be wasted from various sources, such as industrial processes, power stations, and waste incineration plants. Another important aspect of DH systems is their capacity to integrate low-cost energy storage solutions. This feature allows for better balancing of supply and demand, as well as enhancing the overall efficiency of the system. The incorporation of energy storage helps optimize the use of available heat sources and ensures a more stable and reliable heat supply to connected buildings. DH systems also enable flexible interplay between the electricity and heating sectors using combined heat and power (CHP) units, electric boilers, and heat pumps. This synergy further supports the role of DH systems in the energy landscape (Ebrahimi 2020).

A key limitation of third-generation district heating systems is their constrained ability to integrate renewable energy sources such as solar thermal energy. The high distribution temperatures and system design make the incorporation of these technologies challenging. Additionally, these systems often lack the flexibility to adapt to fluctuations in heat demand or supply (Lund, et al. 2018). Fourth-generation district heating (4GDH) systems have been proposed to address these limitations. These systems are characterized by lower distribution temperatures, resulting in reduced thermal losses and improved energy efficiency. This approach enables the integration of renewable energy sources, such as solar power, which helps decrease reliance on fossil fuels. These systems are flexible and can adapt to changing heat demands by incorporating emerging technologies (Sorknæs, et al. 2020).

https://doi.org/10.52202/077185-0213

The incorporation of emerging energy technologies poses a range of modeling difficulties and integration complexities. Predicting the performance of heat pumps and solar photovoltaics under diverse real-world conditions is hampered by the limited availability of operational data (Majidi, et al. 2019). Additional modeling complexity arises from the dynamic interactions between multiple interconnected energy-conversion technologies, which can exhibit transient behaviors. Human behavior and acceptance of new technologies, as well as the implementation of control strategies, introduce further variables that are difficult to quantify.

The integration of distributed energy resources into modern district heating grids results in complex systems, where it is crucial to ensure that all components function properly (Månsson, et al. 2018). Fault management is a valuable tool for identifying vulnerabilities and predicting potential failures. In the context of heating, ventilation, and air conditioning (HVAC) systems, a fault or failure occurs when a system, equipment, or component performs in a way that negatively impacts the thermal comfort or energy efficiency of a building (Li and O'Neill 2019). These faults can be gradual or abrupt and can result in increased energy consumption and decreased thermal comfort. Similar fault management principles can be applied to buildings. HVAC faults, such as sensor errors and equipment failures, are prevalent in commercial buildings; these faults degrade system performance and account for 15-30% of lost energy in buildings (Katipamula and Brambley 2005, Qin and Wang 2005). In a survey with Swedish district heating utilities, Månsson, et al. (2019) investigated the most common faults in customer installations causing high return temperatures and the strategies employed by successful utilities to address these issues. It was found that gaining access to customer installations and proactively identifying and fixing faults, particularly leakages and problems in customers' internal heating systems, were key to maintaining low return temperatures across the system. Neumayer, et al. (2023) conducted a literature review of fault and anomaly detection techniques for district heating substations. This study discussed the prevalence of both conventional methods, such as thresholds and visualization, as well as machine learning approaches. Unsupervised learning is more widely applied in this domain owing to the lack of labeled training data.

As energy systems in buildings become more interconnected, managing the resulting complexity requires sophisticated fault detection methods to maintain optimal performance. Machine learning is gaining popularity for building system fault detection and diagnostics owing to its ability to handle large amounts of sensor data and produce more accurate and reliable results compared to traditional rule-based methods (Nelson and Culp 2022). Bode, et al. (2020) developed a machine-learning fault detection algorithm and investigated its transferability from experimental data to real-world applications. The algorithm was trained on an experimental dataset containing data for typical heat pump failures measured on an outfitted air-water heat pump. Although it performed well in controlled experiments, its effectiveness diminished significantly when applied to real-world data. This underscores the difficulties in adapting machine learning models to different systems with varying data quality and labeling standards. Data-driven fault detection and diagnostics (FDD) methods have been extensively applied to various HVAC components, with air handling units (AHUs), variable air volume (VAV) terminal units, and chillers being the most studied (Chen, et al. 2023). These components are critical to building performance and are prone to faults, such as sensor issues, damper problems, and fouling. Yan, et al. (2016) investigated sensor fault detection in air handling units (AHUs) within HVAC systems, an area critical for maintaining energy efficiency and indoor comfort in commercial buildings. Recognizing the limitations of previous fault detection and diagnosis (FDD) strategies, the authors identified two main issues. Traditional methods often depend on substantial training data and struggle to generalize across varying operational conditions. To address these challenges, an unsupervised learning-based strategy was employed. Their research distinguished itself by effectively pinpointing both individual and multiple sensor faults without the need for extensive user-defined input parameters, thus addressing a prevailing challenge in clustering-based FDD methodologies.

