Energy & Buildings 263 (2022) 112041

Contents lists available at ScienceDirect

Energy & Buildings

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

A simulation-based evaluation of fan coil unit fault effects

Yimin Chen [a,*], Guanjing Lin [a], Zhelun Chen [b], Jin Wen [b], Jessica Granderson [a]

[a] Building Technology and Urban Systems Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[b] Department of Civil, Architectural and Environmental Engineering, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 19 January 2022
Revised 3 March 2022
Accepted 20 March 2022
Available online 22 March 2022

Keywords:
Fault symptom evaluation
Fault effects
Symptom occurrence probability
Symptom intensity
Fan coil unit
HVACSIM+

ABSTRACT

Faults in heating, ventilation and air conditioning (HVAC) systems cause increased energy consumption, degrading thermal comforts, growing operational cost and reduced system lifespan. An effective evaluation of fault effects is critical to the development of various fault diagnostics solutions, the improvement of operation maintenance and the optimization of monitoring systems. In the HVAC area, a majority of research work in evaluating fault effects was to analyze energy consumption impacts or thermal comfort impacts. However, a handful of research has been conducted on evaluating fault effects on various measurements, which are increasingly employed to monitor equipment's operation. Fault effects on various measurements may display different symptom patterns and present changed sensitivities when the equipment operates under various faults, severity levels, as well as operation conditions. However, a long-term observation of fault symptoms under various operation conditions, different fault types and severity levels to evaluate fault effects is extremely challenging. In this paper, a simulation-based framework was proposed to evaluate fault effects in fan coil units (FCUs). Two metrics namely fault symptom occurrence probability (SOP) and fault symptom daily continuous duration (SDCD) were developed to quantify fault symptoms under various FCU faults. A total of 18 common FCU faults at different severity levels were implemented on the developed HVACSIM+ simulation platform to obtain a full year fault inclusive data set for 48 fault simulation cases. The framework, as well as obtained SOP and SDCD distributions will benefit multiple folds such as the development of probability-based fault diagnostics inference approaches, optimization of sensor location, and fault prioritization.

1. Introduction

In buildings, heating, ventilation and air conditioning (HVAC) systems are critical to maintain zone thermal comforts and desired air quality for occupants. However, numerous HVAC systems operate under faulty conditions which cause increased energy consumption, deteriorated zone thermal comfort, decreased air quality, as well as increased maintenance cost and reduced system lifespan [1]. For example, studies show that 15–30% energy consumption is wasted due to faults, malfunctioning and degrading equipment as well as poor maintenance in HVAC systems in commercial buildings [2]. To enhance a reliable HVAC system operation and avoid energy wastes, comprehensive works have been conducted to develop various automated fault detection and diagnostics (FDD) approaches in the past thirty decades [3]. A study shows that the average 8% energy consumption can be saved after applying FDD solutions in commercial buildings [4].

Another research area related to HVAC system faults is to evaluate fault effects. Fault symptoms and impacts are the effects of a fault on various measurements, components, control objectives and operation performance. Fault effects reflect the undesired or unpermitted operating states of a system, i.e., when a fault occurs, equipment operation experiences abnormal changes (discrepancies) compared to the normal operation, and generates residuals which are reflected by either direct measurements (e.g., sensor readings and control signals), or indirect measurements (e.g., rules obtained from comparing multiple measurements). A complete and systematic evaluation on fault effects may benefit many facets. First, the fault effects evaluation enables an effective fault prioritization which would benefit various aspects such as efficient fault correction and system maintenance. For instance, after evaluating the service costs of various faults in rooftop units (RTUs), Breuker et al. ranked RTU faults to facilitate RTU maintenance [5]. Secondly, the faults and effects evaluation has been widely used in developing FDD approaches. The observable fault symptoms on various measurements were employed by system operators to determine system operational abnormalities. This heuristic process evolved

* Corresponding author.
E-mail address: YiminChen@lbl.gov (Y. Chen).

https://doi.org/10.1016/j.enbuild.2022.112041
0378-7788/

Nomenclature

Symbols

ε	Difference value
y_o	Observed value of a measurement
y_{ref}	Nominal reference value of a measurement
y_n	Normal value of a measurement
μ	Mean value of samples
σ	Standard deviation of samples

Abbreviations

FDD	Fault detection and diagnostics
OAT	Outdoor air temperature
RM_TEMP	Zone air temperature
MAT	Mixed air temperature
DAT	Discharge air temperature
RAT	Return air temperature
CVLV_DM	Cooling coil valve control signal
CLG_GPM	Cooling coil water flow rate
CLG_RWT	Cooling coil return water temperature
HVLV_DM	Heating coil valve control signal
HTG_GPM	Heating coil water flow rate
HTG_RWT	Heating coil water return rate
DA_CFM	Discharge air flow rate
OA_CFM	Outdoor air flow rate
DMPR_DM	Outdoor air damper control command signal
SPD	Fan speed
MA_HUMD	Mixed air relative humidity
DA_HUMD	Discharge air relative humidity
RA_HUMD	Return air relative humidity
SOP	Symptom occurrence probability
SDCD	Symptom daily continuous duration

to some fault diagnostics approaches such as rule-based fault diagnostics [6] and the expert system [7]. Additionally, some data-driven FDD approaches, which rely on building automation system (BAS) interval data collected from diverse measurements, may also need an accurate cause-and-effect analysis to facilitate the development of an inference model. For example, in a Bayesian Network (BN)-based fault diagnostic method, both qualitative and quantitative models need to be developed to represent fault cause-and-effect relations in different HVAC subsystems [8]. Lastly, the faults and effects evaluation can improve the design of a monitoring system by optimizing the sensor deployment and hence increase the monitoring efficiency [9].

Compared to other types of HVAC equipment such as chillers, air handling units (AHUs) or variable air volume (VAV) terminal units, there is a lack of efficient monitoring strategies for the operation of a fan coil unit (FCU). FCUs are simple and decentralized air-conditioning devices which are primarily used to locally condition the air in zones. Compared with other HVAC systems, FCUs can be easily and flexibly deployed in buildings where the space is limited to install ducts [10]. Therefore, FCUs are widely used in various types of buildings including offices, hotels, schools, as well as residential apartments in the U.S and in Europe. An evaluation on the effects under FCU faulty operation will significantly improve the monitoring performance of FCUs and facilitate the early detection of FCU faults.

This paper studies the effects of FCU faults utilizing simulated system operation data. It proposes a new evaluation framework to bridge the gap of a lack of systematic evaluation of fault effects on FCU measurements. The framework includes fault symptom characterization, baseline generation, and fault effects evaluation. Specifically, we quantify fault effects in terms of two metrics, namely symptom occurrence probability (SOP) and symptom daily continuous duration (SDCD), to represent the fault symptom occurrence likelihood and intensity respectively.

When evaluating fault effects, a long-term observation on the equipment faulty operation under multiple operating conditions, and under various fault severity levels is required. However, this may be very time-consuming and expensive in real practice. To address this challenge, we employ simulation data to study fault symptoms on various measurements which include sensor readings and control signals in FCUs. The simulation data, which are generated on the HVACSIM+ based FCU model, enable a thorough analysis on the system operation under different faults, fault severity levels and operating conditions. In addition, existing fault effect evaluation methods cannot fully assess impacts a fault has on system measurements, which may be employed to develop FDD approaches, optimize the measurement deployment and prioritize faults, as will be discussed in Section 2 in detail.

The proposed framework, as well as obtained SOP distributions and SDCD distributions on each measurement can be used for evaluating FCU fault effects, developing and validating fault models, as well as optimizing measurement deployment and prioritizing faults.

The research results presented in this paper are based on the HVAC system Fault Data Curation project which aims at building the largest HVAC system fault database in the world. The HVAC system fault data used in this research have been fully validated through the developed protocol to ensure the data quality. The following sections are arranged as: Section 2 reviews past works on the fault effects and impacts analysis. Section 3 presents the proposed method, as well as introduces the simulation process. Section 4 illustrates the process of evaluation and analysis, as well as discusses the applications of the developed method and results. Section 5 concludes the paper and proposes future work direction.

2. Related works

We illustrate some valuable studies which evaluated fault symptoms or impacts. Among those studies, data collected from laboratory experiment tests and simulation platforms were often employed. The assessment on fault impacts or symptoms were carried out through analyzing various measurements connecting to the BAS, or through analyzing different metrics such as energy consumption, operating or maintenance costs as well as occupants' comforts.

Among the laboratory experiment tests, several representative studies are reviewed as below. Comstock et al. investigated eight common faults under different cooling loads in a centrifugal chiller in a laboratory environment [11]. A total of 13 measurements were used to evaluate the measurement sensitivity under chiller's faulty operation. Breuker et al. investigated common faults and corresponding impacts on the rooftop units (RTUs) [5]. In the study, 96 fault tests at 4 load levels and 24 fault severity levels were performed via the experimental tests to evaluate fault impacts on the transient profiles of nine performance indices. The authors analyzed the fault symptoms on the transient profiles. The quantitative changes in RTU cooling capacity, coefficient performance, and two temperature measurements were analyzed. Although the authors concluded some generic rules that described the fault impact directions on various temperature measurements, they did not quantify the symptoms on five temperature measurements.

Cho et al. carried out transient pattern analysis for HVAC FDD [12]. In the study, four types of faults in the VAV HVAC system were imposed in a real test chamber to obtain system faulty operation data. The study concluded the evolution of fault residuals to form patterns which can be used to isolate faults. Additionally, the authors found the temporal patterns in multiple components caused when a fault occurs and recovers to steady state. However, the authors did not report if the patterns will be affected by system operating conditions. Schein et al. extracted 28 rules for AHUs named AHU performance assessment rules (APAR) from the observation of fault symptoms on temperature measurements and normalized control signals [13]. The developed APAR has been widely used by market FDD solutions. Although the authors gave the thresholds for each rule as the judgment on whether the symptom can be observed and the rule is violated, the authors did not provide the quantifiable thresholds under different operating conditions.

Compared with hardware experiments, simulation platforms are more efficient to evaluate HVAC system fault impacts. For example, Chen et al. employed both EnergyPlus and Modelica tools to develop a single duct VAV system model to analyze fault impact [14]. In the study, the simulation scenarios were selected from two aspects of evaluations, quantitative long-term (week/month/year) impacts, as well as chronological short-time (within hours) dynamic impacts, which can be also used to generate a fault onset data set for FDD method testing. In addition, the study also reported relations between some physical faults and control logic sequence and the seasonal operating conditions. However, only fault impacts on energy consumption were quantified and only a few days in each season were used to evaluate fault impacts. Shi et al. developed three steps to evaluate fault impacts using EnergyPlus-based building performance simulation (BPS) to address the challenge of quantitatively translating the symptoms caused by a fault to specific inputs inside a BPS model [15]. However, the authors only used mean values to quantify directly observable symptoms, and used three quartiles (i.e., 25%, 50% and 75%) to approximate indirect estimated symptoms. Li et al. proposed a fault impact analysis framework by incorporating the fault model library with the EnergyPlus simulation tool [16]. In the study, 129 fault modes from 41 groups of fault models were simulated from the medium sized office case. Fault occurrence probability models were integrated into the simulation platform to more accurately evaluate the magnitude of fault impacts in buildings in different climate zones. However, the framework only focused on site energy impact and HVAC energy impact within one year scope, and no analysis on fault symptoms on various measurements was performed.

