Title
Research challenges and directions in HVAC fault prevalence

Permalink
https://escholarship.org/uc/item/6084q8c9

Journal
Science and Technology for the Built Environment, 27(5)

ISSN
2374-4731

Authors
Kim, Janghyun
Trenbath, Kim
Granderson, Jessica
et al.

Publication Date
2021-05-28

DOI
10.1080/23744731.2021.1898243

Peer reviewed

Research challenges and directions in HVAC fault prevalence

Janghyun Kim,[1*] Kim Trenbath,[1] Jessica Granderson,[2] Yimin Chen,[2] Eliot Crowe,[2] Hayden Reeve,[3]
Sarah Newman[3], and Paul Ehrlich[4]

[1]National Renewable Energy Laboratory, Golden, Colorado, USA

[2]Lawrence Berkeley National Laboratory, Berkeley, California, USA

[3]Pacific Northwest National Laboratory, Richland, Washington, USA

[3]Pacific Northwest National Laboratory, Richland, Washington, USA

[4]Building Intelligence Group, Portland, Oregon, USA

Abstract

This study provides a review of the current state of knowledge, gaps, and potential value in research on the prevalence of faults in commercial buildings. Two separate efforts were made in this study: (1) we performed a literature review to determine the extent of currently available fault prevalence data for heating, ventilation, and air-conditioning (HVAC) systems, and (2) we conducted dozens of interviews with subject matter experts and stakeholders to determine the HVAC fault data that would be of greatest value. Through the literature review and interviews, we discovered unmet needs for empirical data on the prevalence of faults at the desired level of granularity, consistency, and scale; this lack of data leads us to recommend future work studying commercial buildings' HVAC fault prevalence, with robust fault taxonomy and a variety of meaningful fault prevalence metrics.

Keywords: fault prevalence, fault incidence, fault detection and diagnostics, literature review, interview

1. Introduction

According to the United States (U.S.) Energy Information Administration, the U.S. commercial building sector consumes approximately 5.2 PWh (17.83 quadrillion Btu) of primary energy annually, and heating, ventilation and air-conditioning (HVAC) systems make up 30% of the total commercial building energy consumption (Goetzler et al. 2017). Faults have a significant impact on U.S. commercial building operations and have been estimated to waste 205 TWh (0.7 quads) of energy annually—worth nearly 14 billion U.S. dollars (Roth et al. 2005). The detection of building HVAC faults has been well studied, with a myriad of publications dating back to the 1980s (Katipamula and Brambley 2005a, 2005b; Kim and Katipamula 2018).

Today, commercially available automated fault detection and diagnostics (FDD) tools are increasingly used to detect the presence of faults for operators and owners (mostly in HVAC systems) and sometimes the root causes of faults, providing visibility and support for corrective action. Dozens of commercial offerings exist (Granderson et al. 2017; Smart Energy Analytics Campaign 2019). These solutions typically apply algorithms to existing data streams from building automation systems or connected equipment, and they are being used by owners to enable significant cost-effective savings. For example, recent publications evaluating the use of commercial analytics technology across hundreds of millions of square feet of monitored buildings indicate savings of approximately 7%–9% of whole-building energy consumption on average (Kramer et al. 2019; Lin et al. 2020).

Although there has been significant growth in the development and deployment of FDD solutions, there has been less work understanding the prevalence of faults within the commercial building population as it was also mentioned in previous studies (Yuill and Braun 2013; Li and O'Neill 2018). Accordingly, this paper provides a review of the current state of knowledge, gaps, and potential value in further research on how often and prevalent faults are in commercial HVAC systems. The structure of this article is organized to provide how uncertain and sparse information on HVAC fault prevalence is by combining and summarizing available information from previous studies (Section 2). Then, confirm these limitations found from literature review with 25 stakeholders with their responses through interviews (Section 3). Interviews are also designed to understand key needs in the FDD community from experts and stakeholders so that future fault prevalence studies could be directed towards the correct pathway. A discussion and recommendations are provided in Section 4, followed by conclusions in Section 5.

2. Literature review

This section presents methodology and review of previous studies, as well as analyses on these studies, to answer the following questions: (1) Was there any common metrics used for quantifying how often faults occur and/or prevalent faults are? (2) What sample spaces were considered for quantifying faults? (3) How were faults categorized/classified while quantifying? (4) Which faults were most considered? (5) How were faults quantified? and (6) How much they differ between studies? A summary table used for deriving findings presented in the following sections is included in the Supplementary Material.

2.1 Methodology

We initially identified[1] a list of 65 literature resources based on specific topics: general review of FDD, assessing various fault types, and including any information related to the occurrence and

---

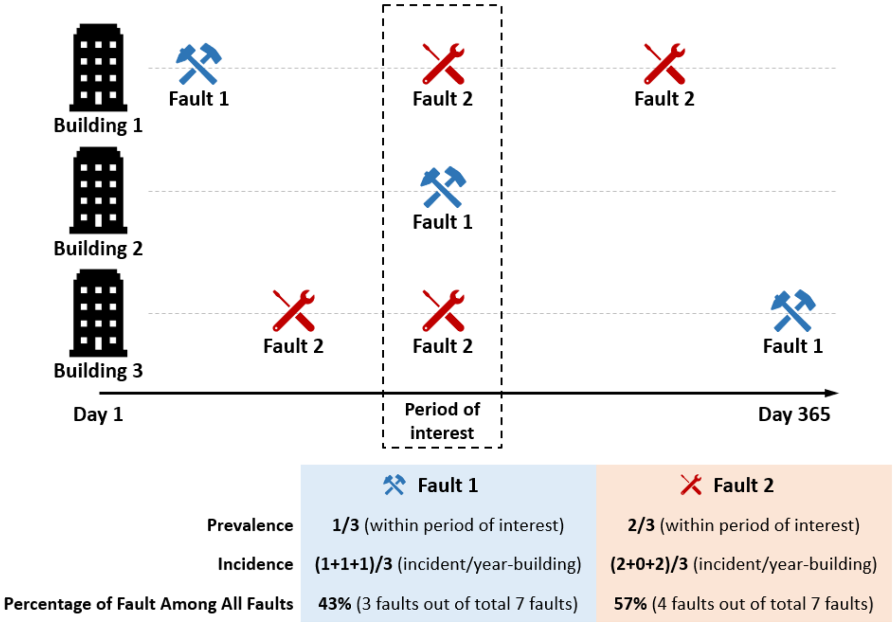
[1] Keywords such as building, HVAC, FDD, fault, incidence, prevalence, occurrence, and frequency were used on search engines such as sciencedirect, taylor francis online, and google scholar.

prevalence of faults. Out of these 65 studies, 41 studies included information related to the occurrence and prevalence of faults. None of the studies included specific definitions of how the occurrence or prevalence were quantified in their studies. For this reason, Table 1 presents definitions of each metric that are used in this study to properly classify information spread out in these 41 studies.

