Introducing Causality to Symptom Baseline Estimation
A Critical Case Study in Fault Detection of Building Energy Systems

Lu, C.J.; Wang, Z.; Mosteiro-Romero, Martín; Itard, L.C.M.

Publication date
2024

Document Version
Final published version

Published in
ASim2024, The 5th Asia Conference of the IBPSA

Citation (APA)
Lu, C. J., Wang, Z., Mosteiro-Romero, M., & Itard, L. C. M. (2024). Introducing Causality to Symptom Baseline Estimation: A Critical Case Study in Fault Detection of Building Energy Systems. In *ASim2024, The 5th Asia Conference of the IBPSA* (pp. 1621-1628). IBPSA.

Introducing Causality to Symptom Baseline Estimation: A Critical Case Study in Fault Detection of Building Energy Systems

C. Lu[1,*], Z. Wang[1], M. Mosteiro Romero[1], L. Itard[1]

[1]Department of Architectural Engineering and Technology, Delft University of Technology, Delft, The Netherlands.

ABSTRACT

Fault detection and diagnosis (FDD) provides several interrelated benefits, including reducing energy waste, enhanced operational efficiency, and maintaining indoor comfort. The initial step in FDD is to detect deviations from normal or expected operation. However, establishing a reliable baseline can be challenging, especially when there is a lack of sufficient system documents or when complex control strategies are involved. This study investigates three feature selection methods for the baseline estimation: expert knowledge-based, correlation-based, and causality-guided, using heating coil valve control estimation as an example. These methods were tested in an office building in the Netherlands. The results show that while the correlation-based method achieved the best estimation, it may lead to false negatives due to features with reverse causality. This study aims to emphasize the necessity of causal analysis in the baseline estimation to achieve reliable FDD in buildings.

KEYWORDS

Feature selection, Causal effect, Fault detection, Air handling unit, Building energy systems

INTRODUCTION

In the European Union, buildings account for around 40% of energy consumption and over one-third of energy-related greenhouse gas emissions. To support the heightened climate objectives of the European Green Deal, the updated Energy Performance of Buildings Directive sets a target of reducing emissions in the building sector by at least 60% by 2030 compared to 2015 levels, with the goal of achieving climate neutrality by 2050. Reducing energy waste in buildings is key to meeting these targets. Heating, ventilation, and air conditioning (HVAC) systems are the main energy consumers in buildings and are prone to a range of faults involving sensors, mechanical components, and control systems. These faults can result in uncomfortable indoor conditions, poor air quality, and significant energy waste. Therefore, developing automatic fault detection and diagnosis (FDD) algorithms for

[*] Corresponding author email: c.j.lu@tudelft.nl

building energy systems is vital to minimizing energy waste and reducing maintenance costs.

Over the past few decades, numerous FDD tools have been developed for building energy systems, which can be broadly categorized into knowledge-based and data-driven approaches (Zhao et al. 2019). While data-driven models have gained significant attention, a major challenge is their need for large amounts of high-quality labelled faulty data for training, which is often difficult to acquire in practice (Xiao et al. 2014). In contrast, knowledge-based approaches rely on the physics rules or engineering knowledge with the advantages of trustworthiness and interpretability. These approaches typically begin by identifying symptoms, or "evidence," through comparisons of real-time operations or energy consumption against a predefined baseline. When the deviation exceeds a certain threshold, the symptom is detected, allowing the diagnosis of related faults. For instance, according to the literature (Zhao et al. 2015, Dey and Dong, 2016, Wang et al. 2024), the symptom "estimated control signal of heating coil valve (HCV) versus actual control signal" is a key symptom of several faults in air handling units (AHUs), as shown in Fig.1.

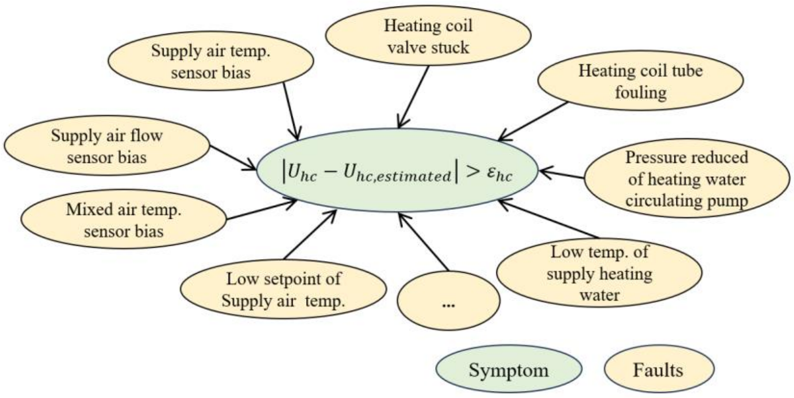

Figure 1. *Realted faults to the symptom "estimated control signal of HCV versus actual control signal" in AHUs, where $U_{hc,etimated}$ represnts the estimated HCV control signal, U_{hc} represnts the actual HCV control signal, and ε_{hc} is the threshold.*