In summary, the above studies have investigated a fault's effects from many angles. However, the following two major gaps, which prevent a complete understanding of a fault's effects as, still exist:

1) When evaluating fault effects or impacts, most research works have been conducted on evaluating final or long-term fault impacts such as annual energy consumption, thermal comforts and operating costs [16–18]. However, little research has been conducted on quantitatively evaluating fault effects on system measurements which are often presented as fault symptoms and are commonly used to assess a system's dynamic operation, as well as to develop FDD approaches.
2) Some studies have used trend data comparison to visualize the fault symptoms on measurements under one operating condition [11,14,19]. However, fault symptoms may be sensitive to various operating conditions such as control sequences, weather conditions, as well as fault severity levels. The uncertainties of observable fault symptoms were never investigated.

To address the above issues, in the proposed simulation-based fault effect evaluation framework, we focused the evaluation on fault symptoms, and quantified fault symptoms on various measurements. Two new metrics (i.e., SOP and SDCD), which can evaluate fault symptom occurrence frequency and intensity on various measurements, will be illustrated in detail in this paper.

3. Methodology

Fig. 1 shows the framework of the methodology as will be illustrated in the following sections. We first illustrate the characteristics of a fault symptom in HVAC system operation in Section 3.1. In Section 3.2, we introduce the method of baseline data generation. Then, we illustrate two metrics that quantify an observable fault symptom as fault symptom occurrence probability in Section 3.3 and fault symptom duration in Section 3.4. Lastly, we present how the FCU faults are simulated, as well as the fault inclusive and exclusive data set used for the fault effect analysis in Section 3.5.

3.1. Overview of fault symptoms

1) Introduction of fault symptom generation methods.

When a fault occurs, fault symptoms could be reflected as deviations of sensor readings or control signal from normal values on various measurements in the system [20]. In a control system, fault symptoms may present different patterns because the system's operation may be complicated. There are several studies discussing the fault symptom patterns [20–23]. For example, in [21] fault symptoms are classified into two categories: semantic symptoms and trend symptoms. A semantic symptom can be obtained by comparing the difference between a measurement's current value and its nominal reference value. This type of symptoms can become more notable and will continue to be observed for a period after a fault occurs. A trend symptom refers to the changing rate of a measurement value. This type of symptom is more significant at the initial phase when a fault occurs but is not observable after a time period when the system reaches another steady state.

In our research, semantic symptoms were analyzed. There are three methods to produce semantic symptoms. In the first method, the normal value is presented by using the nominal reference value (e.g., the temperature setpoint or rated fan speed) required by the control system as shown in Fig. 2 (a). When the difference between observed values and nominal values exceeds a certain threshold, the exceeding value could represent an observable symptom as given in Eq. (1). For example, the zone temperature setpoint of a FCU can be used to determine if abnormal zone temperature can be observed.

$$\varepsilon = y_o - y_{ref} \tag{1}$$

where y_o is the observed value of a measurement (e.g., zone temperature) in the system, y_{ref} is the nominal reference value the measurement (e.g., zone temperature setpoint), and ε is the difference between the nominal value and measured value.

In the second method, the normal value of each individual measurement is obtained through data collected during equipment's fault-free operation as shown in Fig. 2 (b). The difference between a measurement's current value and its normal value is calculated to produce fault symptoms as given in Eq. (2):

$$\varepsilon = y_o - y_n \tag{2}$$

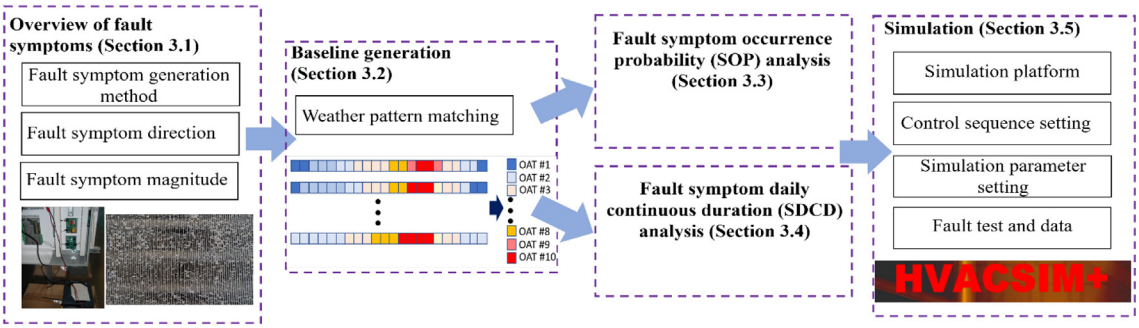

Fig. 1. Framework of the methodology.

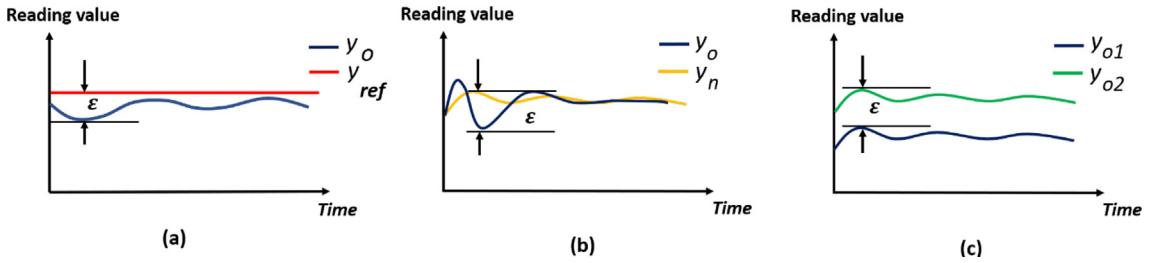

Fig. 2. Demonstration of fault symptom generation methods.

where y_o is the observed value of a measurement (e.g., measured discharge air temperature value from the sensor) in the system, y_n is the normal value of the measurement (e.g., calculated discharge air temperature mean value under equipment's fault-free operation), and ε is the difference between the normal value and measured value.

In the third method, the symptom can be produced by comparing concurrent values collected from two or more different measurements in the system as shown in Fig. 2 (c). The difference between current values collected by different measurements can be calculated according to certain rules as given in Eq. (3). For example, in AHU performance assessment rules [13], the mixed air temperature (MAT) in the AHU is compared with the outdoor air temperature (OAT) to determine if a fault symptom occurs and to indicate an outdoor air damper stuck fault.

$$\varepsilon = f(y_{o1}, y_{o2}, \cdots, y_{oj}) \tag{3}$$

where y_{oj} is j^{th} the observed value of a measurement in the system (e.g., MAT or OAT), and ε is the difference among various measured data sets (e.g., the difference between MAT and OAT).

In this study, we employed the first and second method to produce fault symptoms and evaluate fault effects on various measurements. For the first method, both the zone temperature cooling setpoint and heating setpoint are used as the reference value. For the second method, the normal value of the measurement y_n is calculated from the normalized baseline data by using the z-score method (assuming the distributions of deviations on each measurement to be normal distributed), which has been mostly employed by data-driven methods as:

$$Z = \frac{y_o - \mu}{\sigma}$$

After normalizing the baseline data, the mean value μ and standard deviation σ of each measurement can be obtained as:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2}$$

where y_i is the i^{th} observation and n is the number of observations.

Therefore, the symptom can be obtained when the absolute difference between each observation and the mean value of measurement is higher than the standard deviation as given in Eq. (4).

$$|y_o(i) - \mu| > t \times \sigma \tag{4}$$

where y_o is the observed time series data, t is the threshold value (e.g., 1, 2, ...). In this study, t is set to one as the threshold (with a 68% confidence level) despite the "three-sigma empirical rule" has been often adopted and the value of three standard deviations (with a 99.7% confidence level) has often been used. This is because a lower threshold means a smaller deviation of the measured value can be captured, and then can be determined as an observed symptom event. Therefore, a lower threshold may increase the sensitivity of a measurement considering the HVAC system faults (especially at a mild severity level) may not generate significant measured deviations on measurements.

The method of the baseline data generation will be illustrated in Section 3.2.

2) Fault symptom direction.

A symptom's direction can be labeled as a positive direction or a negative direction to represent the direction of a measured value compared to the baseline value, i.e., a difference (ε) is higher than the baseline value or lower than the baseline value respectively. It is noted that for the same fault, the fault symptom direction of a measurement can be different due to various operating conditions. For example, if an outdoor air damper is stuck at a higher position in a FCU, the mixed air temperature could be higher than the baseline when the OAT is high (e.g., in the summer season), and be lower than the baseline when the OAT is low (e.g., in the winter season). In addition, the MAT may not present a symptom when the OAT is close to the MAT. Therefore, the correct identification of a fault symptom direction under specific operating conditions

should be critical for the FDD process. Otherwise, error FDD results may be triggered when applying certain rules without considering the symptom direction. In this study, a sign ("+" or "-") is associated with the fault symptom's direction to represent positive or negative deviation of the observed value.

3) Fault symptom magnitude.

The symptom magnitude can be presented by using a qualitative description or a quantitative description. In the qualitative description, the symptom magnitude is obtained through heuristic analysis of system operation from building operators' observation [20]. Through this way, fault symptom severity levels can be qualitatively classified by using linguistics variables (e.g., small, medium and large) or by certain vague values. Although this qualitative description of the fault symptom magnitude is relatively obscure, these depictions of fault symptoms are widely used in FDD approaches because obtaining an accurate degree value of a fault symptom is usually difficult and unnecessary in many engineering practices. Therefore, the qualitative representation of the fault symptom magnitude could enable the development of some FDD approaches. For example, in [23], the fault symptom magnitudes were qualitatively described as trend data signature and classified into seven levels for fuzzy-logic based fault diagnostics in the chemical process industry. In the quantitative description, numeric values are used to quantify measurement sensitivities or fault symptom magnitudes. For instance, a sensitivity factor was defined from the measurement residual for the worst fault case and the maximum uncertainty for the specific measurement to evaluate chiller fault impacts [11]. In the study, the upper limit and lower limit of sensitivity factors in the selected measurements were given to demonstrate the fault impacts under various faults in chillers. Similarly, Dash et al. [23] evaluated the fault symptom intensity by calculating the relative sensitivities from each measurement and its threshold.