Table 1. Definitions of each metric

Metric		Definition
Fault occurrence metrics	Fault prevalence	Percentage of units with a given fault at a given severity and at a single point in time
	Fault incidence	How often a given fault occurs within a specified period of time
	Percentage of fault among all faults	Percentage of a specific fault incidence as a subset of a greater collection of faults

In this paper, we broadly call these metrics, “fault occurrence metrics.” The definition of fault prevalence includes the prevalence of faults at a single point in time; however, instantaneous fault prevalence rate is not as practical as the fault prevalence rate over a day or week. Therefore, fault prevalence defined in this study also includes assumptions that faults counted over a short time frame, such as two weeks, are all happening at a single point in time. Figure 1 shows an example of how these metrics are calculated differently in certain sample spaces: three buildings, two faults, and within a one-year period. The prevalence is calculated based on the specific period of interest. Therefore, the prevalence of fault 1 during the period of interest shown in the figure becomes 33% (one faulted building out of three total buildings). And the incidence of fault 2 becomes 1.3 incidents/year-building, because four incidents occurred in three total buildings within a year. The percentage of fault among all faults is mostly used in the literature for differentiating different types of faults among all service records.



	Fault 1	Fault 2
Prevalence	**1/3** (within period of interest)	**2/3** (within period of interest)
Incidence	**(1+1+1)/3** (incident/year-building)	**(2+0+2)/3** (incident/year-building)
Percentage of Fault Among All Faults	**43%** (3 faults out of total 7 faults)	**57%** (4 faults out of total 7 faults)

Figure 1. Example of fault occurrence metrics calculations

Figure 2 shows all 41 studies that captured at least one of these fault occurrence metrics; studies are grouped based on the information that each study includes. As shown in the figure, some of these studies referred to metrics from other studies, because their focus was not on measuring or quantifying fault occurrence metrics. They were either a review study for understanding the current knowledge (Braun 2003; Comstock et al. 2002; Comstock and Braun 1999; Hunt et al. 2010), a study estimating impacts of faults (Codes and Standards Enhancement [CASE] 2011; Djunaedy et al. 2011; Roth et al. 2004, 2005), or a study evaluating FDD tools (Farahmand et al. 2017; Heinemeier 2012; Wen and Li 2011; Zhao et al. 2017).

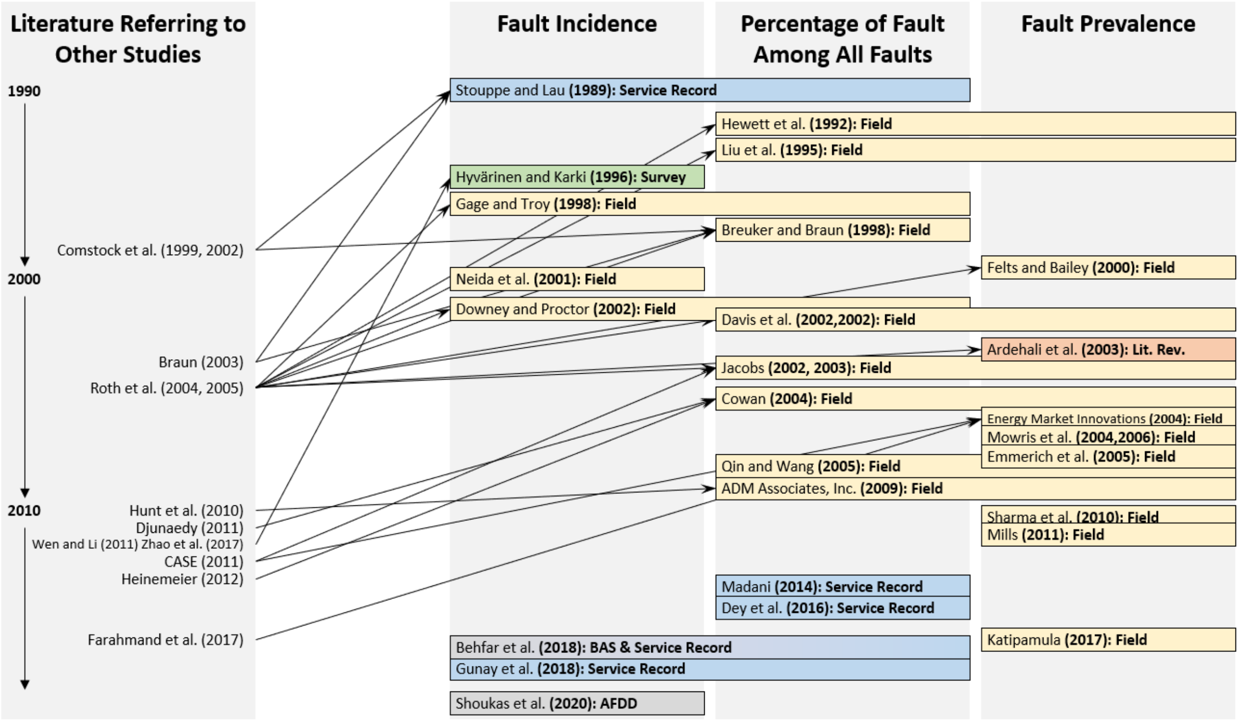


Figure 2. 41 studies (26 unique studies) capturing fault occurrence metrics (arrows are pointing towards the source reference)

As shown in Figure 2, each study captures different fault occurrence metrics depending on the purpose of the study. Additionally, the method (field measurement, service record, survey, building automation system [BAS] data, automated FDD tool data, or literature review) used for capturing fault occurrence metrics is indicated with text next to the citation and is also marked with different color. Although a total of 41 studies are mentioned in Figure 2, five of them are pairs of overlapping studies. For example, Mowris et al. (Mowris 2006; Mowris et al. 2004) has two studies: a conference paper that summarizes the findings and a technical report with more detail of data and calculations. These study pairs describe the research in total, and therefore are counted as a single unique study in this literature review. Pairs can be seen in Figure 2 as a single line item and two citations; 26 unique studies are found from the literature review.