The use of extensive sensor networks can be expensive and complex, particularly in older buildings. Sensors can also have issues such as failures, calibration needs, and communication disruptions, leading to unreliable data. In addition, some faults may not be directly detected by sensors, such as gradual

https://doi.org/10.52202/077185-0213

performance decay or complex interactions between system components. When direct sensor measurements are insufficient or unavailable, a fault model aims to generate simulated faulty data by modeling the faulty behavior of a component or system under different possible failure conditions (Li and O'Neill 2018). Lu, et al. (2021) conducted an extensive fault impact analysis to assess the robustness of high-performance control sequences recommended in ASHRAE Guideline 36. A virtual testbed of a medium-sized office building was created, incorporating both airside and plant-side sequences of operation. A total of 359 common fault scenarios across various categories, including sensors, ducts/pipes, dampers/valves, HVAC equipment, controls, schedules, and design/construction aspects, were simulated under diverse seasonal conditions and fault intensities. Key performance indicators related to energy use, indoor environment, control performance, and power systems were chosen to quantify the effects of individual faults. The implications of this fault simulation dataset are multifaceted. This approach not only provides a more extensive and well-structured collection of fault symptoms that correspond with Guideline 36 but also enables the recognition of faults that exert the most adverse effects on system performance. This prioritization of fault characteristics can guide the development of more effective fault-handling strategies, which is crucial for the future advancement of reliable automated control systems.

While existing research has explored fault detection in HVAC systems, there is a lack of focus on specific radiator faults, particularly in the context of multi-energy source buildings connected to district heating networks. This study aims to fill this gap by investigating the effectiveness of an unsupervised learning approach for identifying stuck valve faults and faulty thermostatic radiator valves (TRVs). An unsupervised data-driven fault detection approach using a one-class support vector machine (OCSVM) is proposed, and its performance is evaluated using various metrics. The proposed approach aims to address one of the challenges that may arise in modern district heating networks on the consumer side, namely, the need for effective fault detection. As these networks become more integrated with various energy systems, ensuring that faults can be detected quickly and accurately becomes crucial.

2 MODEL DESCRIPTION

A baseline model is created using a typical Swedish office building, following the guidelines for materials and general recommendations. The building under study is a one-floor office building in Västerås, Sweden. It is composed of a single thermal zone, representing the area adjacent to the exterior walls, and is directly influenced by the outdoor weather conditions, solar radiation, and heat transfer through the building envelope. The total floor area of the building is 68 m^2. The building envelope thermal properties is reported in Table 1. Specifically, the occupancy density per effective area in the office building is 0.05, which is derived from the Swedish building standard Sveby, applicable from Monday to Friday, between 8am and 17pm (Sveby 2013). For the office building, a metabolic rate of 1.2 met was used, as specified corresponding to sedentary activity typically found in office environments (Ahmed, et al. 2017). In summer, the convective heat gain is 148 W, and the radiative heat gain is 130 W. During winter, the convective heat gain is 128 W, and the radiative heat gain is 132 W. The internal heat gain from electricity usage (lighting and appliances) is set to 50 kWh/m^2 according to Sveby's standard annual value. Infiltration rates vary with wind conditions and outdoor temperatures. For simplicity, this model assumes a constant average infiltration rate of 0.1 air changes per hour.

Table 1: Building envelope thermal properties.

Surface type	Surface area (m^2)	Heat transfer coefficient (W/m^2K)
Exterior wall	56.3	0.510
Interior wall	69	0.508
Ground floor	67.3	0.039
Ceiling	67.3	4.153
Window	11.2	2.89

https://doi.org/10.52202/077185-0213

As illustrated in Figure 1, the HVAC system model, which features a comprehensive array of components, is developed in TRNSYS Simulation Studio and serves as a dependable solution for heating. It includes a ground-source heat pump, solar photovoltaic-thermal (PVT) panels, hydronic baseboard radiators, and district heating as the primary heat source. The district heating substation supplies hot water, which is circulated through baseboard radiators installed throughout the zone. As a supplementary heating system, the building has a ground source heat pump that utilizes a network of underground pipes to transfer heat to and from the Earth. The PVT panels on the roof generate additional thermal energy for heating. When the stored water temperature reaches a certain point, it supplies the heat pump, which extracts thermal energy and distributes it indoors through radiators. The setpoint temperature for the zone is programmed differently for weekdays and weekends. On working days, the temperature is set to 21°C from 6am to 5pm. Outside of those hours, the setback temperature is 19°C. On weekends, when the office is unoccupied, temperature control is set to activate heating once the zone temperature drops below 18°C.The integration of various components allows for an effective heating solution that aligns with the principles of next-generation district heating systems (Gong, et al. 2023). The specifications of each component are listed in Table 2.