Intuitively, the more severe a fault is, the stronger a fault symptom (i.e., a higher magnitude level of a fault symptom) on a measurement would be. As a consequence, the fault symptoms could be more likely to be observed or captured by the FDD methods. However, in an HVAC system, fault symptom magnitude could also be affected by multiple factors such as weather conditions, system control sequence or internal load conditions as discussed in the Introduction Section. Therefore, more metrics need to be employed to accurately quantify fault symptom presence as well as the measurement sensitivity.

In this study, apart from the above-mentioned two characteristics of fault symptoms, we proposed two additional metrics to evaluate fault symptoms as illustrated in Section 3.3 and Section 3.4.

3.2. Baseline generation

To reliably produce observable fault symptoms, it is critical to generate the baseline data set from the fault-free data under various operation conditions (e.g., control sequence, weather conditions and zone load), which match operation patterns from the faulty data set. In addition, when evaluating the fault symptom occurrence frequency (as explained in Section 3.3), the symptom occurrence probability needs to be more accurately calculated by considering the baseline data distribution. In this study, a weather-based pattern matching (WPM) method, which was developed in a FDD approach, was employed to generate the baseline data from the fault-free data [24]. The WPM method was employed by using the OAT as an indicator to match FCU operation patterns and generate the baseline data. This is because OAT is one critical driver which affects building thermal load, HVAC equipment operation and energy consumption [25]. When performing the WPM method, the OAT during the system's one year operation period was first equally binned. Then, FCU fault-free operation data within the same binned OAT window was grouped to generate the baseline data.

The number of binned OAT windows affects the generation of the baseline data. In this study, the determination of the number of binned OAT windows included two considerations as 1) the FCU operation performance within each binned OAT window should be similar; and 2) the sample size within one binned OAT window should be large enough to reach a statistical significance. After evaluating the FCU's operation, the OAT during the FCU's one year operation period was equally binned to ten windows in this research to generate the baseline data. Consequently, ten baseline subsets were formed under each binned OAT window. Then the baseline data within each binned OAT window was normalized to obtain the mean value and the standard deviation for each measurement as illustrated in Section 3.1. The operation time ratio within each binned OAT window will be given in Section 4.1. Fault symptoms on each measurement will be generated by comparing the fault data and the baseline within each binned OAT window.

3.3. Fault symptom occurrence probability

Different measurements may have different sensitivities according to faults types, severity levels and operating statuses. In this study, the fault symptom occurrence probability (SOP) is proposed to quantify the symptom occurrence frequency, i.e., what is the likelihood that a fault symptom could present on a measurement when a fault occurs. The SOP can be calculated by counting the number of observations of fault symptoms during a range of time periods when a fault occurs.

In this study, two steps are employed to calculate the fault SOP. First, the SOP under each binned OAT window is calculated as given in Eq. (5).

$$P(fault\_sym|OP_{OAT}) = \frac{\sum num\_fault\_sym}{\sum OP\_time} \tag{5}$$

where *num_fault_sym* is the time when a fault symptom is observed, and OP_time is the total operational period within each binned OAT window.

Secondly, the total probability distribution of the fault symptom occurrence is calculated. There are multiple probability weighting approaches that can be used to calculate the total probability [26] In this study, we employ the Bayesian approach [27] to calculate the total probability distribution of a symptom under each fault type and one fault severity level as given in Eq. (6).

$$P(fault\_sym|OP_{OAT}) = \sum_{i}^{num\_bin\_window} P(fault\_sym|OP_{OAT})_i P(OP_{OAT})_i \tag{6}$$

where $P(fault\_sym|OP_{OAT})_i$ is sSOP the under the i^{th} binned OAT window as given in Eq. (5), $P(OP_{OAT})_i$ is the operating ratio of the i^{th} binned OAT window under all operating time, *num_bin_window* is the total number of binned OAT windows.

Then, the range of the fault SOP under various fault severity levels can be obtained for each type of fault as illustrated in Section 4.

3.4. Fault symptom daily continuous duration

Fault symptom duration is the time period of a fault that may affect the measurement of a sensor in a dynamic control system [9]. The analysis of fault symptom duration is critical to identify symptom patterns and can be used for multiple applications such as the evaluation of fault detectability, sensor location optimization and fault isolation [9,28–30]. In this study, the fault symptom

daily continuous duration (SDCD) is proposed to evaluate a fault symptom duration in FCUs. The SDCD represents whether a fault symptom could be constantly present during an operating day. A higher SDCD that a fault triggers, the measurement can solidly generate a symptom without being affected by various operating conditions. For example, if the outdoor air damper of a FCU is stuck at a certain position, the outdoor air flow, which is measured by the outdoor air flow sensor, should be continuously present as a symptom during the operating day no matter what the operating condition is. However, the discharge air temperature measured by the discharge air temperature sensor may not present continuous abnormality as discharge air temperature is affected by the control of the cooling coil valve or heating coil valve.

The SDCD is calculated as given in Eq. (7).

$$SDCD = Max(Contiuous\_OP\_time\_daily) \tag{7}$$

where *Contiuous_OP_time_daily* is the continuous sample of a fault symptom collected in an operating day.

The overall SDCD for a specific measurement can be obtained by evaluating the whole year operation of an equipment, i.e., the percentage of the SDCD in a year as given in Eq. (8).

$$PctSDCD = \frac{\sum_i^n SDCD_{thresh\_i}}{\sum_{\Delta t}^{operating\_day} \Delta t} \tag{8}$$

where $\sum_i^n SDCD_{thresh}$ is the number of days that the SDCD is more than a certain time threshold (e.g., 60 min in this study), and *operating_day* is the total number of operating days.

3.5. Description of simulation

In this section, we illustrate how the fault simulation platform was set up and what control sequences and parameters were applied to operate the FCU.

3.5.1. Description of the simulation platform

A FCU (as illustrated in the left part in Fig. 3) is a common terminal equipment to condition zones in residential and commercial buildings in the U.S. and Europe. In this study, we employed the FCU model which was originally developed as a tool for evaluating FDD approaches [19]. A vertical four pipe hydronic FCU was modeled through the HVACSIM+ software [31]. Compared with other fault impact studies which often used EnergyPlus, the FCU fault model developed on the HVACSIM+ software tool has several advantages as: 1) the platform includes more detailed dynamic component models such as the damper model and the valve model; 2) the platform allows the HVAC and control systems to be simulated with a much finer time step (as low as 2.5 s), so that the dynamic operating performance of the equipment can be captured more accurately; and 3) various measurements can be easily modeled and embedded in the simulation platform to provide complete measures to evaluate equipment's dynamic operation.

In this study, the FCU model includes a fan that operates at three speed levels as high, medium and low. The FCU is controlled to maintain zone air temperature to the thermostat heating and cooling setpoints. The equipment physical configuration schematic and measurements are illustrated in the right part of Fig. 3. In this schematic, the various types of measurements are color labeled such that the red color text represents the sensor reading, and the blue color text represents the control signal.

In this study, a total of 17 measurements including 14 sensors and three control signals were used to monitor and control the FCU's operation. The measurements included in the dataset are summarized in Table 1. It is noted that some measurements (e.g., cooling coil entering water temperature and cooling coil returning water temperature) are not often deployed in real practice. However, we included those measurements when developing the FCU model to more accurately capture the equipment operating profile and validate the model. In this study, we also analyzed those measurements to show that some measurements can be valuable to reflect the equipment operational performance and hence can be considered when designing the FCU monitoring system.

Both the FCU fault-free model and fault inclusive model were validated in the Iowa Energy Center during the model development process. The detailed fault model validation process can be found in [19].

3.5.2. Description of control sequence

The occupied operation mode was set to 6:00 to 18:00 from Monday to Friday. Four control sequences, which are normally used in the FCU control in fields, were used for cooling coil valve control, heating coil valve control, fan control, and outdoor air damper control, respectively.

During the occupied mode, the zone cooling setpoint was set to 22.2 °C (72 °F) and heating setpoint was set to 20 °C (68 °F). Two PID control loops were used to adjust both cooling coil valve position and heating coil valve position respectively. If the zone temperature was above 21.67 °C (71 °F), i.e., 0.56 °C (1 °F) below the cooling setpoint, the FCU switched to the "cooling" mode. The cooling coil valve PID loop was enabled and the cooling valve position was controlled by the PID controller. If the zone temperature was below 20.56 °C (69 °F), i.e., 0.56 °C (1 °F) above the heating setpoint, the FCU switched to the "heating" mode. The heating coil valve PID loop was enabled and the heating valve position was controlled by the PID controller.

In the FCU model, a 3-speed fan with "Automatic On/Off" (Auto) mode was adopted. The operation of fan on/off status and fan speed is controlled according to the cooling PID output and heating PID output. There are three speed levels: 1) low-speed condition: the PID outputs (the cooling/heating coil valve position) is >0% and <40%; 2) medium speed condition: the PID outputs (the cooling/heating coil valve position) is >= 40% and < 80%; and 3) high-speed condition: the PID outputs (the cooling/heating coil valve position) is >= 80% and <100%. A 10% dead band was set at each speed switchover level. When there is no heating or cooling demand, the supply air fan stops running.

The outdoor air damper was controlled to maintain a minimum damper position at 30% open position during the operation of the FCU.

3.5.3. Other simulation settings

Apart from the control sequences for FCU operation, the weather data and the zone load profile were defined to simulate the outer and inner operation conditions. In this study, the TMY weather data file for Des Moines, IA, U.S., where the FCU model was validated, was used as the weather inputs. The internal load density was set to be varied to simulate a typical load pattern in a zone in commercial buildings on weekdays. The hourly zone load density during the occupied hours within a weekday is given in Fig. 4.

3.5.4. Description of fault tests and data

In the study, 18 types of faults which include two sensor related faults (i.e., zone air temperature sensor positive bias fault and negative bias fault), six actuator related faults (i.e., outdoor air damper stuck fault, outdoor air damper leakage fault, cooling coil valve stuck fault, cooling coil valve leakage fault, heating coil valve stuck fault, and heating coil valve leakage fault), seven static part related faults (i.e., fan outlet blockage fault, heating coil fouling airside fault, heating coil fouling waterside fault, cooling coil fouling airside fault, cooling coil fouling waterside fault, filter restriction fault and outside air inlet block fault), and three control related faults

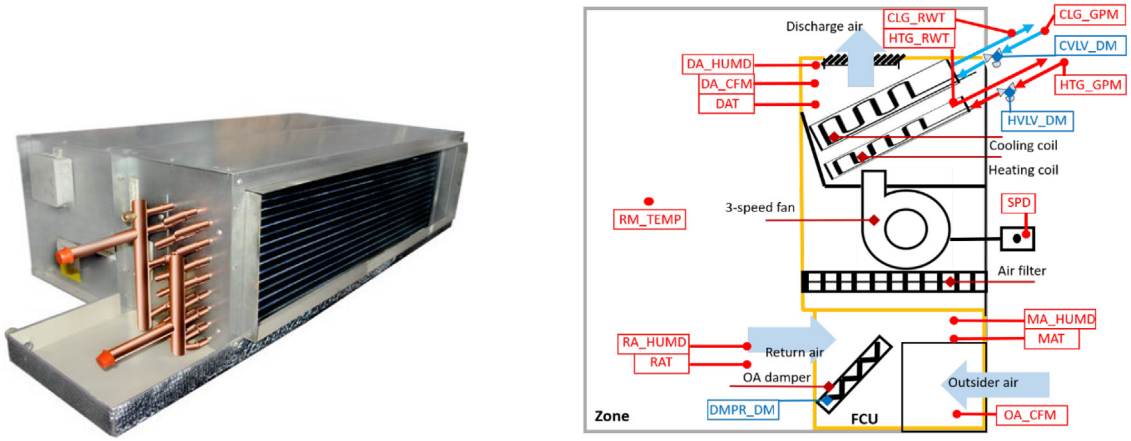

Fig. 3. Left: layout of FCU; right: schematic of a fan coil unit in the simulation.