2.2 Review of 26 unique studies

This section summarizes 26 studies that capture at least one of the fault occurrence metrics. The studies are grouped based on the common system type where faults mainly occurred.

Faults in heating and cooling systems

Stouppe and Lau (1989) characterized failures in air-conditioning (AC) and refrigeration systems in commercial buildings by collecting 8 years of data from an insurance company, including 15,760 failures. The authors documented failures in motors, fans, valves, and compressors, and summarized probable age at failure, major cause of failure, and failure prevention measures for these components. Based on various failures of service records that include the age at component failure and incidence, we can infer the percentage of failure among all failures. However, prevalence of individual failures at a certain point in time cannot be derived because the study does not specify the time when failure occurred for the 15,760 service records. The regional coverage of these service records is also unavailable.

Hewett et al. (1992) quantified energy savings that could be achieved through efficiency tune-ups on commercial unitary cooling equipment in a New England utility company's service territory in the U.S. The study conducted field measurements on 25 AC systems in 9 different sites, and focused on faults related to airflow, refrigerant charge, and duct leakage in smaller commercial buildings. The fault prevalence can be derived in this study by counting the number of units that were under faulty operation among the entire sample size (e.g., refrigerant leakage found in 18 out of 25 AC units). There is not enough information for deriving the fault incidence from this study because field inspections were performed relatively instantaneously rather than units being monitored under a longer period. Because only the small number of samples were selected due to the budget restriction, the study acknowledges that the samples might not be representing the condition of HVAC systems of the whole customer. However, the selection of HVAC system types was properly selected within the samples to represent most system types of the whole customer.

Breuker and Braun (1998) characterized common faults in rooftop units (RTU) and estimated their impact on energy consumption. Around 6,000 service records were gathered from a database owned by a service company that primarily services RTUs in commercial retail buildings. The study focused on illustrating the percentage of faults among all faults within the RTU system. The fault incidence cannot be derived from this study because the period took for collecting the service records is not specified. The regional coverage of these 6,000 service records is also not shown in this study, making it difficult to differentiate regional or climatic impact on fault occurrence metrics.

Felts and Bailey (2000) characterized the performance of 250 RTUs installed in small commercial buildings in northern California in the U.S. The authors summarized key insights on economizer operation, short cycling, and unit oversizing in terms of improving equipment operating efficiency. The entire monitoring was conducted within a 3-month period in the summer season and each

RTU was monitored for a three to 5-day period. The measurements included power, power factor, supply air, return air, mixed air, and outdoor air temperatures. Faults were measured by using a performance analysis tool detecting faults based on several sensor points. For example, low refrigerant charge faults were quantified by measuring the temperature difference between inlet and outlet of the evaporator coil. While the goal of the study is to represent the whole 450,000 RTU customers in northern California in the U.S., the study acknowledges that the number of samples covered in the measurements is not statistically representative. Prevalence of faults were quantified in this study, however, only summarized results (e.g., prevalence of oversizing of RTU) are presented, instead of individual monitoring results of the 250 RTUs.

Downey and Proctor (2002) focused on quantifying fault prevalence in AC systems in both commercial and residential sectors. Field measurements on 13,258 ACs in California in the U.S. were targeted, and the study focused on faults such as incorrect refrigerant charge and incorrect airflow. An assessment tool was developed in this study which takes evaporator inlet/outlet temperatures and refrigerant temperature/pressure as inputs to compare against manufacturer’s recommended airflow and refrigerant charge level. Systematic procedure was suggested to technicians to minimize any measurement biases during routine installation, repair, and maintenance visits. While measurements were taken over a relatively long period (26 months), the study summarizes findings with the fault prevalence metric (e.g., 57% of the entire AC systems had incorrect refrigerant charge level), and therefore the fault incidence cannot be inferred from this study.

Davis et al. (Davis, Baylon, et al. 2002; Davis, Francisco, et al. 2002) developed a procedure used to evaluate RTU performance in small commercial buildings and presented field measurement findings of applying energy efficiency measures on 30 RTUs in Oregon in the U.S. The existing tool developed by Downey and Proctor (2002) was used for detecting refrigerant charge faults. The coil cleaning was also considered in the field measurements because the calculation of the tool is based on clean heat exchangers on both condenser and evaporator. Incorrect airflows across evaporator and/or economizer were quantified by measuring the pressure drop, converting pressure drop to the airflow, and comparing the airflow against recommended airflow. This study only includes prevalence of faults related to refrigerant, evaporator airflow, and economizer operation.

Jacobs (2002, 2003) presented the underlying causes of faults or sub-optimum performance in commercial small package HVAC systems via field measurement of 215 units at 75 sites in California in the U.S. Physical inspections, series of one-time tests, and/or short-term monitoring (for two to three weeks) of unit performance were conducted up to four HVAC units per building in this study. Incorrect airflows were quantified by measuring the pressure drop across a plate installed at the filter location and converting it to the airflow. Refrigerant charge fault was also quantified with the same tool developed by Downey and Proctor (2002). The study provides projections of statewide energy savings when faults are properly addressed. Prevalence of faults (e.g., refrigerant charge, low airflow, economizer problems, etc.) were quantified; however, the other fault occurrence metrics cannot be derived.

Cowan (2004) characterized operational problems by combining five previous field measurement projects that include a total of 503 RTUs in 181 commercial buildings across five states in the U.S.—Oregon, Washington, Idaho, Montana, and California. One of the five previous field studies is the study done by Jacobs (2002, 2003) described previously and all these projects performed field evaluations on RTUs. The study acknowledges the protocol for evaluating large portions of RTUs were not defined strictly and the procedure evolved over time which affected the quality of the data. The study summarizes key RTU problem areas (refrigerant charge, economizer, airflow, thermostats, and sensors), quantifies fault prevalence related to these areas, and estimates potential energy savings.