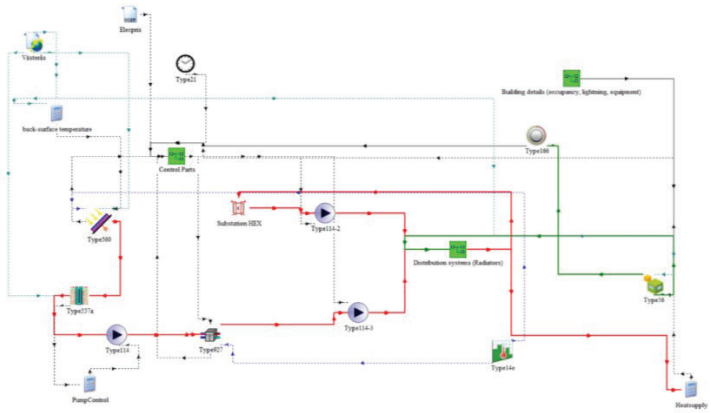

Figure 1: Simulation model in TRNSYS Simulation Studio.

Table 2: Components and parameters of the system.

TRNSYS source	Parameter	Value
Type 560- PV/T collector	Collector area	65.8 m^2
	Absorber plate thickness	0.003 m
	Thermal conductivity of the absorber	720 kJ/hmK
	Number of tubes	240
	Tube diameter	0.012 m
	Bond width	0.012 m
	Bond thickness	0.003 m
	Bond thermal conductivity	162 kJ/hmK
	PV efficiency at reference condition	17 %
Type 557-Vertical Ground Heat Exchanger	Storage volume	4500 m^3
	Borehole depth	100 m
	Header depth	5 m
	Number of boreholes	8
	Borehole radius	0.1 m
	Storage thermal conductivity	4.68 kJ/hmK
	Storage heat capacity	2016 kJK/m^3
	Fill thermal conductivity	1.3 W/mK
	Pipe thermal conductivity	1.5 W/mK
	Gap thermal conductivity	5.04 W/mK

https://doi.org/10.52202/077185-0213

	Gap thickness	0.01 m
	Heat capacity of layer	2016 kJK/m³
	Thickness of layer	1000 m
Type 927-Water-to-Water heat pump	Rated cooling capacity per heat pump	2 kW
	Rated cooling power per heat pump	1.6 kW
	Rated heating capacity per heat pump	10 kW
	Rated heating power per heat pump	3.2 kW
Type 91-Substation heat exchanger	Primary supply temperature	80 °C
	Primary flow rate	2400 kg/h
	Heat exchanger effectiveness	0.9
Type 1231- Hydronic baseboard radiators	Design capacity	12 kW
	Design surface temperature	55 °C
	Number of pipes	4
	Pipe inside diameter	0.015 m
	Air pressure exponent	0.2

The temperature setpoint triggers the heating system to turn on or off. Additionally, a decision must be made regarding the heating source: whether sufficient heat should be provided by the heat pump or district heating system. The operation of the heating system is governed by a rule-based controller that evaluates the cost of heating for each hour and determines which system to run based on that calculation. The running cost of the ground source heat pump was calculated based on the average hourly grid electricity consumption of the heat pump compressor. The heat pump compressor can be powered by either electricity from the grid or PVT electricity. If PVT electricity is accessible, it is utilized to power the compressor first. However, if there is insufficient PV electricity, the compressor operates on grid electricity. Several factors were considered when determining the cost of grid electricity, including the spot price, cost of electricity certificates, variable network fee, energy tax, value-added tax, and annual fixed network fee for maintaining the electricity grid (Monghasemi, et al. 2023). District heating prices in Sweden have both fixed and variable components. The fixed component is typically associated with administrative and operating costs, whereas the variable component is related to the actual production and distribution of heat based on consumption. Customers pay an annual fixed fee to connect to the district heating network, which covers administrative costs. They also pay a variable price per kWh of heating used. This rate fluctuates seasonally, with higher prices in the winter heating months and lower prices in the summer months. The local district heating company sets specific fees and rates (Lygnerud, et al. 2023). The rule-based controller then compares the cost of operating the heat pump to the cost of operating the district heating system and selects the system with the lowest cost. If the cost of operating the heat pump is lower, the heat pump is activated, and the district heating system is deactivated. Conversely, if the cost of operating the heat pump is higher, the district heating system is activated, and the heat pump is deactivated.