Table 1
Summary of FCU data measurements.

No.	Measurement Name	Description	Unit
1	RM_TEMP	Zone temperature	°C
2	MAT	Mixed air temperature	°C
3	DAT	Discharge air temperature	°C
4	RAT	Return air temperature	°C
5	CVLV_DM	Cooling coil valve control signal	Open (0–1)
6	CLG_GPM	Cooling coil water flow rate	m^3/s
7	CLG_RWT	Cooling coil return water temperature	°C
8	HVLV_DM	Heating coil valve control signal	Open(0–1)
9	HTG_GPM	Heating coil water flow rate	m^3/s
10	HTG_RWT	Heating coil return water temperature	°C
11	DA_CFM	Discharge air flow rate	m^3/s
12	OA_CFM	Outdoor air flow rate	m^3/s
13	DMPR_DM	Outdoor air damper control signal	% Open
14	SPD	Fan speed	rev/s
15	MA_HUMD	Mixed air relative humidity	%
16	DA_HUMD	Discharge air relative humidity	%
17	RA_HUMD	Return air relative humidity	%

(i.e., cooling control reverse fault, heating control reverse fault, and control stable fault), were simulated. Among these 18 fault types, 13 faults were simulated at different fault severity levels as described in Appendix I of this paper. Consequently, a total of 48 fault cases were simulated in this study. Each fault case was simulated for one-year operation to obtain a complete operating data set under different weather conditions. Detailed descriptions for each type of fault and implementation methods can be found in the Appendix I of this paper.

Both fault-free (i.e., fault exclusive) data and faulty (i.e., fault inclusive) data were generated in .csv format files. The faulty data for each fault case was stored in one .csv file and was used to evaluate the fault symptoms. The fault-free data was used to generate the baseline data.

4. Results

In this section, we present the results of fault symptom analysis. We first provide the primary parameters obtained from the fault-free data as fundamentals in Section 4.1. Then we provide an example fault case to show the analysis scenario, and illustrate the total SOP results from all fault test cases as given in Section 4.2. We provide SDCD results to show the fault symptom intensity followed by one example as given in Section 4.3. Finally, we discuss some potential applications of the fault effect evaluation in Section 4.4.

4.1. Description of baseline data

In this study, the simulation time step was set to 5 s, which is a common time interval used by field direct digital controllers to update their output. The simulation output rate was set to 1-minute interval, which is the common data sampling rate in the BAS. Consequently, for the fault-free test case and each fault test case, the simulation generates 187,920 operating minutes (i.e., the number of samples under a 1-min sampling rate) and 12 operating hours within 261 operating days for the fault-free case and each fault test case in one year.

The OAT in the operating time period is equally binned into 10 windows with a bin size of 6 °C. Table 2 provides the median OAT value, operation duration and operation time ratio in each binned OAT window. It can be seen that the operation duration from the #5 window to the #9 window accounts for 74.5% of operation minutes.

Among 17 measurements listed in Table 1, the symptom on the zone temperature measurement is generated based on method #1, i.e., zone temperature is compared with the zone temperature setpoint to generate symptoms. In this study, we extend 0.82 °C to the setpoint (i.e., the cooling setpoint plus 0.82 °C and the heating setpoint minus 0.82 °C) as the baseline to generate the fault symptom on the zone temperature. For the HVAC systems in commercial buildings or residential buildings where the zone temperature setpoint is not required to be accurately maintained, this may avoid too many observations of zone temperature abnormalities. Consequently, the positive symptom is recorded when zone temperature is higher than 23.7 °C, and the negative symptom is recorded when zone temperature is lower than 18.9 °C.

Symptoms on other 16 measurements are generated from method #2, i.e., the measurement data in the fault inclusive data set is compared with the baseline (i.e., mean value of the measurement under each binned OAT window) generated through the fault-free data set.

4.2. SOP analysis

Here, we employ the "heating valve leakage" (VLVLeak_Heating) fault as an example to demonstrate the SOP analysis scenario. This fault was simulated under three severity levels as 20%, 50%, and 80% were simulated. Figs. 5–7 provide SOP distribution results for each measurement under each binned temperature window.

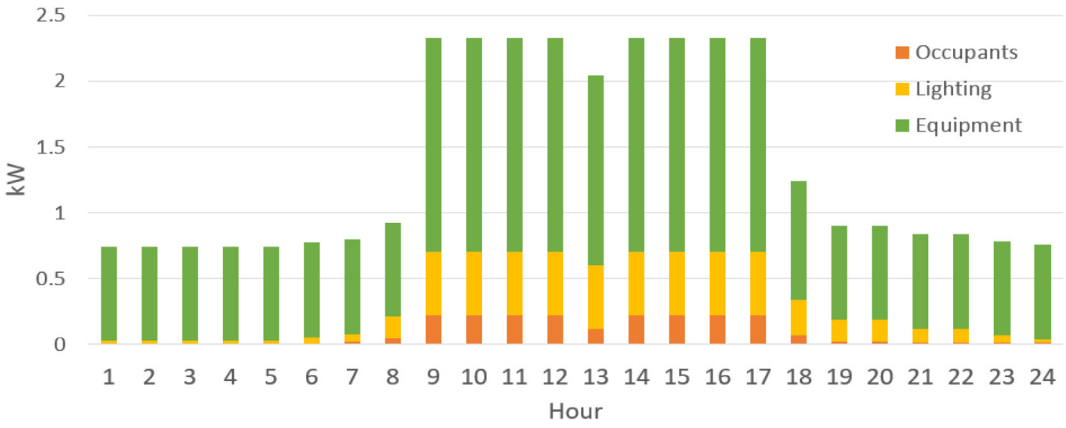

Fig. 4. Hourly zone load setting (kW).

Table 2
Primary parameters for fault symptom evaluation.

Bin No.	Bin #1	Bin #2	Bin #3	Bin #4	Bin #5	Bin #6	Bin #7	Bin #8	Bin #9	Bin #10
Median OAT (°C)	−20.3	−14.4	−8.5	−2.5	3.4	9.4	15.7	21.3	27.2	33.1
Operation duration (minutes)	2445	7580	11,919	17,665	31,987	21,548	21,127	32,737	32,640	8272
Operation time ratio (%)	1.3	4.0	6.3	9.4	17.0	11.5	11.2	17.4	17.4	4.4

The darker circle (or the smaller circle) shows the lower SOP for the specific measurement, and the lighter circle (or the larger circle) shows the higher SOP for the specific measurement.

From Figs. 5–7, it can be seen that for some measurements, the SOP values in the high temperature windows are high. For example, under the fault severity level at 20% leakage, the SOP value of CVLV_DM at a positive direction dramatically increases in the #7 window (i.e., binned OAT at 15.7 °C). This is because when the OAT is high, the cooling coil valve position should be increased to compensate for the extra cooling need caused by the leaking heating valve fault. Hence, this fault causes a simultaneous heating and cooling operation status.

The total SOP for each measurement under a specific fault severity level can be calculated by using Eq. (6). For the faults which were simulated on various severity levels (e.g., for the cooling coil stuck fault, there are five severity levels), the total SOP ranges can be obtained to indicate SOP values calculated under different severity levels. The total SOP distribution results for each measurement under 18 fault types are illustrated in two enhanced heatmaps as Figs. 8 and 9. In both enhanced heatmaps, the blue color cells and pink color cells represent a single SOP value, which indicates two conditions as 1) there was only one fault severity level in such a fault type (e.g., the Control_CoolingReverse fault), and 2) the SOP values are the same under different severity levels (e.g., the RM_TMP for the OABlock fault), for each measurement. In addition, we also color labeled the cells with red, orange, and green respectively to categorize the total SOP ranges, which were obtained when a SOP value varies under different fault severity levels. In both figures, the total SOP range values are categorized into three levels as 1) the minimum value is higher than 50% (in red color). For example, the MAT under the SensorBias_RMTemp_Neg fault (the SOP (+) ranges from 68% to 84%); 2) the difference between minimum value and maximum value is higher than 40% (in orange color). For example, the cooling water flow rate (CLG_GPM) under the SensorBias_RMTemp_Pos fault (the SOP (+) ranges from 33% to 57%); and 3) the maximum value is lower than 50% and difference between minimum value and maximum value is lower than 40% (in green color). For example, the MAT under the FilterRestriction fault (the SOP (+) ranges from 4% to 21%).

Additionally, from Figs. 8 and 9, it can be seen that for the VLVLeak_Heating fault, the total SOP value may be different for various measurements. Three measurements such as HVLV_DM, DMPR_DM, MA_HUMD, the total SOP values are relatively lower. This is due to two reasons. First, the fault symptoms on some measurements can be hardly observed. For example, the total SOP value on the DMPR_DM measurement (either for the positive symptom or negative symptom) is zero because the damper control loop in the simulation does not include any measurements in the FCU, and hence, is relatively isolated. Second, the operation time period when the occurrence of the symptom is relatively short. For example, a high SOP value for the HVLV_DM as can be seen in the # 1, #2, #3 and #4 windows (i.e., when the OAT is relatively low) as shown in Figs. 5 to 7. However, the operation time ratios for those four binned OAT windows are only 1%, 4%, 6% and 9% respectively. This causes the total SOP value to be relatively low.

Five measurements (i.e., DAT, CVLV_DM, CLG_GPM, CLG_RWT, and HTG_GPM) have significantly high SOP values (i.e., the minimum SOP is higher than 50%). This indicates that fault symptoms on those measurements can be more easily observed. For example, for the DAT measurement, the positive symptom total SOP ranges from 50% to 91% under different severity levels, i.e., the discharge air temperature would be more likely higher than the baseline when the heating coil valve is leaking. When the valve leaking is more severe, the symptom will be more observable. While for this measurement, the negative symptom total SOP is at 0% which means that it is impossible to observe the discharge air temperature to be lower than the baseline when the heating coil valve is leaking.