A report written by Energy Market Innovations (2004) includes information on the AirCare Plus Program, which was initially led by Pacific Gas and Electric Company (PG&E). This program provides no-cost diagnostic HVAC tune-up services to commercial customers, including field assessments. Although the report itself did not include values of any fault occurrence metrics, studies (Codes and Standards Enhancement 2011; Farahmand et al. 2017) referring to this program (shown in Figure 2) include prevalence (although it incorrectly uses the term incidence) of faults related to sensors and economizers. The program is promising from the standpoint of quantifying fault occurrence metrics, however, there is not enough public information (e.g., sample space coverage) available to fully understand the data collected in this program.

Mowris et al. (2004; 2006) implemented an evaluation, measurement, and verification program that verifies the refrigerant charge and airflow in AC units in both commercial and residential sectors in California in the U.S. Training and pre-/post-interviews were conducted on technicians from participating contractors to efficiently evaluate the equipment and to minimize measurement biases. Incentives were given for the first 12,000 AC unit evaluations and the total number of units evaluated by the contractors resulted in 12,453 AC units. Measurement procedures for quantifying faults were not clearly described in the earlier work (Mowris et al. 2004), however, the later field study (Mowris et al. 2006) describes more details on how temperature, pressure, airflow, and power were measured in multiple locations to derive the performance of the AC unit. Through this program, the refrigerant charge and airflow of these AC systems were adjusted to optimize system performance. The study included prevalence of these faults; however, the other fault occurrence metrics cannot be inferred.

ADM Associates, Inc. (2009) performed field measurements of AC unit performance in residential buildings and assessed the effects of proper system servicing. The field measurements focused on 109 packaged AC systems in residential buildings in southern California in the U.S. Standard set of measurement points was pre-defined before taking actual measurements from each AC unit and specific measurement protocols under certain operating conditions were also described in detail. Additional screening of the measured data was also performed using equations based on physics (e.g., air side measurement verification against psychrometric equations). AC units were selected among participants in the utility's demand response program which might include sampling bias. The prevalence of various faults (e.g., inefficient compressor, refrigerant flow restriction, condenser fouling, evaporator fouling, refrigerant charge, and insufficient airflow) were quantified in this study.

Mills (2011) conducted a meta-data analysis on a combined data gathered from the commissioning community, actual monitoring-based commissioning projects, and projects in literature which covered 643 non-residential buildings from 37 commissioning providers. The study analyzed how much commissioning cost, how much energy was saved, and how long the payback took in past commissioning projects on new or existing buildings by looking into real commissioning data. Because commissioning of existing buildings involves fixing deficiencies in buildings, deficiencies (or faults) around major building components (heating and cooling, lighting, envelope, plug loads, etc.) are also quantified in this study. However, the deficiencies were described only based on the system level (e.g., heating and cooling, lighting, plug load, etc.) and specific reasons (or root cause) were not provided. While prevalence of deficiencies are quantified, incidence of deficiencies cannot be derived from the available information.

Madani (2014) studied common and costly faults that occur in heat pump systems in both commercial and residential sectors. 37,000 fault reports from manufacturers, as well as 8,659 fault reports from an insurance company in Sweden, were collected to characterize faults in specific components (e.g., fan, controller, valve, compressor, refrigerant circuit, etc.) in the heat pump system. Only the percentages of individual faults among all faults are presented in this study.

Dey et al. (2016) developed a method based on Bayesian Belief Network (BBN) which can be applied along with the rule-based FDD to not only detect faults with rule-based FDD but also to diagnose faults with the BBN method. The proposed method was applied to an actual university building in Texas in the U.S. and percentages of faults among all faults were also quantified from maintenance records of 1 year. Faults that occur in heating coils, cooling coils, mixing boxes, controllers, sensors were quantified.

Pacific Northwest National Laboratory has developed a commercial building retuning approach. Katipamula (2017) documented and analyzed trend data for 99 buildings across 26 states in the U.S. In this study, the author classified retuning measures based on energy savings potential and level of effort, performed metadata analysis for correlating measures with building metadata (e.g., region, vintage, building type, size, etc.), and documented the prevalence of various types of measurements. Because some of these measures are solutions for faults (e.g., fixing broken dampers), prevalence of faults can be inferred from this study.

Gunay et al. (2018) studied the frequency of faults by collecting building maintenance records and applying a text-mining technique to extract information on failure patterns in building systems and components. The basis for the text mining was 26,992 HVAC-related service records collected over 7 years for 44 buildings and 2 years of service records for the central heating and cooling plant in a university campus in Canada. The number of warning or failure instances (fault incidence) during the sampling period as well as percentage of individual warnings and failures among all service records were quantified in this study.

Shoukas et al. (2020) collected massive amounts of AFDD data from four different companies to understand how AFDD operates on RTUs, the types and frequencies of faults identified, and how building operators interact with these systems. This is a well-design study with appropriate sample space classifications which combined data covering 28,000 RTUs, five different building types, and multiple climate zones in the U.S. While most of the other studies that quantify fault incidence rates provide the number of fault incidents during a certain period, this study provides the duration (in hours) of faults during the monitoring period. This type of format is common in automated FDD tools where the duration of fault is logged until it is properly addressed by the building operator. However, because the fault incidence was defined in Section 2.1 based on the incident and not with time, an assumption of converting 24 hours of duration into 1 incident was made to combine results with other studies. This assumption is definitely not a correct conversion because the study also mentioned faults that occur in the economizer were not fixed for 80 days in average. This assumption should be noted to readers and the actual study should be referred for more accurate information.

Liu et al. (1995) focused on air handling units (AHU), especially improving supply air temperature control and recommissioning terminal boxes for improving building efficiency by the request of the building owner. Field measurements were performed on a hospital building in Texas in the U.S. that include 3 AHUs and 210 terminal boxes (out of total 248 terminal boxes). The total number of samples represents most of the terminal boxes; however, the field measurement is only done on one building. Faults in the terminal boxes were detected by comparing discharge temperatures between heating and cooling modes without any biased approach for quantifying faults. While the fault prevalence of faults in AHU can be derived, the fault incidence cannot be inferred with the available information.

Qin and Wang (2005) conducted a site survey over 14 days in a commercial building in Hong Kong. Strategies of automated FDD with hybrid approach were studied and applied to detecting faults in 261 variable air volume terminals. The study presents a summary table that includes the percentages of faults among all faults and how many times each fault occurred. Based on this information, fault incidence can also be inferred for the sampling period.