3 FAULT CONSTRUCTION

Fault construction allows faults to be simulated in the model by introducing faults into the simulation. It maps parameters to inputs, sets up fault occurrence distributions, and creates performance models that represent device or component failures that cause inefficient energy use or discomfort. Fault construction during the simulation of a building's HVAC system involves introducing errors or malfunctions into the simulation model to study their effects on system performance (Li and O'Neill 2019). This study focuses on radiator faults, specifically considering two main fault types: stuck valves and TRV faults. Stuck valve faults occur when a valve in the radiator system becomes stuck at a specific position: fully open, partially open, or fully closed. This can lead to uneven heat distribution, overheating, or underheating. Stuck valve faults can be caused by various factors, such as mechanical wear, corrosion, or debris accumulation, and can significantly impact the system's ability to regulate heat effectively. TRV faults, on the other hand, involve malfunctions in the thermostatic radiator valves, which are designed to control the flow of hot water through the radiators based on the desired room temperature. TRV faults can occur owing to various factors, such as sensor failures, calibration issues,

https://doi.org/10.52202/077185-0213

or mechanical problems. These issues can lead to inaccurate temperature readings, causing malfunctioning in the system. The fault scenarios considered in this study are listed in Table 3.

Table 3: Fault scenarios for radiators.

Fault type	Fault identifier	Description
Stuck valve	SV30-15	Valve 30% closed, occurring for 15% of simulation time
	SV30-30	Valve 30% closed, occurring for 30% of simulation time
	SV50-15	Valve 50% closed, occurring for 15% of simulation time
	SV50-30	Valve 50% closed, occurring for 30% of simulation time
Faulty TRVs	TV30-15	TRVs 30% deviated, 15% of simulation time
	TV30-30	TRVs 30% deviated, 30% of simulation time.
	TV50-15	TRVs 50% deviated, 15% of simulation time.
	TV50-30	TRVs 50% deviated, 30% of simulation time.

Fault detection relies on monitoring a range of variables that can indicate potential system issues. However, certain critical data points may not be readily available through standard sensors or may not be explicitly modeled in simulation platforms. Table 4 presents the data types monitored in this study for fault detection.

To simulate the impact of faults on the system, the behavior of specific components within the model was modified. Stuck valve faults were replicated by adjusting the parameters related to the valve flow characteristics, effectively simulating a restriction in flow through the affected radiators. TRV faults, on the other hand, were simulated by introducing an offset to the desired room-temperature setpoint that the TRV component uses for control. A positive offset represents a TRV that perceives the room as colder than it is, leading to potential overheating, whereas a negative offset simulates a TRV that senses a warmer room, potentially causing underheating. The timing and duration of fault occurrences were defined based on fault tags, which specify the percentage of total simulation time during which the fault is active. For each fault scenario, such as "SV30-15" or "TV50-30, " the fault duration was defined as a percentage of the total simulation time. In the case of "SV30-15" the stuck valve fault is set to occur for 15% of the simulation time, while in "TV50-30" the faulty TRV fault is active for 30% of the simulation time. Faults are not necessarily consecutively introduced for the entire specified duration. Instead, the faults were introduced intermittently, with the total aggregated duration of the fault occurrences equal to the specified percentage. To achieve this intermittent fault behavior, random start times were generated within the simulation period for each fault occurrence. The duration of each individual fault occurrence was also randomized, ensuring that the total aggregated duration of all fault occurrences matched the specified percentage defined in the fault tag. For example, in the "SV30-15" fault scenario, the stuck valve fault might occur in three separate intervals of 5% duration, each randomly distributed throughout the simulation period. The total aggregated duration of these three fault occurrences would then sum to the specified 15% of the total simulation time.

https://doi.org/10.52202/077185-0213

Table 4: Monitoring variables to be used as features for fault detection.