Among the measurements that have high total SOP value, the narrow total SOP range of a measurement for a specific fault indicates that this measurement may have a similar occurrence probability for various fault severity levels. For example, the measurements of CVLV_DM, CLG_GPM, and HTG_GPM have posi-

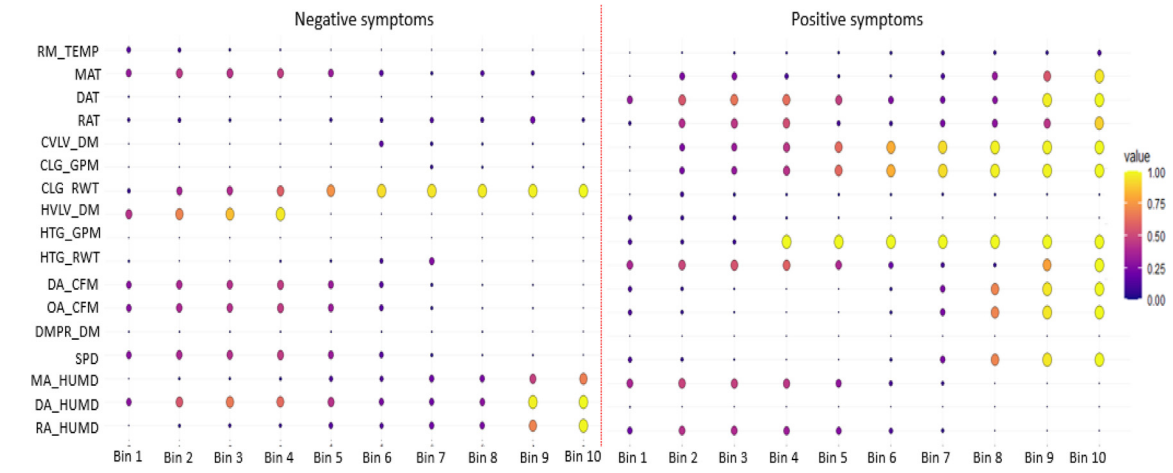

Fig. 5. SOP under each binned OAT window, VLVLeak_Heating fault (20% leakage).

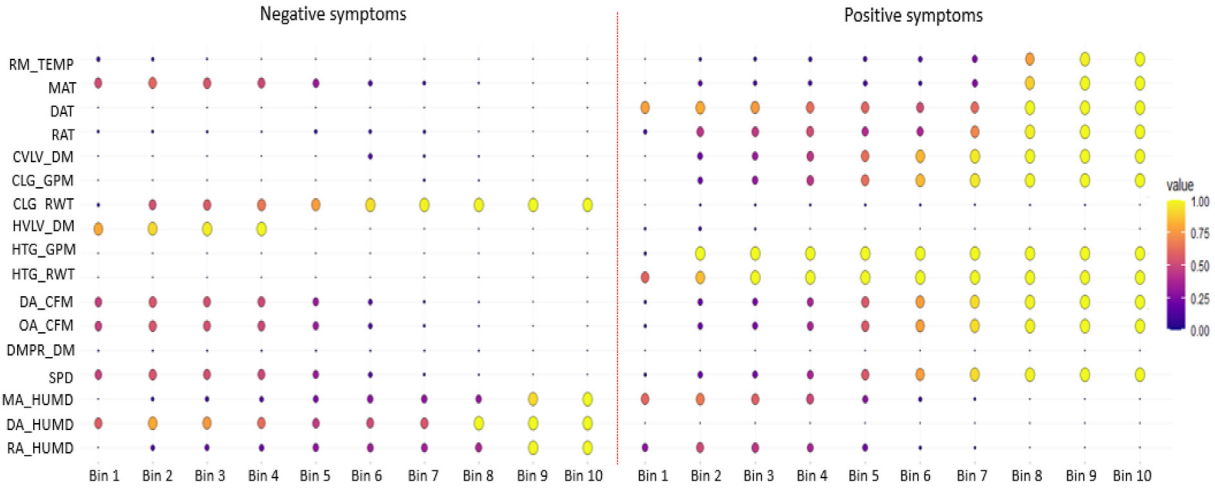

Fig. 6. SOP under each binned OAT window, VLVLeak_Heating fault (50% leakage).

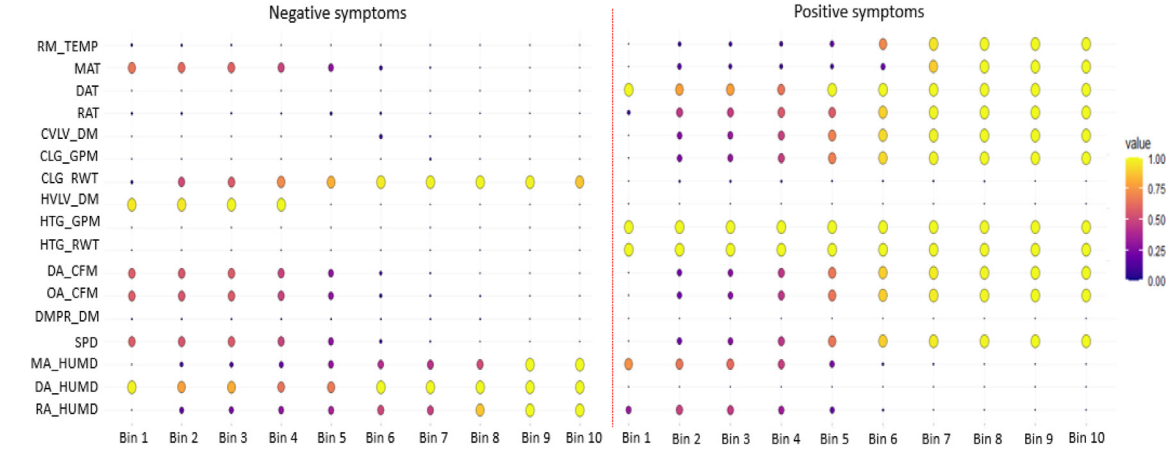

Fig. 7. SOP under each binned OAT window, VLVLeak_Heating fault (80% leakage).

tive symptoms SOP range from 74% to 77%, 73% to 77%, and 86% to 97% respectively. Therefore, when using the SOP values to develop fault diagnostics methods, those measurements may not be used as a general way to isolate faults regardless of the fault severity levels. The SOP value may vary in a very wide range depending on different severity levels. For example, the SOP values for the positive symptom on the DAT can range from 50% to 91%. This means that the observability of fault symptoms on DAT are very sensitive to

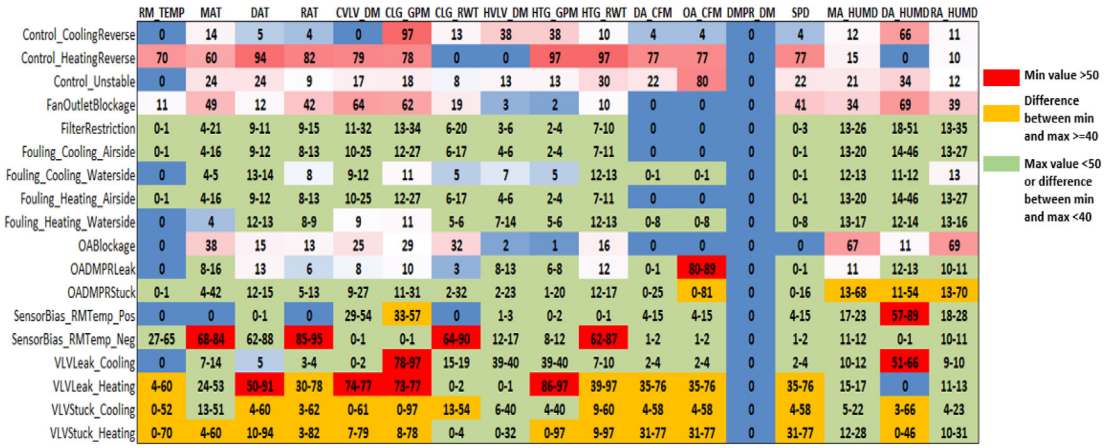

Fig. 8. Positive symptom (+) SOP range for each type of fault (%).

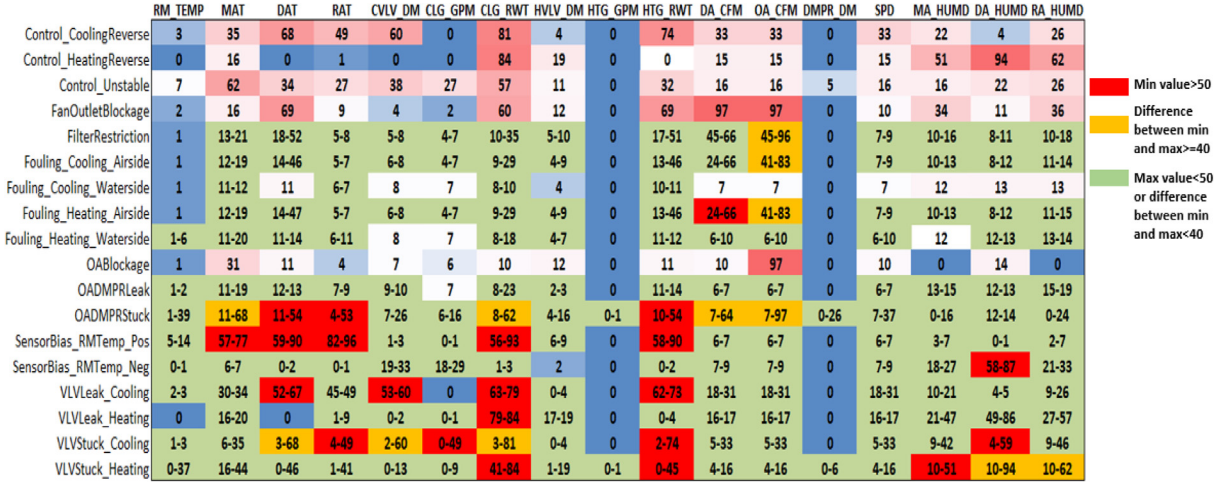

Fig. 9. Negative symptom (-) SOP range for each type of fault (%).

fault severity levels. But the overall symptom occurrence is high and hence this measurement can be used for diagnostic inference.

In addition, nine measurements (i.e., RM_TEMP, MAT, RAT, HTG_RWT, DA_CFM, OA_CFM, SPD, DA_HUMD, RA_HUMD) have wide ranges of total SOP values, considering different fault severity levels. For example, for the RM_TEMP, the total SOP value ranges from 4% to 60%. This not only indicates that the observability of fault symptoms on those measurements are very sensitive to fault severity levels, but also shows that the usage of those measurements in the diagnostic inference should be careful as will be discussed in Section 4.4.