The predefined baselines represent the normal operating conditions of building energy systems and can be established from various sources, such as documentation or fault-free data. For instance, Taal and Itard (2020) set the baselines for the efficiency of a heat exchanger and the efficiency of the thermal energy regeneration of an aquifer thermal energy storage system at 87% and 100%, respectively, according to design specifications and Dutch regulations. Van Koetsveld van Ankeren (2024) set the baselines based on the design document for control faults diagnosis in AHUs. However, not all baselines can be established from documentation, as some systems may lack sufficient system documents or involve complex control strategies that cannot be simply described. In such cases, alternative methods, such as statistical analysis or machine learning (ML) models based on historical fault-free data, may be required to estimate the baseline. For instance, Zhao et al. (2015, 2017) used polynomial functions to estimate many operation baselines in AHUs, including the control signal of heating and cooling coil valves, the energy consumption of fans, and the differential pressure across the filter. Chitkara (2022) and Gunderi (2022) utilized eXtreme Gradient Boosting (XGBoost) model, to estimate coil valve control signals for the FDD task in AHUs. Chen et al. (2022) proposed a weather and schedule-based pattern matching method to automatically generate the baseline for cross-level fault diagnosis in the HVAC systems.

Selecting the appropriate features is crucial for accurate baseline estimation in building energy systems. With extensive sensing capabilities in modern buildings, the initial datasets are often high-dimensional, not all the features are conducive to the baseline estimation. A compact and informative subset of features can significantly reduce complexity and enhance the performance of data-driven models (Lu et al. 2022). However, previous studies have typically relied on either expert knowledge or correlation-based feature selection. Expert knowledge may not always recognize all relevant features, causing inaccurate estimation. Correlation-based feature selection neglects the causal relationships between features and the target variable (Yu et al. 2020). This oversight can result in "reverse causality," potentially causing misdiagnoses. When the selected features are causally linked to the target variable, data-driven baseline estimations will undoubtedly become more reasonable and robust.

This study focuses on HCV control signal baseline estimation as a pioneering case study. We compare three feature selection methods: expert knowledge-based, correlation-based, and causality-guided feature selection. The goal is to illustrate the limitations of the popular correlation-based feature selection and to emphasize the importance of incorporating causality into the feature selection process for more robust baseline estimation in FDD task of building energy systems.

METHODOLOGY

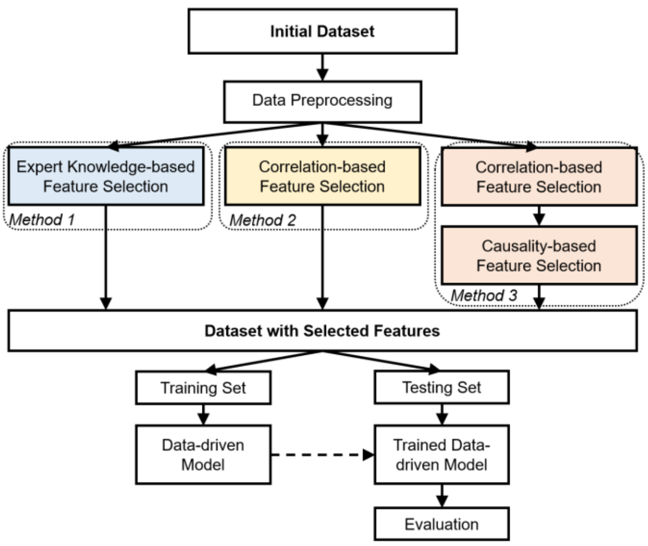

Figure 2. *Flowchart for symptom baseline estimation in this study.*

In this paper, we investigate the impact of feature selection on symptom baseline estimation and the subsequent FDD task. The flowchart for symptom baseline estimation of this study is illustrated in Fig. 2. The first step is data preprocessing, which primarily involves examining the initial dataset, removing features with constant values or missing data. Next, three feature selection methods are applied, including expert knowledge-based feature selection, correlation-based feature selection, and the proposed causality-guided feature selection. Finally, data-driven models are used to estimate the symptom baseline based on the dataset with the selected features from each method.