Monitoring variable	Unit	Description
T_{amb}	°C	Ambient temperature
T_{air}	°C	Indoor air temperature
Q_{sens}	kJ/h	Heating load
T_{in}	°C	Inlet water temperature
T_{out}	°C	Outlet water temperature
m_{rad-dh}	kg/h	Radiator mass flow rate (flow from substation)
m_{rad-hp}	kg/h	Radiator mass flow rate (flow from heat pump)
T_{surf}	°C	Radiator surface temperature
Q_{rad-dh}	kJ/h	Radiator heat transfer rate (flow from substation)
Q_{rad-hp}	kJ/h	Radiator heat transfer rate (flow from heat pump)
E_{hp}	kJ/h	Heat pump electricity consumption
Q_{dh}	kJ/h	District heating energy consumption
γ	-	On/off controller
γ_{hp}	-	Control signal for heat pump
γ_{dh}	-	Control signal for district heating

4 FAULT DETECTION

An OCSVM was employed for fault detection. The data preparation part has already been addressed in the fault construction stage. The features were normalized using min–max scaling to ensure that no single feature dominated the distance calculations owing to its scale. Once the data were prepared, the model was trained using non-faulty data instances. Based on these data, the OCSVM constructed a model that delineated the boundary of normal behavior in the feature space, effectively distinguishing between regular operation and potential anomalies. To optimize the performance of the OCSVM model, hyperparameter tuning using a grid search with 5-fold cross-validation was conducted. The hyperparameters tuned included the kernel type, outlier fraction, and kernel coefficient. The best hyperparameters were selected by maximizing the average accuracy of the model for predicting the normal class. Upon completion of model training, the OCSVM was applied to new datasets in the testing phase to evaluate its efficacy in anomaly detection. The role of the model was to assess whether new data points resided within the established normal boundary. In this section, the dataset comprises both faulty and non-faulty datasets collectively. Data points that fell outside this boundary were flagged as anomalies, suggesting possible faults.

5 RESULTS AND DISCUSSION

In this section, the results of applying the fault detection strategy are presented. Several performance metrics were analyzed to evaluate the capability of the model to address the faults within the system. To assess the performance of the fault detection mechanism, various evaluation metrics were employed.

Figure 2 illustrates the difference between the fault-free and faulty datasets for different severities when the stuck valve occurs. This reveals the impact of stuck valve faults on the energy demand in the system. The severity of the fault, determined by the percentage the valve is stuck closed, and the duration of the fault both contribute to increased energy consumption. As the severity and duration of the fault increase, the energy demand increases accordingly. This higher energy demand in faulty situations is attributed to the system requiring more energy to achieve the desired set-point temperature of the zone. When a valve is partially closed, it restricts the flow of the heat transfer fluid, making it more difficult for the system to maintain the target temperature. Consequently, the system must consume more energy to compensate for the reduced heat-transfer efficiency caused by the faulty valve.

https://doi.org/10.52202/077185-0213

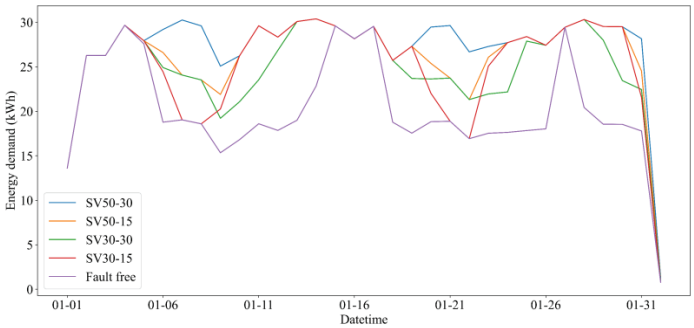

Figure 2: Energy demands in the presence and absence of stuck valve faults.

Figure 3 shows the energy consumption over time for different TRV fault scenarios. While most faulty scenarios exhibit a higher energy demand than the fault-free case, there is no consistent trend correlating the severity and duration of TRV faults with energy consumption. This can be attributed to the nature of the TRV faults, which involve the malfunctioning of the thermostatic radiator valves responsible for regulating the flow of hot water based on the desired room temperature. Unlike stuck valve faults that physically restrict flow, TRV faults can lead to inaccurate temperature sensing, causing the system to either overheat or underheat the rooms. When a TRV underestimates the required temperature, it may cause the system to provide more heat than necessary, leading to increased energy consumption. Conversely, when a TRV overestimates the temperature, it may restrict the flow of hot water to the radiator, potentially resulting in lower energy consumption compared to the fault-free scenario. There are instances where the TV30-30 scenario results in lower energy consumption than the fault-free case. This observation suggests that the TRV might have underestimated the required heat during these periods, leading to reduced energy consumption. The occurrence of such exceptions, in which faulty TRVs can lead to lower energy consumption, highlights the complexity of their impact. This is in contrast with the more predictable effects of stuck valve faults, where increased severity and duration consistently result in a higher energy demand.