4.3. SDCD analysis

As the FCU operation simulation data is output at a 1-minute time interval, the SDCD analysis is to test how many continuous minutes that a fault symptom can be observed. As illustrated in Section 3.4, SDCD represents whether a fault symptom could be constantly present during an operating day. For each fault, the mean SDCD ranges of positive symptom and negative symptom for all fault severity levels were calculated by averaging the SDCD values calculated from 261 operating days.

Figs. 10 and 11 illustrate the mean SDCD values of two symptom directions for each measurement respectively. Using the same approach when presenting the total SOP results as illustrated in Figs. 8 and 9, we use blue cells and pink cells to indicate the single mean SDCD value, as well as color labeled cells to categorize the SDCD range values when various fault severities were performed in Figs. 10 and 11. We categorized the SDCD range values into four levels according to the minimum value of the mean SDCD and maximum value of the mean SDCD. These four levels are: 1) the minimum value is higher than 121-min (in red color). For example, the MAT under the SensorBias_RMTemp_Neg fault (the mean SDCD (+) ranges from 475-min to 558-min); 2) the minimum value is between 61-min to 120-min (in orange color). For example, the DAT under the OADMPRLeak fault (the mean SDCD (+) ranges from 72-to 73-min); 3) the minimum value is lower than 60-min and the difference between minimum and maximum value is higher than 120-min (in yellow color). For example, the MAT under the FilterRestriction fault (the mean SDCD (+) ranges from 24-min to 127-min); and 4) the minimum value is lower than 60-min and the difference between minimum and maximum value is higher than 120-min (in green color). For example, the MAT under the OADMPRLeak fault (the mean SDCD (+) ranges from 42-min to 106-min).

It can be seen that for some measurements, the mean SDCD is very low. For example, for the damper position control signal (DMPR_DM), the mean SDCD of negative symptom (i.e., the damper control signal is lower than 30% as it should be) is higher than 60-min only for the "outdoor air damper stuck" fault under certain

fault severity levels, but not for other faults. After checking the operation, it was found that this symptom occurs when the outdoor air damper is stuck at a higher position during the winter season (when the OAT is very low). Consequently, this caused the FCU operation to be terminated according to a specific control sequence when the MAT was lower than 0 °C. Therefore, the DMPR_DM output zero value which was lower than the baseline. This indicates that the outdoor air damper control signal cannot be strongly affected by most faults, and hence, the symptom cannot continuously present. On the contrary, the mean SDCD values for some measurements are higher for most faults. That is to say, the symptom can continuously present during a day if a fault occurs. For example, the MAT presents a higher mean SDCD value at most faults. However, the range of mean SDCD value for some faults at different severity levels may be very wide. For example, for the cooling coil stuck fault, the mean SDCD of MAT can range from 2-min to 376-min depending on various fault severity levels. This suggests that the continuous symptom presence may be sensitive to certain fault severity levels. For the minor fault, the fault symptom may not be observed continuously.

In addition, the evaluation of the SDCD can be carried out by analyzing the percentage of days that the SDCD value is higher than a predefined threshold value. For example, the percentage of days, which the SDCD value is higher than 60-min, can be calculated to determine if the fault symptom is a strong symptom or a weak symptom.

Here, we use the "cooling coil stuck" (VLVStuck_Cooling) fault as an example to demonstrate the SDCD analysis. This fault was imposed at five severity levels as the valve was stuck at 0%, 20%, 50%, 80% and 100% position respectively. Fig. 12 shows the percentage of operating days that the SDCD values are higher than 60-min under each fault severity level. It can be seen that under the VLVStuck_Cooling fault, four measurements such as DAT, MAT, HTG_RWT, and DA_HUMD tend to steadily present a higher percentage of operation days (i.e., higher than 50% operating days) under the SDCD value is higher than 60-min. For example, for the DAT measurement, the percentage of operating days reaches 87% for the positive symptom (discharge air temperature is higher than the baseline) when the valve is stuck at 0% position. This value is 67%, 82%, 84% and 84% for negative symptoms (discharge air temperature is lower than the baseline) when the valve is stuck at 20%, 50%, 80% and 100% position respectively. However, for some measurements, the percentage of days that the SDCD is higher than 60-min is higher only under some fault severity levels. For example, for the SPD measurement, only when the cooling coil valve is stuck at 0% position, the percentage of days that the SDCD value is higher than 60-min can reach 75%.

4.4. Discussion

In this section, we discuss applications of the developed framework, as well as the obtained SOP and SDCD values to facilitate the development of FDD approaches as well as FCU fault prioritization.

4.4.1. SOP scalability

In the study, the total SOP distribution for each measurement is calculated by aggregating SOP values in each binned OAT window. A similar total SOP distribution can be calculated by considering the operation time ratio under different zone loads if the zone load presents a wide range distribution.

In addition, the total SOP distribution can be re-calculated when the operating time ratio within each binned OAT window is significantly different from what is presented in this study. For example, if a building is located in a hot climate zone where a higher percentage of operation time period is present during the OAT is very high, the total SOP distributions should be re-calculated by using the new percentage of time period in each binned OAT window as given in Eq. (6). For this purpose, we provide the entire SOP distribution table, which lists all the SOP distributions under each binned OAT window for each type of fault in Appendix II on the website. Users can generate the new total SOP distributions by adjusting the operation time ratio as given in Eq. (6) according to the climate zone where their buildings are located.

4.4.2. Usage for probability-based fault diagnostics

The total SOP distribution table can be used to develop inference approaches such as Bayesian Network (BN), fuzzy logic or fault tree in the fault diagnostic process. For example, in the BN diagnostic method, this table can be used as the conditional probability distributions in developing the BN parameter model [32]. It is noted that the total SOP value may highly rely on the fault severity level as shown in Section 4.2. The total SOP values for some measurements have a wide range depending on different fault severity levels. The usage of high total SOP values could cause the fault diagnostic approach to be very sensitive to the measurement, and as a result, could lead to mis-diagnosis. Therefore, two approaches are suggested to address this issue. First, the low total SOP values can be employed at the initial step to weaken the sensitivity of measurement and modified after the real diagnostic result is obtained to evaluate the fault diagnostic method. For example, for the RM_TEMP measurement, the total SOP value ranges from 4% to 60% for the VLVLeak_Heating fault as given in Section 4.2. An initial total SOP value of 5% to 15% can be adopted to test the diagnostic method. Second, a discretization processing

	RM_TEMP	MAT	DAT	RAT	CVLV_DM	CLG_GPM	CLG_RWT	HVLV_DM	HTG_GPM	HTG_RWT	DA_CFM	OA_CFM	DMPR_DM	SPD	MA_HUMD	DA_HUMD	RA_HUMD
Control_CoolingReverse	0	3	37	15	0	720	1	265	264	72	33	33	0	33	70	486	75
Control_HeatingReverse	508	424	694	606	585	584	3	0	720	720	574	574	0	574	87	0	75
Control_Unstable	0	143	31	13	25	27	3	32	32	57	40	514	0	40	135	100	87
FanOutletBlockage	56	253	64	226	397	382	84	27	19	52	0	0	0	268	237	407	275
FilterRestriction	5-6	24-127	60-69	47-84	77-212	84-220	36-110	27-43	17-32	47-62	0-4	0-5	0	6-22	74-155	124-339	80-225
Fouling_Cooling_Airside	5	23-90	61-71	42-77	69-173	76-181	33-93	29-47	18-35	46-65	0-6	0-6	0	6-13	70-121	98-309	76-168
Fouling_Cooling_Waterside	5	25-27	74-78	38-41	62-83	70-76	28-30	49	37	69-73	6-14	6-14	0	6-14	68-69	74-80	74
Fouling_Heating_Airside	5	23-90	60-71	42-77	69-172	76-180	33-93	29-47	18-35	46-65	0-6	0-6	0	6-13	70-121	98-309	76-168
Fouling_Heating_Waterside	5	24-25	69-74	37-45	60-61	69	31-32	51-100	37-49	70-73	7-62	7-62	0	7-62	68-93	80-96	73-92
OABlockage	5	211	74	87	171	195	177	20	0	74	0	0	0	0	450	78	463
OADMPRLeak	4	42-106	72-73	26-30	54-56	65-68	16-19	58-86	44-57	68-69	7-13	541-600	0	7-13	65-66	83-86	59-65
OADMPRStuck	4-11	24-241	69-78	25-87	61-184	69-212	15-173	19-146	10-120	69-101	0-134	0-547	0	0-114	69-455	75-384	74-468
SensorBias_RMTemp_pos	6-8	0-1	2-10	0-1	196-377	223-399	0-1	10-25	6-17	2-9	31-116	31-116	0	31-116	99-139	373-605	113-169
SensorBias_RMTemp_neg	2-3	475-558	320-619	610-683	1-10	0-10	375-645	84-114	60-85	324-617	8-17	8-17	0	8-17	61-62	5-13	58-59
VLVLeak_Cooling	0-1	2-17	35-41	9-15	0-14	561-720	1-4	271-273	265-271	57-71	20-34	20-34	0	20-34	54-64	352-481	61-68
VLVLeak_Heating	22-433	125-383	307-651	161-564	484-570	482-569	3-10	0-9	640-720	233-720	231-558	231-558	0	231-558	88-99	0	76-91
VLVStuck_Cooling	0-383	2-376	34-435	8-435	0-448	0-720	1-385	44-273	34-271	66-435	32-428	32-428	0	32-428	27-138	23-486	30-157
VLVStuck_Heating	4-508	19-424	51-694	19-606	44-585	55-584	3-23	0-236	0-720	51-720	231-574	231-574	0	231-574	82-187	0-333	75-204

Min value>121 min

61 min<Min value<120 min

Min value<60 min, and difference between min and max>120 min

Min value<60 min, and difference between min and max<120 min

Fig. 10. Positive symptom (+) mean SDCD range for each type of fault (minutes).