Hyvärinen and Karki (1996) identified common faults in various types of HVAC and refrigeration systems in commercial buildings. Although engineering judgements from a total of 71 experts were used for prioritizing faults in AHUs, heat pumps, and chillers (in terms of the occurrence frequency of faults), the study does not include values or quantification of fault incidence. Instead, the study provided three metrics—high, medium, and low—of how frequently faults occur in HVAC and refrigeration systems.

Faults in commercial refrigeration systems

Gage and Troy (1998) focused on large commercial refrigeration systems in supermarkets, and studied methods for reducing refrigerant emissions. The study includes service records over a 1-year period from 110 supermarkets, to help understand the type of faults that are common in refrigeration systems. Based on services performed on various components (e.g., condenser, expansion valve, etc.) within the period, the fault incidence can be derived on a yearly basis for faults that occur within each component. The study also includes a summarized table that shows the percentages of faults among all faults collected during the sampling period.

Behfar et al. (2018) especially focused on system characteristics and operating faults in supermarkets. Data sources such as experts surveys, facility management system messages, service calls, and service records were gathered to investigate equipment characteristics (system type, condenser type, control type, etc.) and common operating faults (refrigerant leakage, failed evaporator, failed condenser fans, failed compressors, etc.). This is another well-designed study with proper sample space classifications but where the focus is only on supermarkets. While data collected through service calls and records mostly provided information on the percentage of faults among all faults, data collected through building management systems provided fault incident rates of various faults in the refrigeration systems across 18 buildings with 2 years of measurement period.

Faults in other building systems

Neida et al. (2001) quantified energy and cost savings potential when occupancy sensors were used for controlling commercial lighting systems. Field measurements were taken from 60 different organizations across 24 states in the U.S. to properly represent the diversity within the commercial building stock. Although the focus of the study is not specifically related to quantifying lighting system faults, the study quantified the percentage of lights left on when the spaces were unoccupied. This can be translated to the prevalence of a lighting control fault.

Ardehali et al. (2003) focused on control systems in commercial buildings by conducting a literature review of case studies (from a total of 118 buildings) that documented the correlation between inefficiencies in buildings and problems associated with controls and direct digital control systems. The regional coverage of 118 buildings is not included, and neither is the duration of measurements taken. The study only summarizes the percentage of various control related faults among all faults.

Emmerich et al. (2005) investigated the impact of commercial buildings’ envelope airtightness on building energy usage. The study referred to a data set capturing airtightness levels for 166 buildings—144 buildings in the U.S. and 22 buildings in the United Kingdom. Based on this data set, only 6% of buildings were meeting the target airtightness level. Because infiltration through the building envelope can cause significant increase in building energy usage, the airtightness of the building is considered to be a fault in this literature review.

Sharma et al. (2010) studied detection methods as well as prevalence of sensor faults in real world data sets. Generic faults in any sensors such as frozen sensor reading and extreme noise in readings were considered by analyzing in the component level. And information on the sensor application (e.g., in what system) was not provided. The data set used for quantifying fault prevalence covered a maximum 6 months of measurement period, however, no additional information (e.g., where and how it was collected) on the data set was provided in the study.

In general, studies that quantified at least one of the fault occurrence metrics provide insights on how frequent and prevalent faults can be in buildings, but most of the studies lack a common and systematic procedure of quantifying fault occurrence metrics that can be properly compared between studies and be understood as true fault occurrence metrics. Further, information on the entire correlation between fault occurrence metrics and various parameters (e.g., climate zone, building type, equipment type, equipment age, maintenance level, etc.) is very sparse in these studies. Within this context, the aim of this literature review is to summarize these gaps from the 26 unique studies; these gaps are summarized in the following subsections.

2.3 Sample space coverage

Samples (e.g., RTU, AHU, insurance claims, service records, etc.) were selected in the previous studies before the samples were evaluated and faults of interest were quantified. However, inappropriate samples lead to biased or skewed analysis results. Thus, it is necessary to understand the type and range of sample spaces covered in these previous studies.

Figure 3 shows the counts of individual studies (on the left) according to the building types (if available) that were included in their entire samples and shows the U.S. commercial building stock characteristics (on the right) such as the number of buildings, electricity consumption, and natural gas consumption for different building types based on Commercial Buildings Energy Consumption Survey (Energy Information Administration 2002). To make an impact with limited resources, the number of samples for each building type can be selected (e.g., more offices and less church) based on the energy consumption as shown in the building characteristics figure. The previous studies as a whole are covering relevant numbers of each building type as shown by the correlation between the number of buildings/energy consumption for a particular building type and the number of studies that address that building type. However, only some of the previous studies (Felts and Bailey 2000; Neida et al. 2001; Davis, Baylon, et al. 2002; Davis, Francisco, et al. 2002; Jacobs 2003; Mowris et al. 2004; 2006; Katipamula 2017) include wide range of building types in each of their study.