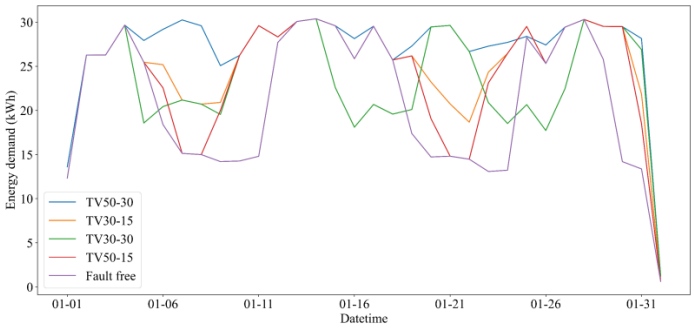

Figure 3: Energy demands in the presence and absence of TRV faults.

5.1 Performance metrics

Table 5 lists the key performance metrics of the models. It was observed that the model demonstrated varying performances across different fault scenarios in the heating system. For stuck valve faults, the model achieved excellent performance in detecting faults with 50% severity, regardless of the fault duration. This suggests that the model is effective in identifying more severe stuck valve faults, even if they persist for shorter or longer durations. On the other hand, for SV faults with 30% severity, the model's recall and F1-score were relatively lower, especially for shorter fault durations. This indicates that the model missed some instances of less severe stuck valve faults, particularly when they occurred for a brief period.

The recall and F1-score values for TV faults are consistently lower than those for SV faults, suggesting that the model may have more difficulty detecting incorrect temperature sensor faults in radiator

https://doi.org/10.52202/077185-0213

thermostats. One possible explanation for the lower performance of the model in detecting TRV faults could be the nature and characteristics of these faults. TRV faults, which occur when the radiator thermostat incorrectly senses the temperature, can have less obvious indications than stuck valve faults. The impact of TRV faults on the overall heating system behavior is more gradual, making them harder to detect using the model. Another factor to consider is the complexity of the temperature dynamics in the heating system. The relationship between the TRV faults and their impact on the system temperature may be more intricate and nonlinear compared to the direct impact of stuck valve faults on the flow of heat. This complexity could make it more challenging for the model to capture and differentiate faults associated with TRV faults.

Table 5: Performance metrics of the OCSVM model for different fault scenarios.

Fault	Accuracy	Precision	Recall	F1-score
SV30-15	0.94	0.92	0.69	0.79
SV30-30	0.88	0.95	0.63	0.76
SV50-15	0.99	0.94	1	0.97
SV50-30	0.98	0.96	1	0.98
TV30-15	0.85	0.97	0.51	0.66
TV30-30	0.79	0.98	0.59	0.74
TV50-15	0.85	0.97	0.53	0.69
TV50-30	0.81	0.98	0.62	0.76

5.2 Confusion matrix

To gain a more comprehensive understanding of the performance of the OCSVM model in detecting faults, the confusion matrices for each fault scenario were aggregated based on the fault types. Specifically, the confusion matrix for stuck valve faults with different severities (30% and 50%) and durations (15 and 30) were summed to create an overall confusion matrix for SV faults. Similarly, confusion matrices for thermostatic radiator valve (TRV) faults with different severities and durations were aggregated to obtain an overall confusion matrix for TRV faults. By aggregating the confusion matrices within each fault category, the model's performance in detecting SV and TRV faults was assessed separately. This approach allows for a more focused analysis of the strengths and weaknesses of the model in identifying specific types of faults by considering the variations in severity and duration within each category.

Figure 4 presents a confusion matrix for the detection model performance on aggregated stuck valve faults, which offers valuable insights into the classification capabilities of the model. The matrix clearly demonstrates the capacity of the model to accurately pinpoint defective instances, as indicated by the substantial number of correct faulty detections. Similarly, the model shows a strong aptitude for precisely categorizing non-faulty or acceptable instances, as evidenced by the considerable number of accurate OK detections. The model incorrectly classified 336 non-faulty instances as faulty, which could potentially lead to unnecessary maintenance or intervention. Conversely, 90 faulty instances were misclassified as non-faulty, indicating missed fault detections that could result in the potential performance degradation of the system if left unaddressed.

https://doi.org/10.52202/077185-0213

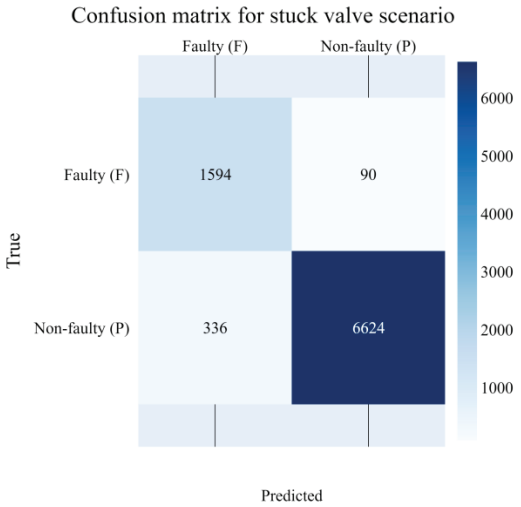

Figure 4: Confusion matrix for aggregated stuck valve faults.