	RM_TEMP	MAT	DAT	RAT	CVLV_DM	CLG_GPM	CLG_RWT	HVLV_DM	HTG_GPM	HTG_RWT	DA_CFM	OA_CFM	DMPR_DM	SPD	MA_HUMD	DA_HUMD	RA_HUMD
Control_CoolingReverse	22	117	493	167	445	0	312	33	0	540	44	44	1	44	119	33	188
Control_HeatingReverse	1	79	0	4	5	0	574	147	0	0	79	79	1	79	373	694	456
Control_Unstable	55	353	104	83	59	32	312	65	7	102	78	78	39	78	108	31	166
FanOutletBlockage	11	102	409	51	25	11	351	83	3	406	720	720	1	74	242	63	265
FilterRestriction	9-11	61-119	118-345	30-44	36-48	22-37	68-221	38-69	1-2	115-340	306-476	306-718	1	39-63	70-116	58-65	72-133
Fouling_Cooling_Airside	10-11	56-109	94-311	30-37	39-50	25-38	60-190	34-64	1-2	91-307	152-473	268-598	1	37-60	72-95	58-69	74-107
Fouling_Cooling_Waterside	11	51-53	72-77	32-35	49-51	38	56-64	29	1	69-74	34	34	1	34	83-84	71-76	87-88
Fouling_Heating_Airside	10-11	56-108	94-312	30-37	38-50	25-38	60-188	34-65	1-2	91-308	152-473	268-598	1	37-60	72-96	58-69	74-108
Fouling_Heating_Waterside	12-47	50-94	77-94	32-67	51	39	54-114	28-50	0-1	74-83	33-56	33-56	1	33-56	83-85	67-71	87-97
OABlockage	8	222	76	28	50	36	64	85	3	73	75	720	1	75	2	72	3
OADMPRLeak	11-13	49-87	81-86	36-46	54-67	39-40	59-155	15-21	0	76-92	31-32	31-32	1	31-32	95-105	70-71	105-132
OADMPRStuck	8-294	51-498	73-385	28-353	50-191	36-110	54-454	29-123	1-9	71-391	34-395	34-720	1-175	34-210	1-116	70-78	2-166
SensorBias_RMTemp_pos	7-9	383-515	385-618	556-715	12-25	3-10	360-600	44-69	1-2	381-606	33-38	33-38	1	33-38	26-53	2-10	20-49
SensorBias_RMTemp_neg	14-20	28-32	5-14	3-7	107-182	94-145	12-23	18-19	0	5-14	35-43	35-43	1	35-43	119-172	307-609	136-209
VLVLeak_Cooling	17-22	97-111	356-488	169-190	381-445	0	317-361	4-28	0-1	424-537	23-37	23-37	1	23-37	60-115	32-40	57-188
VLVLeak_Heating	1-5	82-109	0	5-41	6-18	1-5	486-573	123-145	0	0-21	80-94	80-94	1	80-94	145-339	308-615	180-416
VLVStuck_Cooling	10-22	27-117	24-493	22-181	16-445	0-362	25-352	3-34	0-1	20-540	22-44	22-44	1	22-44	60-304	31-430	57-339
VLVStuck_Heating	1-276	80-302	0-333	4-304	5-87	0-52	300-574	13-147	0-9	0-325	33-81	35-81	1-48	33-81	77-373	49-694	77-456

Min value>121 min
61 min<Min value<120 min
Min value<60 min, and difference between min and max>120 min
Min value<60 min, and difference between min and max<120 min

Fig. 11. Negative symptom (-) mean SDCD range for each type of fault (minutes).

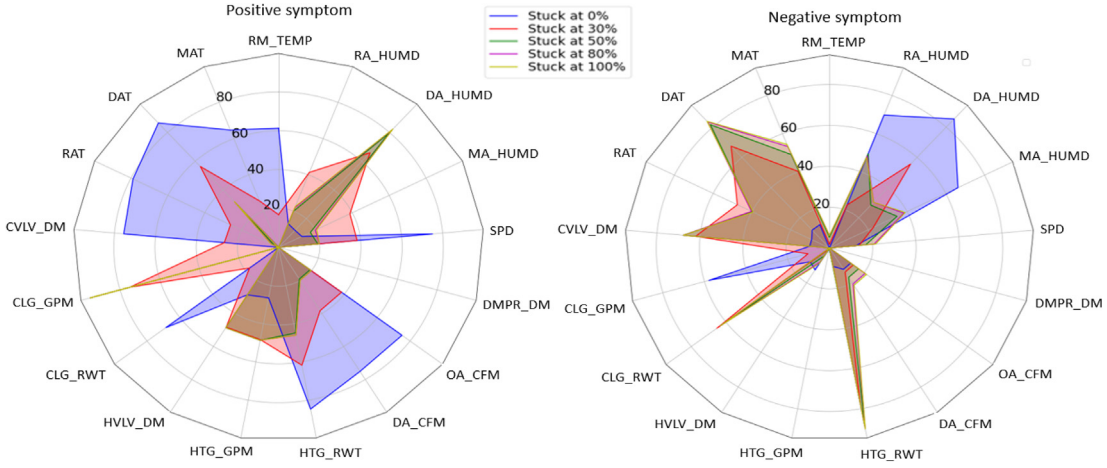

Fig. 12. Percentage of operating days that the SDCD is higher than 60-min (using VLVStuck_Cooling fault).

can be used to discretize the total SOP value [33]. For example, the total SOP value of a measurement can be divided into three levels as "weak" (say, the total SOP value from 0% to 25%), "moderate" (say, the total SOP value from 26% to 50%), and "strong" (say, the total SOP value from 51% to 100%). By this means, the RM_TEMP measurement under the VLVLeak_Heating fault can be considered as a measurement showing a "moderate" symptom.

4.4.3. Fault ranking according to SDCD results

The fault ranking based on the energy consumption impact and zone comfort impact was usually carried out. In this study, we rank the fault according to the fault impacts on system operation performance, i.e., fault effects on various measurements. Here, we use the SDCD result and three steps to rank the fault. First, we classify three SDCD levels such that the SDCD is<60-min (i.e., the level 3 measurement group), the SDCD is between 61-min to 120-min (i.e., the level 2 measurement group), and the SDCD is>121-min (i.e., the level 1 measurement group). Second, we analyze how many measurements may fall into that level. Last, we rank the fault according to the number of measurements following the order of level 3 to level 1. This reflects that fault impacts on measurements based on a temporal scale. Therefore, the more numbers of measurements in the longer duration of symptom occurrence may indicate that the fault affects the system operating performance more significantly. We use the SDCD value calculated under the minor fault severity level to rank the fault. For example, for the VLVStuck_Heating fault, the mean SDCD values of positive symptom on the RM_TEMP measurement is from 4-min to 508-min at different fault severity levels, as can be seen in Fig. 10. Here, we use 4-min and this measurement falls into the level 3 measurement class. Additionally, if a measurement falls into two classes for the positive symptom and the negative symptom respectively, we classify the measurement to a higher level. For example, for the MAT under the Control_HeatingReverse fault, the SDCD values for the positive symptom and negative symptom are 424-min and 79-min respectively. Therefore, we classify this measurement into the level 1 group. Table 3 provides the fault ranking result.

It can be seen that the Control_HeatingReverse fault may cause the most severe fault impact on the system operating performance because under such a fault, the mean SDCD values for 16 measurements are higher than 121-min (i.e., level 1 measurement). However, the VLVStuck_Cooling fault has a minor impact on the system operating performance because under such a fault the mean SDCD value for only one measurement (i.e., HTG_RWT) is higher than 61-min. Two reasons lead to a lower ranking result of the VLVStuck_Cooling fault. First, we select the mean SDCD value from the minor fault severity level (i.e., valve stuck at a 20% position) to rank the faults. Consequently, under this fault

Table 3
Fault ranking result (according to the lowest SDCD value of each measurement).

Rank No.	Fault name	Number of level 1 measurements (>121 min)	Number of level 2 measurements (61 to 120 min)	Number of level 3 measurements (<60 min)
1	Control_HeatingReverse	16	0	1
2	VLVLeak_Heating	15	0	2
3	FanOutletBlockage	10	1	6
4	VLVLeak_Cooling	9	2	6
5	Control_CoolingReverse	8	3	6
6	SensorBias_RMTemp_Pos	8	2	7
7	SensorBias_RMTemp_Neg	7	4	6
8	OABlockage	7	2	8
9	Control_Unstable	5	6	6
10	VLVStuck_Heating	4	3	10
11	FilterRestriction	2	9	6
12	Fouling_Cooling_Airside	2	7	8
13	Fouling_Heating_Airside	2	7	8
14	OADMPRLeak	1	5	11
15	OADMPRStuck	0	7	10
16	Fouling_Cooling_Waterside	0	6	11
17	Fouling_Heating_Waterside	0	5	12
18	VLVStuck_Cooling	0	1	16

severity level, the operational behavior is very close to the normal operational behavior. This causes minor and unobservable fault symptoms. Secondly, the equipment operation is relatively robust when this fault occurs because the control sequence can compensate for the negative effects caused by the fault. For example, the control sequence can increase the fan speed to provide more cooling needs in the cooling mode, or increase the heating coil valve position to compensate for the unnecessary cooling in the heating mode.

5. Conclusions and future work

In this paper, we illustrate a simulation-based evaluation framework to systematically analyze fan coil unit (FCU) fault effects which are presented as fault symptoms on various measurements. In the framework, we discussed fault symptom generation methods commonly used in monitoring HVAC systems. When analyzing fault symptoms, apart from the fault symptom direction and magnitude which were previously investigated, we employed two novel metrics, namely fault symptom occurrence probability (SOP) and fault continuous symptom daily duration (SDCD) to quantify the fault symptom occurrence likelihood and intensity on measurements under various faults. By using both metrics to analyze fault symptoms, fault effects on various measurements in a FCU can be completely evaluated.

We imposed 18 types of faults with different severity levels on the developed FCU simulation platform to generate 48 fault simulation cases. For each case, the simulation was carried out to generate one-year simulation results so that fault inclusive data cover all possible inner and outer operation conditions. From the analysis of SOP and SDCD distributions, we demonstrate that both metrics can benefit multiple applications such as the development of probability-based FDD approaches and fault prioritization.

Our future works include: 1) using the developed method to analyze fault symptoms on more HVAC systems and equipment so that measurement sensitivities can be obtained for different type of HVAC systems; 2) employing the Monte Carlo simulation to simulate faults under different operation modes, climate conditions, and system configurations so that the SOP and SDCD distributions can be more accurately calculated to reflect faulty operation under various real operation conditions; and 3) evaluating fault impact propagation in a completed HVAC system so that hierarchical distribution features of fault effects on various measurements can be obtained.

CRediT authorship contribution statement

Yimin Chen: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft. **Guanjing Lin:** Conceptualization, Writing – review & editing. **Zhelun Chen:** Data curation, Formal analysis, Writing – review & editing. **Jin Wen:** Data curation, Writing – review & editing. **Jessica Granderson:** Project administration, Funding acquisition.

Declaration of Competing Interest

The 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.

Acknowledgements

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Office, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Dr. Ran Liu is highly appreciated for his assistance in explaining the FCU model.

Appendix II. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.enbuild.2022.112041.

References

[1] J. Granderson, G. Lin, A. Harding, P. Im, Y. Chen, Building fault detection data to aid diagnostic algorithm creation and performance testing, Sci. Data 7 (2020) 65, https://doi.org/10.1038/s41597-020-0398-6.
[2] S. Katipamula, M.R. Brambley, Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I, HVACR Res. 11 (2005) 3–25, https://doi.org/10.1080/10789669.2005.10391123.
[3] W. Kim, S. Katipamula, A review of fault detection and diagnostics methods for building systems, Sci. Technol. Built Environ. 24 (1) (2018) 3–21, https://doi.org/10.1080/23744731.2017.1318008.
[4] G. Lin, H. Kramer, J. Granderson, Building fault detection and diagnostics: achieved savings, and methods to evaluate algorithm performance, Build. Environ. 168 (2020) 106505, https://doi.org/10.1016/j.buildenv.2019.106505.
[5] M. Breuker, J. Braun, Common faults and their impacts for rooftop air conditioners, HVACR Res. 4 (3) (1998) 303–318, https://doi.org/10.1080/10789669.1998.10391406.
[6] J. Schein, S. Bushby, A hierarchical rule-based fault detection and diagnostic method for HVAC systems, HVACR Res. 12 (1) (2006) 111–125.