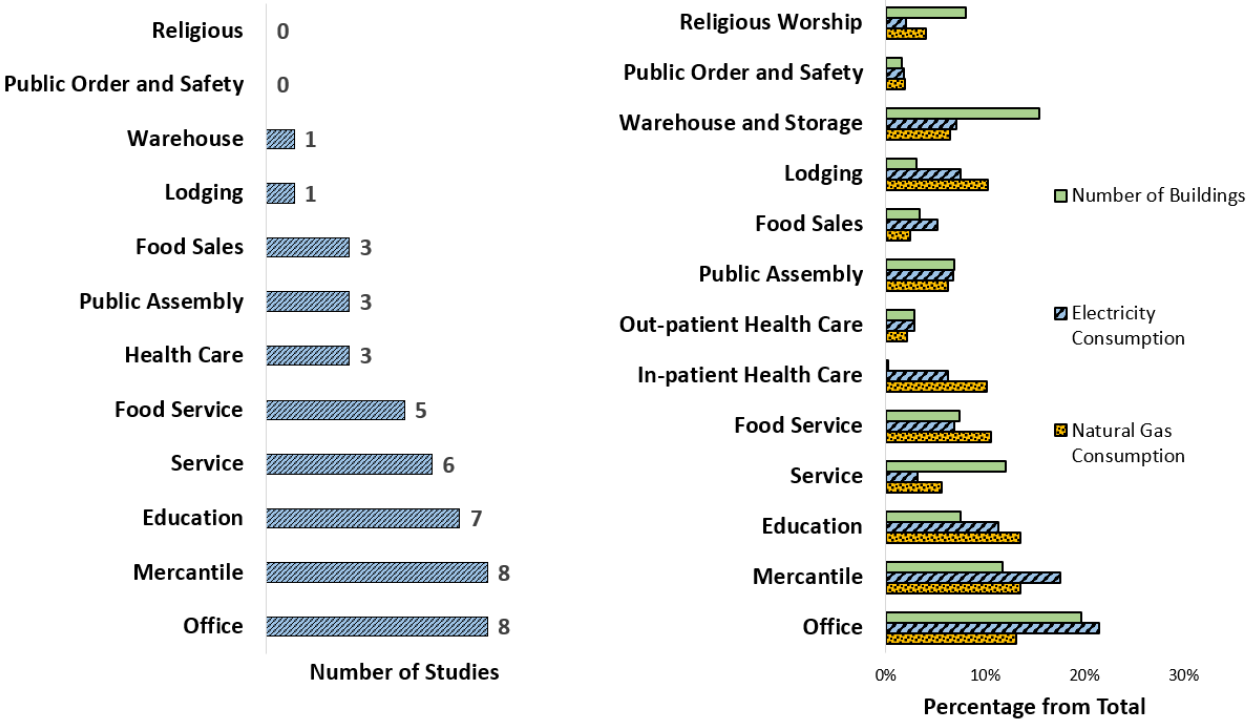


Figure 3. Sample space covered in 26 unique studies: building type

Figure 4 shows the comparison between counts of the U.S. based studies (on the left) on each U.S. state (if available) based on the sample buildings' locations and the corresponding annual source energy consumption (on the right) in each state for the entire commercial buildings in the U.S. based on Commercial Buildings Energy Consumption Survey (Energy Information Administration 2017). In total, the studies cover the majority of the states that correspond to most of the climate zones defined by ICC (2012) and focus on regions with higher impact (California, Texas, Florida) in terms of energy consumption. While some of these studies collectively cover a wide range of regions across the country (Neida et al. 2001; Emmerich et al. 2005; Katipamula 2017), most of the other studies mostly focus on local regions where quantified results can be climate specific. Variation in climate affects building system operation significantly, especially for heating and cooling systems (differences in equipment runtime result in different rates of recorded fault occurrence), thus, appropriate climate zones should also be considered for a proper sampling.

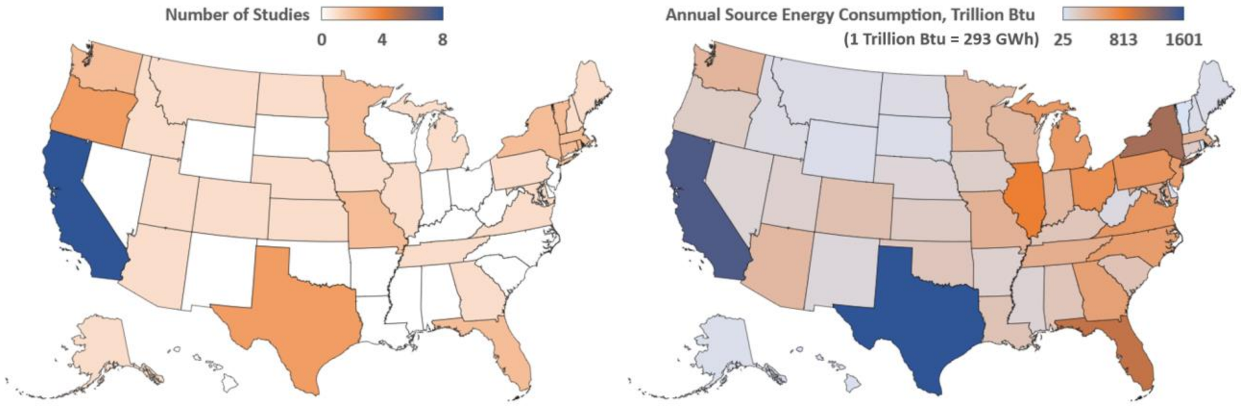


Figure 4. Sample space covered in 26 unique studies: U.S. states

Studies that cover fault occurrence metrics gathered data from two main data sources: field measurements and service records. Field measurements are typically obtained through system monitoring (Felts and Bailey 2000; Hewett et al. 1992; Katipamula 2017; Neida et al. 2001) or technician inspection (ADM Associates, Inc. 2009; Cowan 2004; Davis, Baylon, et al. 2002; Davis, Francisco, et al. 2002; Downey and Proctor 2002; Jacobs 2003; Katipamula 2017; Liu et al. 1995; Mowris 2006; Qin and Wang 2005). Only one of these field measurements (Liu et al. 1995) was conducted by the request of the building owner while the other measurements were initiated by the local utility's incentive program. The field measurement done by ADM Associates, Inc. (2009) includes relatively detailed protocols and procedures to efficiently evaluate equipment and to minimize measurement biases. However, uncertainties and biases in these field measurements still exist and especially when the purpose of evaluation is being incentivized and reimbursed as pointed out from Close (2010).

Service record sources include reports from building maintenance records (Farahmand et al. 2017; Gunay et al. 2018), insurance companies (Madani 2014), service companies (Breuker et al. 2000; Gage and Troy 1998), and manufacturers (Madani 2014; Stouppe and Lau 1989). While faults documented through service records provide definite needs of equipment fix or replacement, this type of data cannot provide prevalence of soft faults—where the fault severity increases over time—that evolve slowly over time (e.g., condenser fouling) without causing significant harm to the operation of the system.

Other sampling characteristics (e.g., number of samples, type of samples, and the data collection period) also vary significantly between different studies. The data collection period varied from 4 hours of inspection (Davis, Baylon, et al. 2002; Davis, Francisco, et al. 2002) to 10 years of recording fault reports (Madani 2014). Further, although the type of samples varied between service record, component, system, and building, the number of samples varied from 1 (Dey et al. 2016) to 371 (Shoukas et al. 2020) buildings, 15 (Shoukas et al. 2020) to 25,800 (Shoukas et al. 2020) vapor compression systems, and 100 (Behfar et al. 2018) to 177,240 (Behfar et al. 2018) service records.