Figure 5 presents a confusion matrix for the performance of the OCSVM model on aggregated TRV faults, providing insights into the model's classification capabilities for this fault type. The matrix highlights the ability of the model to correctly identify faulty instances, as indicated by the high number of true faulty detections. However, it is important to note that the model exhibits a higher number of misclassifications compared to stuck valve faults. The model incorrectly classified 1458 normal instances as faulty, which could lead to significant false alarms and unnecessary maintenance. This higher false positive rate suggests that the model may struggle to differentiate normal operation from TRV faults. However, the model correctly classified many normal instances, indicating its accuracy in identifying a typical performance. The low false negative rate suggests that the model rarely overlooks actual TRV faults, which is crucial for timely detection before further degradation.

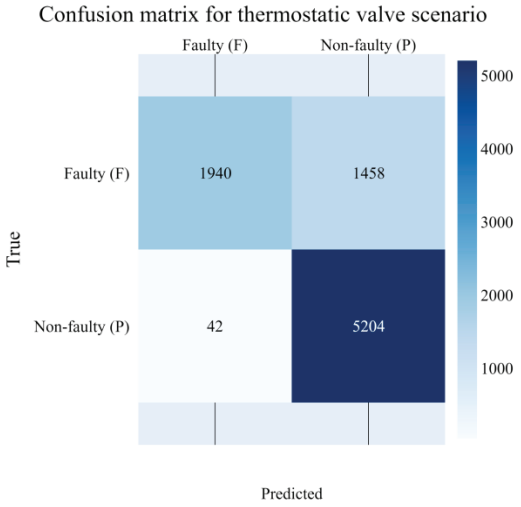

Figure 5: Confusion matrix for aggregated thermostatic valve faults.

Overall, the OCSVM model demonstrated its ability to detect faults in the heating system, with particularly strong performance in identifying severe stuck valve faults.

https://doi.org/10.52202/077185-0213

6 CONCLUSIONS

This study developed an unsupervised data-driven approach using a one-class support vector machine for fault detection in the radiator system of a commercial building. The results demonstrated that the model was effective in identifying severe stuck valve faults, achieving high recall and F1-scores. However, its performance was lower in detecting less severe stuck valve and TRV faults. This highlights the importance of considering fault characteristics, such as severity, when implementing fault detection systems. The confusion matrix analysis provided valuable insights into the model's strengths in correctly classifying normal and faulty instances, as well as its weaknesses in misclassifying some instances. The proposed approach can help maintenance personnel proactively identify and address faults to improve system efficiency and indoor comfort. Incorporating domain knowledge to guide feature selection and expand training data through the simulation of additional fault scenarios may help enhance model performance.

NOMENCLATURE

C	cost	SEK
E	electricity consumption	SEK
G	total tilted solar radiation	kJ/h
M	mass flow rate	kg/h
Q	heat transfer rate	kJ/h
T	temperature	℃

Abbreviations

4GDH	Fourth-generation district heating
AHU	Air handling unit
CHP	Combined heat and power
DH	District heating
FDD	Fault detection and diagnostic
OCSVM	One-class support vector machine
PVT	Photovoltaic thermal
SV	Stuck valve
TRV	Thermostatic radiator valve
VAV	Variable air volume

Greek symbols

γ	On/off controller signal

Subscript

amb	ambient
air	indoor air
dh	district heating
hp	heat pump
pvt	photovoltaic thermal collector
rad	radiator
sens	heating load
surf	surface
in	indoor
out	outlet

REFERENCES

Ahmed, K., J. Kurnitski and B. Olesen, "Data for occupancy internal heat gain calculation in main building categories," *Data in brief,* 15, 1030-1034 (2017).

Bode, G., S. Thul, M. Baranski and D. Müller, "Real-world application of machine-learning-based fault detection trained with experimental data," *Energy,* 198, 117323 (2020).