Case Study

This study utilizes a building manage system (BMS) dataset from the Kropman office building in Breda, the Netherlands, which is a living lab of the Brains4Building project. The building has an approximate floor area of 1500 m^2. This study mainly focuses on the central AHU system under heating mode, equipped with one heat recovery wheel (HRW) and heating coil, distributing three zones (four rooms). The simplified scheme of the HVAC system in the living lab is illustrated in Fig. 3. The HCV openness in the AHU is controlled to maintain the supply air temperature at its set point, which is determined by the room temperature. The rooms are also equipped with radiators, meaning the AHU primarily functions to heat the incoming fresh air.

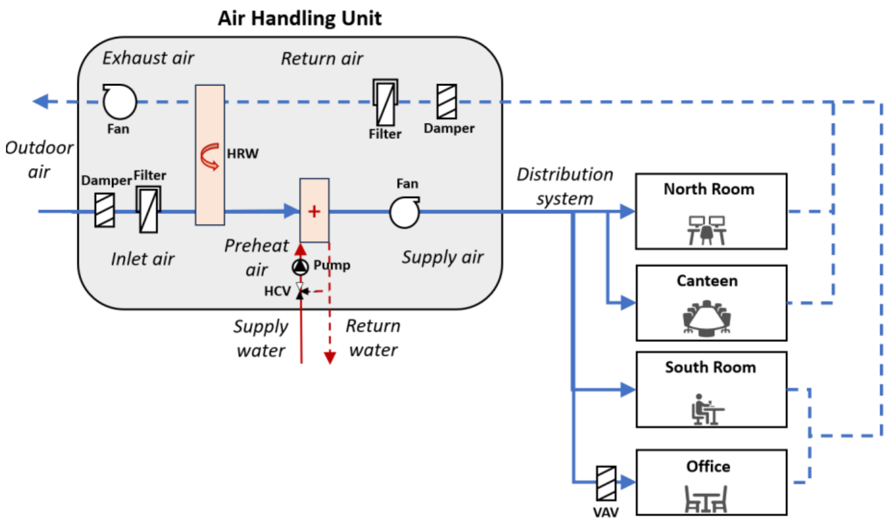

Figure 3. *Simplified scheme of the HVAC system in the living lab.*

The BMS platform is InsiteView-BMS, which coordinates sensor-based measurements, actuators, and monitoring data across all operational levels in the building to enable effective control. The initial dataset comprises 55 features. We select data from the working hours (9 AM to 16 PM) over two weeks (ten working days, from January 23rd to February 3rd in 2023). After examining the dataset and removing features with constant values or missing data, 46 features remain for further analysis.

Three feature selection methods for HCV control signal baseline estimation are applied in this study, described as follows.

- Method 1: Expert knowledge-based feature selection. The method selects the features based on domain knowledge, including heat and mass transfer analysis, control scheme analysis, and installation experience.
- Method 2: Correlation-based feature selection. The method selects the features based on the association strength with the target variable, which is also known as the filtering method. There are some criteria to evaluate the association strength, such as Pearson correlation coefficient (PCC) and mutual information (MI). PCC is widely used to measure the linear correlation between variables. MI can measure not only the linear but also the nonlinear correlation. In this study, a greedy forward selection strategy is applied (Min H and Ren W 2015), using MI as the criterion to identify the optimal subset of 10 features for estimating the baseline of the HCV control signal.
- Method 3: Causality-guided feature selection. The method introduces causality analysis to examine the features selected by Method 2. Features exhibiting reverse causality need to be excluded. Reverse causality occurs when the direction of cause and effect between two variables is misunderstood or misrepresented. In the baseline estimation for FDD tasks, only the cause variables of the target variable should be included, while effect variables should be excluded to ensure accurate predictions. In this study, the casual relationships are examined by expert knowledge.

In this study, we utilize linear regression to estimate the HCV control signal baseline using the features selected by Method 1 (Zhao et al. 2015) and employ XGBoost to estimate the baseline using the features selected by Method 2&3 (Chitkara 2022 and Gunderi 2022). For training, we implement 10-fold cross-validation to ensure model robustness. Additionally, the grid search of three key hyperparameters is conducted to optimize XGBoost, including "n_estimators", "learning_rate", and "max_depth". Finally, R^2 and Root Mean Square Error *(RMSE*) are used as the evaluation metrics. The testing set consists of one fault-free day and one faulty day, allowing us to assess the model performance under both normal and fault conditions.