2.4 Categorization of faults

In this section, the hierarchy of the compiled faults is analyzed to understand how faults are commonly classified. Figure 5[2] includes classification of 112 faults from the 26 studies based on available information from these studies. In this figure, fault types (outermost circle in the figure) are associated with component types, equipment types, and system types (innermost circle in the figure), representing how faults are classified in the literature. Portions highlighted in gray indicate unspecified fault or equipment type where information was not available from the literature.

---

[2] An interactive plot (in html format) of this figure is also included as a part of the supplementary material. It is possible to zoom-in on a certain pie from this interactive plot by opening the html file on a web browser and clicking on a certain pie to see the raw categorization of faults in detail.

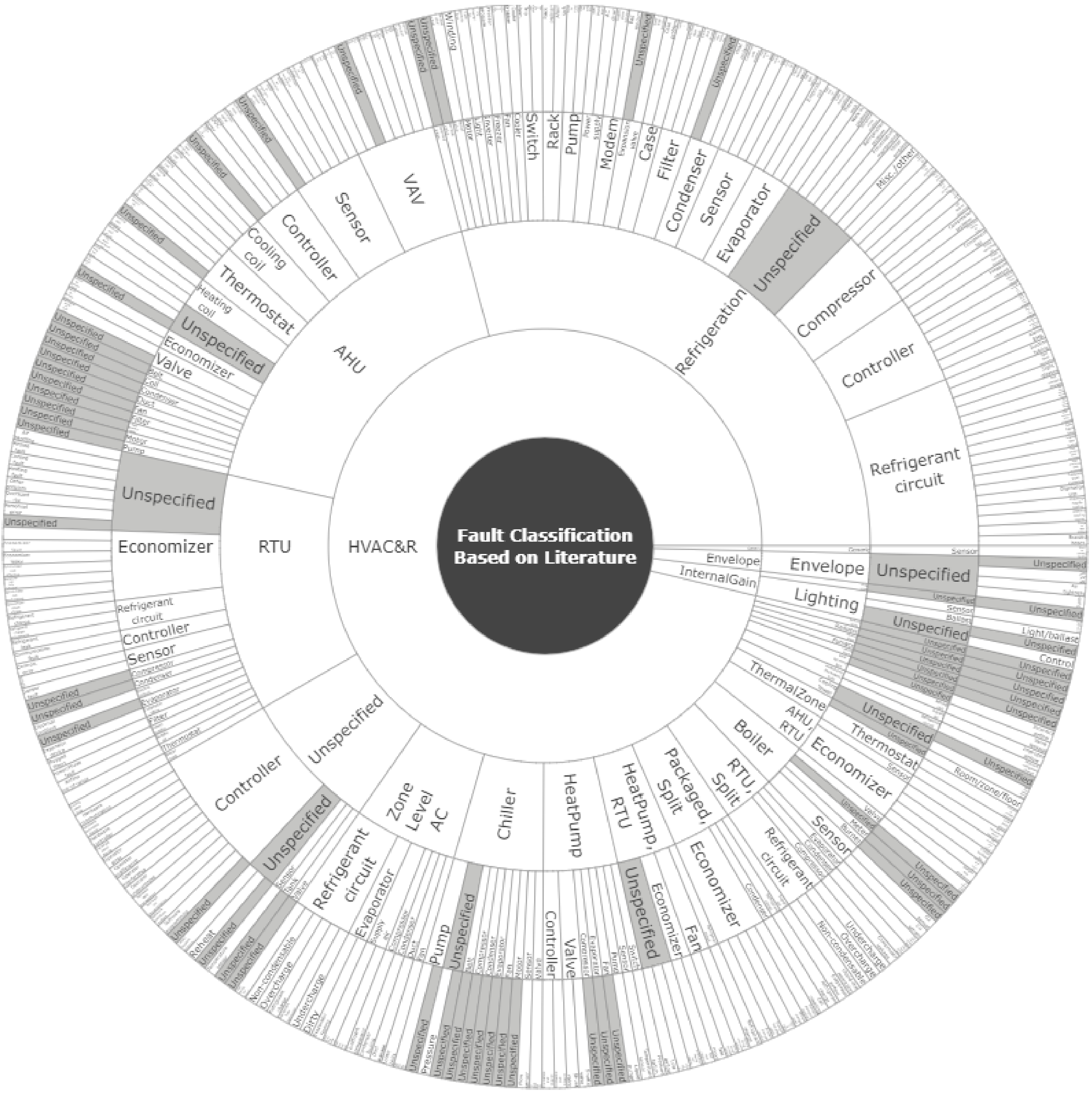


Figure 5. Classification of faults from 26 unique studies

Most of the faults shown in this figure map to HVAC and refrigeration systems, but others[3] map to building envelope, thermal zone, and lighting. Because significant portions of the literature fail to properly specify equipment, component, and fault type, best estimations were made from the information given in studies to properly assign equipment/component/fault types. For example, ADM Associates, Inc. (2009) inspected 109 AC units in both residential and commercial building sectors; however, the detailed equipment type (i.e., whether it is based on cooling only unit or heat

---

[3] Although the focus of this study is on HVAC systems, information gathered during the search process for other systems in buildings are also documented.

pump) are not specified. Thus, the equipment type for this study is denoted as zone level AC to indicate both correct information and limitation of the study. Similarly, Jacobs (2002, 2003) conducted field measurements on a total of 215 "small packaged HVAC systems" at 75 sites. Although the "small packaged HVAC system" category is defined as "single packaged RTUs or residential heat pumps with cooling capacity of 10 tons or less," the measurement results were not attributed to either RTU or heat pump configurations. Equipment types that were not properly differentiated in the study were therefore combined. Faults denoted as "unspecified" mostly map to service records for which only the equipment or component type of the fault is reported.

As it is clearly shown in Figure 5, a comprehensive schema or taxonomy is required for quantifying faults so that data across various systems and faults can be properly aggregated, compared, and exchanged between different analyses. This will also benefit the process of improving the value of quantified data, minimizing the effort for quantifying fault occurrence metrics, and facilitating energy efficiency in buildings.