Chen, Z., 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 and

https://doi.org/10.52202/077185-0213

T. Herr, "A review of data-driven fault detection and diagnostics for building HVAC systems," *Applied Energy,* 339, 121030 (2023).
Connolly, D., H. Lund, B. V. Mathiesen, S. Werner, B. Möller, U. Persson, T. Boermans, D. Trier, P. A. Østergaard and S. Nielsen, "Heat Roadmap Europe: Combining district heating with heat savings to decarbonise the EU energy system," *Energy Policy,* 65, 475-489 (2014).
Ebrahimi, M., "Storing electricity as thermal energy at community level for demand side management," *Energy,* 193, 116755 (2020).
Gong, Y., G. Ma, Y. Jiang and L. Wang, "Research progress on the fifth-generation district heating system based on heat pump technology," *Journal of Building Engineering,* 71, 106533 (2023).
Katipamula, S. and M. R. Brambley, "Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I," *HVAC&R Research,* 11, 3-25 (2005).
Li, Y. and Z. O'Neill, "An innovative fault impact analysis framework for enhancing building operations," *Energy and Buildings,* 199, 311-331 (2019).
Li, Y. and Z. O'Neill, "A critical review of fault modeling of HVAC systems in buildings," *Building Simulation,* 11, 953-975 (2018).
Lu, X., Y. Fu, Z. O'Neill and J. Wen, "A holistic fault impact analysis of the high-performance sequences of operation for HVAC systems: Modelica-based case study in a medium-office building," *Energy and Buildings,* 252, 111448 (2021).
Lund, H., N. Duic, P. A. Østergaard and B. V. Mathiesen, "Future district heating systems and technologies: On the role of smart energy systems and 4th generation district heating," *Energy,* 165, 614-619 (2018).
Lygnerud, K., T. Nyberg, A. Nilsson, A. Fabre, P. Stabat, C. Duchayne and V. Gavan, "A study on how efficient measures for secondary district heating system performance can be encouraged by motivational tariffs," *Energy, Sustainability and Society,* 13, 38 (2023).
Majidi, M., B. Mohammadi-Ivatloo and A. Soroudi, "Application of information gap decision theory in practical energy problems: A comprehensive review," *Applied Energy,* 249, 157-165 (2019).
Månsson, S., P.-O. Johansson Kallioniemi, M. Thern, T. Van Oevelen and K. Sernhed, "Faults in district heating customer installations and ways to approach them: Experiences from Swedish utilities," *Energy,* 180, 163-174 (2019).
Månsson, S., P.-O. J. Kallioniemi, K. Sernhed and M. Thern, "A machine learning approach to fault detection in district heating substations," *Energy Procedia,* 149, 226-235 (2018).
Monghasemi, N., A. Vadiee, K. Kyprianidis and E. Jalilzadehazhari, "Rank-Based Assessment of Grid-Connected Rooftop Solar Panel Deployments Considering Scenarios for a Postponed Installation," in *Energies*. vol. 16, ed., p.^pp. (2023).
Nelson, W. and C. Culp, "Machine Learning Methods for Automated Fault Detection and Diagnostics in Building Systems—A Review," in *Energies*. vol. 15, ed., p.^pp. (2022).
Neumayer, M., D. Stecher, S. Grimm, A. Maier, D. Bücker and J. Schmidt, "Fault and anomaly detection in district heating substations: A survey on methodology and data sets," *Energy,* 276, 127569 (2023).
Qin, J. and S. Wang, "A fault detection and diagnosis strategy of VAV air-conditioning systems for improved energy and control performances," *Energy and Buildings,* 37, 1035-1048 (2005).
Sorknæs, P., P. A. Østergaard, J. Z. Thellufsen, H. Lund, S. Nielsen, S. Djørup and K. Sperling, "The benefits of 4th generation district heating in a 100% renewable energy system," *Energy,* 213, 119030 (2020).
Sveby, "Brukarindata kontor: Version 1.1," *Indatasammanställning*, Stockholm, Sweden (2013).
Yan, R., Z. Ma, G. Kokogiannakis and Y. Zhao, "A sensor fault detection strategy for air handling units using cluster analysis," *Automation in Construction,* 70, 77-88 (2016).

ACKNOWLEDGEMENT

The authors extend their sincere appreciation to the project developers affiliated with the NEEDs—innovative approaches for energy efficiency improvement for positive or self-balanced districts (project No. 52686-1). By sharing their domain-specific experiences and insights, they have indirectly contributed to shaping this paper and its findings. An acknowledgement can be located at the end of the text to indicate the sponsor of the study presented and/or to acknowledge additional contributors to the paper.

 https://doi.org/10.52202/077185-0213