RESULTS

Table 1 presents the results of feature selection across three methods. For Method 1, referred to Zhao et al. (2015), the preheat air temperature (T_{pa}) and the supply air flow rate (F_{sa}) are selected. In Method 2, the optimal subset includes supply water temperature (T_{sw}), supply air temperature setpoint (T_{set}), absolute humidity at supply air distribution system (AH_{sad}), outdoor air temperature (T_{oa}), supply air temperature (T_{sa}), related humidity at supply air distribution system (RH_{sad}), related humidity of return air (RH_{ra}), exhaust air temperature (T_{ea}), related humidity of supply air (RH_{sa}), and inlet air temperature (T_{ia}). In Method 3, T_{sw}, T_{sa}, RH_{sad}, and RH_{sa} are excluded from the optimal subset due to concerns about reverse causality. Even though these features are highly related to the HCV openness, they are directly affected by the

HCV openness. HCV openness influences them rather than the other way around. Including them in the model would introduce reverse causality, where the predictors are the effects of the target variable (HCV control signal), leading to biased estimates. To avoid this, they are excluded from the feature selection process to ensure the model accurately captures the true causal relationships.

Table 1. *Feature selection results*

Methods	Selected Features
1	T_{pa}, F_{sa}
2	T_{sw}, T_{set}, AH_{sad}, T_{oa}, T_{sa}, RH_{sad}, RH_{ra}, T_{ea}, RH_{sa}, T_{ia}
3	T_{set}, AH_{sad}, T_{oa}, RH_{ra}, T_{ea}, T_{ia}

Table 2 presents the evaluation results for HCV control signal estimation based on the features selected by the three methods. Figure 4 illustrates the HCV control signal estimation on a fault-free day. Method 1 clearly performs poorly, likely due to the limited number of selected features and the limitations of linear regression. This poor performance may lead to false positive detection of symptoms, potentially triggering incorrect fault diagnoses. Method 2 achieves the best performance using the optimal feature subset. The performance of Method 3 slightly decreases due to the exclusion of features with reverse causality.

Table 2. *Evaluation results of HCV control signal estimation*

Methods	Training		Testing	
	R^2	RMSE	R^2	RMSE
1	0.38	7.8	0.23	14.2
2	0.81	4.3	0.61	10.1
3	0.63	6.0	0.58	10.5

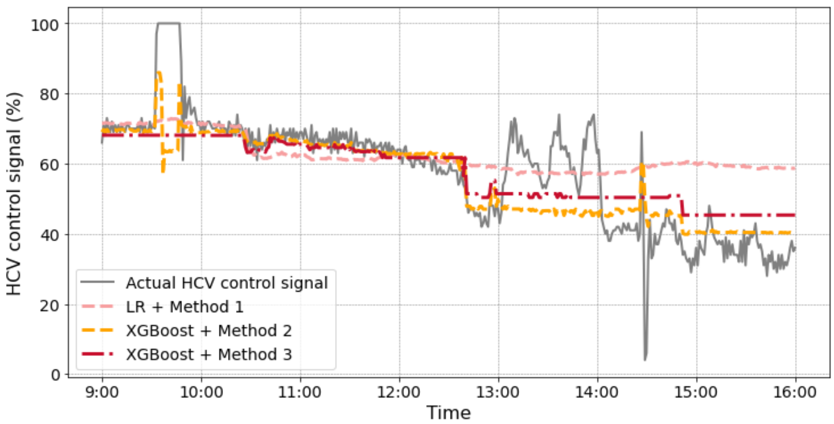

Figure 4. *HCV control signal estimation on the testing set (a faul-free day).*

Figure 5 presents the HCV control signal estimation on a faulty day. Around 10 AM, the HCV control signal becomes stuck at 10%, and from approximately 11 AM onwards, it remains stuck at 40%. The estimation using Method 2 remains the most accurate, as it closely follows the changes in the faulty signal, contrary to the expectation that it should act as a reliable baseline. This behavior indicates that

Method 2 is not reliable for fault detection, as it may lead to false negatives, where actual faults go undetected. In contrast, Methods 1 and 3 exclude features affected by reverse causality, making their estimations relatively unaffected by these changes.