Appendix I. Description of FCU faults and implementation methods

Fault Description				Method of Fault Imposition	Number of Cases
Component Type	Fault Name	Fault Abbreviation	Fault Severity		
Fan	Outlet blockage	FanOutletBlockage	Outlet resistance + 2400% (corresponds to an 80% flow rate reduction)	Increase air flow pressure resistance	1
Heating coil	Fouling airside	Fouling_Heating_Airside	(1) Severe: air flow resistance increases by 200%, heat transfer rate decreases by 10%; (2) Middle: air flow resistance increases by 50%, heat transfer rate decreases by 5%; and (3) Minor: air flow resistance increases by 10%, heat transfer rate decreases by 0%;	Increase air flow pressure resistance, decrease heat transfer coefficient	3
	Fouling waterside	Fouling_ Heating _Waterside	(1) Severe: water flow rate at fully open valve reduces by 50%, heat transfer rate reduces by 50%; (2) Middle: water flow rate at fully open valve reduces by 30%, heat transfer rate reduces by 30%; and (3) Minor: water flow rate at fully open valve reduces by 10%, heat transfer rate reduces by 10%;	Increase water flow pressure resistance, decrease heat transfer coefficient	3
Cooling coil	Fouling airside	Fouling_Cooling_Airside	(1) Severe: air flow resistance increases by 200%, heat transfer rate decreases by 10%; (2) Middle: air flow resistance increases by 50%, heat transfer rate decreases by 5%; (3) Minor: air flow resistance increases by 10%, heat transfer rate decreases by 0%;	Increase air flow pressure resistance	3
	Fouling waterside	Fouling_Cooling_Waterside	(1) Severe: water flow rate at fully open valve reduces by 50%, heat transfer rate reduces by 50%; (2) Middle: water flow rate at fully open valve reduces by 30%, heat transfer rate reduces by 30%; and (3) Minor: water flow rate at fully open valve reduces by 10%, heat transfer rate reduces by 10%;	Increase water flow pressure resistance, decrease heat transfer coefficient	3
Filter	Restriction	FilterRestriction	Outlet resistance + 23.45%, +56.25%, and + 400% (corresponding to 10%, 20%, and 50% flow rate reduction at the same pressure difference)	Increase air flow pressure resistance	3
Outdoor air inlet	Blockage	OABlockage	Face area −80%	Decrease damper face area	1
Outdoor air damper	Leaking	OADMPRLeak	Face area + 20%, +50%, and + 80%	Increase damper face area	3
	Stuck	OADMPRStuck	Stuck at 0%, 20%, 50%, 80% and 100%	Assign a fixed simulated controlled device position	5
Zone air temperature sensor	Sensor bias	SensorBias_RMTemp_Pos	+2°C and + 4 °C	Add bias to sensor output	2
Zone air temperature sensor	Sensor bias	SensorBias_RMTemp_Neg	−4°C and −2°C	Add bias to sensor output	2
Heating valve	Stuck	VLVStuck_Heating	Stuck at 0%, 20%, 50%, 80% and 100%	Assign a fixed simulated controlled device position	5
	Leaking	VLVLeak_Heating	20%, 50%, 80% of the max flow	Assign a water flow rate when fully closed	3
Cooling Valve	Stuck	VLVStuck_Cooling	Stuck at 0%, 20%, 50%, 80% and 100%	Assign a fixed simulated controlled device position	5
	Leaking	VLVLeak_Cooling	20%, 50%, and 80% of the max flow	Assign a water flow rate	3

Appendix I (*continued*)

Fault Description				Method of Fault Imposition	Number of Cases
Component Type	Fault Name	Fault Abbreviation	Fault Severity		
Control	FCU controller cycling	Control_Unstable	Unstable control	when fully closed Decrease all proportional bands to their 10% respectively	1
Control	Heating control	Control_HeatingReverse	Reverse control sequence	Modify simulated control strategy to allow reversed action	1
Control	Cooling control	Control_CoolingReverse	Reverse control sequence	Modify simulated control strategy to allow reversed action	1

[7] S. Kaldorf, P. Gruber, Practical experiences from developing and implementing an expert system diagnostic tool / Discussion, ASHRAE Trans. 108 (2002) 826.

[8] Y. Chen, J. Wen, T. Chen, O. Pradhan, Bayesian Networks for Whole Building Level Fault Diagnosis and Isolation, in: Int. High Perform. Build. Conf., West Lafayette, IN, 2018: pp. 1–10. https://docs.lib.purdue.edu/ihpbc/266..

[9] G. Zhang, G. Vachtsevanos, A Methodology for Optimum Sensor Localization/ Selection in Fault Diagnosis, in: 2007 IEEE Aerosp. Conf., 2007: pp. 1–8. doi: 10.1109/AERO.2007.352878..

[10] B. Thornton, A. Wagner, Variable Refrigerant Flow Systems, U.S. General Services Administration, San Francisco, CA, 2012. https://www.15000inc.com/wp/wp-content/uploads/VRF-Report-by-GSA.pdf (accessed June 1, 2021)..

[11] M.C. Comstock, J.E. Braun, E.A. Groll, The sensitivity of chiller performance to common faults, HVACR Res. 7 (3) (2001) 263–279, https://doi.org/10.1080/10789669.2001.10391274.

[12] S.-H. Cho, H.-C. Yang, M. Zaheer-uddin, B.-C. Ahn, Transient pattern analysis for fault detection and diagnosis of HVAC systems, Energy Convers. Manag. 46 (18-19) (2005) 3103–3116, https://doi.org/10.1016/j.enconman.2005.02.012.

[13] J.M. House, H. Vaezi-Nejad, J.M. Whitcomb, An expert rule set for fault detection in air-handling units / Discussion, ASHRAE Trans. 107 (2001) 858.

[14] Y. Chen, S. Huang, D. Vrabie, A simulation based approach to impact assessment of physical faults: large commercial building hvac case study, in: 2018 Build. Perform. Model. Conf. SimBuild Co-Organ. ASHRAE IBPSA-USA Chic. IL USA, 2018..

[15] Z. Shi, W. O'Brien, Using Building Performance Simulation for Fault Impact Evaluation, in: Proc. ESim 2018, Montréal, QC, Canada, 2019: pp. 124–132..

[16] Y. Li, Z. O'Neill, An innovative fault impact analysis framework for enhancing building operations, Energy Build. 199 (2019) 311–331, https://doi.org/10.1016/j.enbuild.2019.07.011.

[17] S. Ginestet, D. Marchio, O. Morisot, Evaluation of faults impacts on energy consumption and indoor air quality on an air handling unit, Energy Build. 40 (1) (2008) 51–57, https://doi.org/10.1016/j.enbuild.2007.01.012.

[18] J. Verhelst, G. Van Ham, D. Saelens, L. Helsen, Economic impact of persistent sensor and actuator faults in concrete core activated office buildings, Energy Build. 142 (2017) 111–127, https://doi.org/10.1016/j.enbuild.2017.02.052.

[19] S. Pourarian, J. Wen, D. Veronica, A. Pertzborn, X. Zhou, R. Liu, A tool for evaluating fault detection and diagnostic methods for fan coil units, Energy Build. 136 (2017) 151–160, https://doi.org/10.1016/j.enbuild.2016.12.018.

[20] R. Isermann, Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance, Springer Science & Business Media, Germany, 2005..

[21] L. Ma, Y. Ma, K.Y. Lee, An intelligent power plant fault diagnostics for varying degree of severity and loading conditions, IEEE Trans. Energy Convers. 25 (2010) 546–554, https://doi.org/10.1109/TEC.2009.2037435.

[22] C. Wen, Y.a.N. Tao, C.a.I. Wen, Y. Hong-yan, W.a.N. Zhong-hai, W.-P. Sung, T.-Y. Han, The extraction and application of fault characteristic vector for lower vacuum of condenser in 1000MW unit, MATEC Web Conf. 175 (2018) 02003, https://doi.org/10.1051/matecconf/201817502003.

[23] S. Dash, R. Rengaswamy, V. Venkatasubramanian, Fuzzy-logic based trend classification for fault diagnosis of chemical processes, Comput. Chem. Eng. 27 (2003) 347–362, https://doi.org/10.1016/S0098-1354(02)00214-4.

[24] Y. Chen, J. Wen, L.J. Lo, Using weather and schedule based pattern matching and feature based PCA for whole building fault detection — Part I Development of the method, ASME J. Eng. Sustain. Build. Cities (2021) 1–23, https://doi.org/10.1115/1.4052729.

[25] X. Liang, T. Hong, G.Q. Shen, Improving the accuracy of energy baseline models for commercial buildings with occupancy data, Appl. Energy 179 (2016) 247–260, https://doi.org/10.1016/j.apenergy.2016.06.141.

[26] R.T. Cleman, R.L. Winkler, Combining probability distributions from experts in risk analysis, Risk Anal. 19 (1999) 187–203.

[27] B.A. Olshausen, Bayesian probability theory, The Redwood Center for Theoretical Neuroscience, Helen Wills Neuroscience Institute at the University of California at Berkeley, Berkeley CA, 2004..

[28] S. Baek, D.-Y. Kim, Fault prediction via symptom pattern extraction using the discretized state vectors of multisensor signals, IEEE Trans. Ind. Inform. 15 (2) (2019) 922–931, https://doi.org/10.1109/TII.2018.2828856.

[29] H. Kim, W.C. Yoon, S. Choi, Aiding fault diagnosis under symptom masking in dynamic systems, Ergonomics. 42 (11) (1999) 1472–1481, https://doi.org/10.1080/001401399184820.

[30] J.M. Kościelny, K. Zakroczymski, Fault isolation method based on time sequences of symptom appearance, IFAC Proc. 33 (11) (2000) 499–504, https://doi.org/10.1016/S1474-6670(17)37408-6.

[31] D.R. Clark, W.B. May, HVACSIM+ building systems and equipment simulation program - user's guide, National Bureau of Standards, Building Equipment Division, Washington, DC, 1985. https://www.osti.gov/biblio/6127383 (accessed November 20, 2021)..

[32] 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.

[33] J. Rohmer, Uncertainties in conditional probability tables of discrete Bayesian Belief Networks: a comprehensive review, Eng. Appl. Artif. Intell. 88 (2020) 103384, https://doi.org/10.1016/j.engappai.2019.103384.