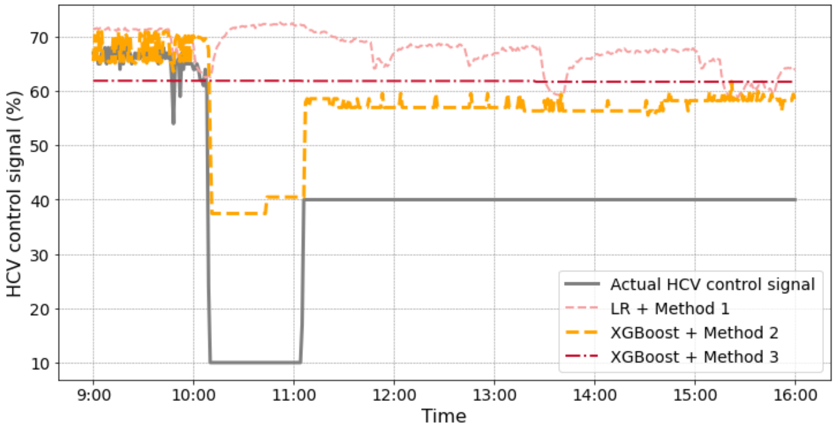

Figure 5. *HCV control signal estimation on a fauly day.*

CONCLUSION AND IMPLICATIONS

In this study, we analyzed three feature selection methods for estimating the HCV control signal baseline. The results showed that the expert knowledge-based method performed poorly due to the limited number of selected features and linear regression constraints, might causing false positive detection. The correlation-based method achieved the best baseline estimation performance, but reverse causality issues could result in false negatives during symptom detection. The causality-guided method, which excluded features with reverse causality, ensured a more reliable baseline for fault detection, albeit with slightly reduced estimation accuracy. The results emphasize the necessity of causal analysis in the feature selection for the symptom baseline estimation to avoid biased fault diagnosis. In the future, advanced causality discovery algorithms could be applied to better determine causal relationships and further improve both the accuracy and reliability of baseline estimation.

ACKNOWLEDGEMENTS

This study is supported by the Brains4Buildings project, sponsored by the Dutch grant program for Mission-Driven Research, Development, and Innovation (Grant number: MOOI32004). The authors would like to thank Srinivasan Gopalan, Karzan Mohammed, and Dr. Rick Kramer from TU Eindhoven for their efforts in data collection and collaboration. Special thanks are also extended to Dr. Shalika Walker, John Verlaan, and other engineers from Kropman B.V. for providing the living lab.

REFERENCES

Chen Y, Wen J, Pradhan O, Lo J, and Wu T. 2022. Using discrete Bayesian networks for diagnosing and isolating cross-level faults in HVAC systems, *Applied Energy*. 327 : 120050.

Chitkara, S. 2022. Continuous Monitoring and Automated Fault Detection and Diagnosis of Large Air-Handling Units, *P.D.Eng. Thesis*, Eindhoven University of Technology (The Netherlands).

Dey D and Dong B. 2016. A probabilistic approach to diagnose faults of air handling units in buildings, *Energy and Buildings*. 130 : 177-187.

Gunderi K. 2022. Fault detection and diagnosis of low delta-T syndrome in air handling unit cooling coils, *P.D.Eng. Thesis*, Eindhoven University of Technology (The Netherlands).

Lu C, Li S, and Lu Z. 2022. Building energy prediction using artificial neural networks: A literature survey, *Energy and Buildings*. 262 : 111718.

Min H and Ren W. 2015. Global mutual information-based feature selection approach using single-objective and multi-objective optimization. *Neurocomputing*. 168 : 47-54.

Taal A and Itard L. 2020. P&ID-based automated fault identification for energy performance diagnosis in HVAC systems: 4S3F method, development of DBN models and application to an ATES system, *Energy and Buildings*. 224 : 110289.

Van Koetsveld van Ankeren L. 2024. Real-time Monitoring of Building Energy Systems Bayesian Network-based Fault Detection and Diagnosis of an Air-handling Unit in a Dutch University Building, *Master's Thesis*, Delft University of Technology (The Netherlands).

Wang Z, Lu C, Taal A, Gopalan S, Mohammed K, Meijer A, and Itard L. 2024. Bayesian network-based fault detection and diagnosis of heating components in heat recovery ventilation, *Proceedings of the 17th RoomVent conference*, Stockholm: RoomVent 2024.

Xiao F, Zhao Y, Wen J, and Wang S. 2014. Bayesian network based FDD strategy for variable air volume terminals, *Automation in Construction*. 41 : 106-118.

Yu K, Guo X, Liu L, Li J, Wang H, Ling Z, and Wu X. 2020. Causality-based feature selection: methods and evaluations, *ACM Computing Surveys*. 53(5) : 111.

Zhao Y, Wen J, and Wang S. 2015. Diagnostic Bayesian networks for diagnosing air handling units faults - Part II: Faults in coils and sensors, *Applied Thermal Engineering*. 90 : 145-157.

Zhao Y, Wen J, Xiao F, Yang X, and Wang S. 2017. Diagnostic Bayesian networks for diagnosing air handling units faults – Part I: Faults in dampers, fans, filters and sensors, *Applied Thermal Engineering*. 111 : 1272-1286.

Zhao Y, Li T, Zhang X, and Zhang C. 2019. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future, *Renewable and Sustainable Energy Reviews*. 109 : 85-101.