2.5 Most frequently studied faults and reported fault occurrence metrics

Figure 6 and 7 present ranges of fault prevalence and incidence rates for various faults derived from the 26 studies as well as number of studies quantifying each fault. In these figures, instead of using raw fault descriptions from the literature, grouping of faults was made to compare rates of the same group of faults between different studies. However, the grouping has its own limitation based on the available information. For example, faults that are described as fouling on evaporators were grouped under "Evaporator fouling" while faults that were only described as improper airflow were grouped as "Improper airflow". These are not completely separate faults because one of the reasons for improper airflow in the AHU can be caused by the fouling on the evaporator. Additionally, faults are also classified with the raw equipment level information in these figures. Although there are many faults where the fault prevalence or incidence are quantified only once within all 26 studies (shown as a single tick mark in the figures), there are several faults for which prevalence and/or incidence were quantified by multiple studies. These are shown in the figure as box plots with ranges (e.g., mean, minimum, maximum, first and third quartiles) specified.

As shown in the figures, both fault prevalence and incidence vary significantly between different studies. For example, the fault prevalence of improper charge in refrigerant circuits is reported between 30% and 70% between 8 different studies. It is difficult to make a true comparison between the metric across the studies, because the severity of the fault (e.g., percentage of refrigerant charge deviation compared to normal amount) is unknown and potentially inconsistent. Faults quantified from field measurements supported under utility's incentive program can also skew (increase) prevalence and incidence rates compared to ground truth rates because the detection of faults are incentivized or reimbursed.

Additionally, the classification of faults used in these studies is also limited in terms of verifying the root cause of the fault. For example, the fault commonly classified as "improper airflow" in the figures can result in insufficient or excessive heating and cooling capacity affecting energy consumption and thermal comfort. However, the root cause can stem from different sources; an incorrect supply air fan setting or poor practice in air balancing can cause higher airflow, whereas poor practice in air balancing, duct leakage, or fouling on the evaporator or in other parts of the duct can cause lower airflow.

Maintenance levels can also have a significant effect on fault occurrence metrics. Some buildings have a building manager and regularly scheduled maintenance, but others may only receive maintenance when there is an issue. Most of these studies do not document quality of maintenance. Soft faults can often be easy to identify and mitigate through regular maintenance (e.g., scheduled evaporator cleaning can eliminate evaporator fouling). Thus, classification of the building system and fault, fault severity definitions, and documentation of maintenance level are keyways in which prior studies vary in their content.

The summarized findings in Figure 6 and 7 are not to inform readers about true rates of fault prevalence and incidence. Instead, this figure is to emphasize how sparse, uncertain, and deviating these values are and why a more comprehensive and accurate methodology is required for quantifying fault prevalence and incidence.

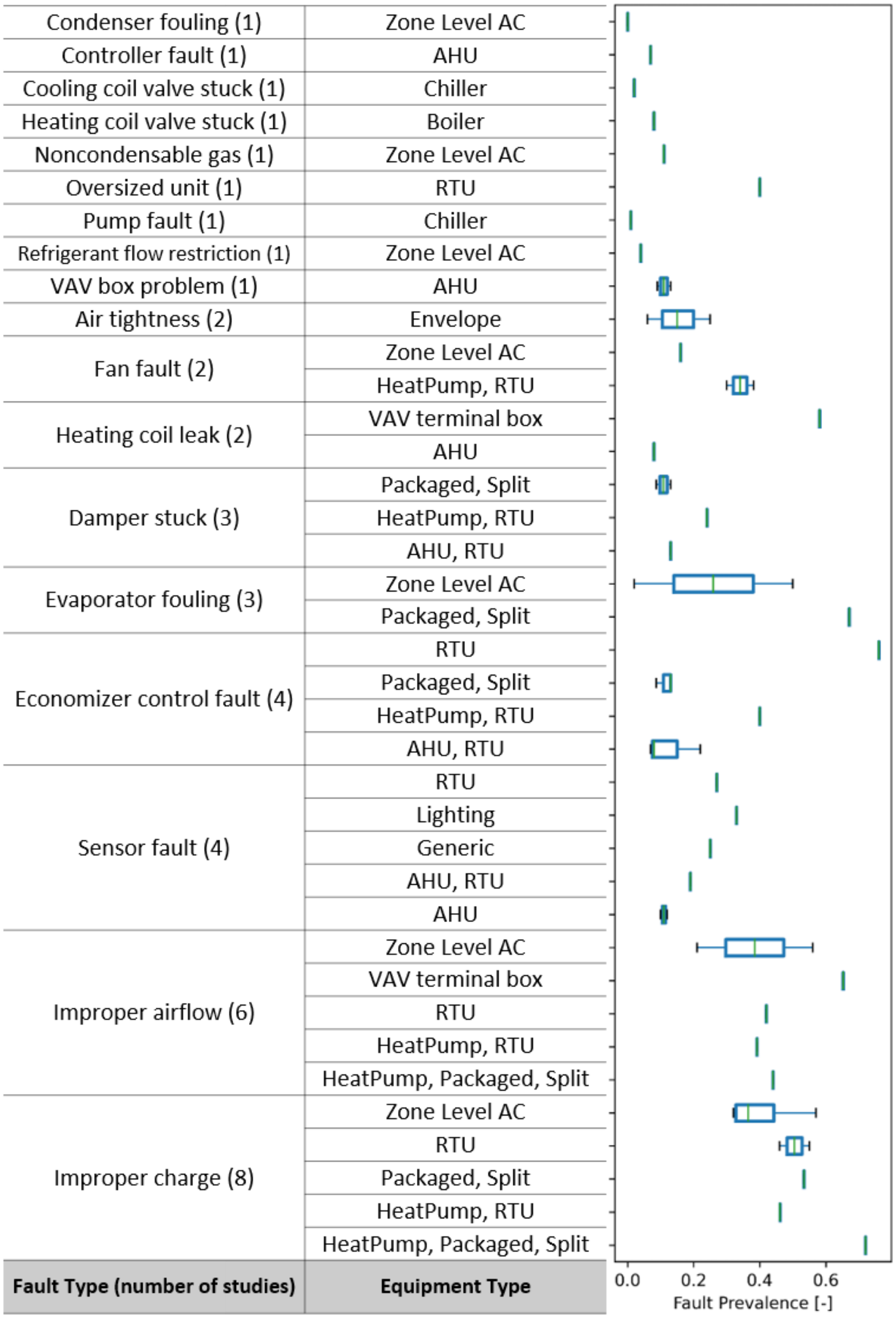
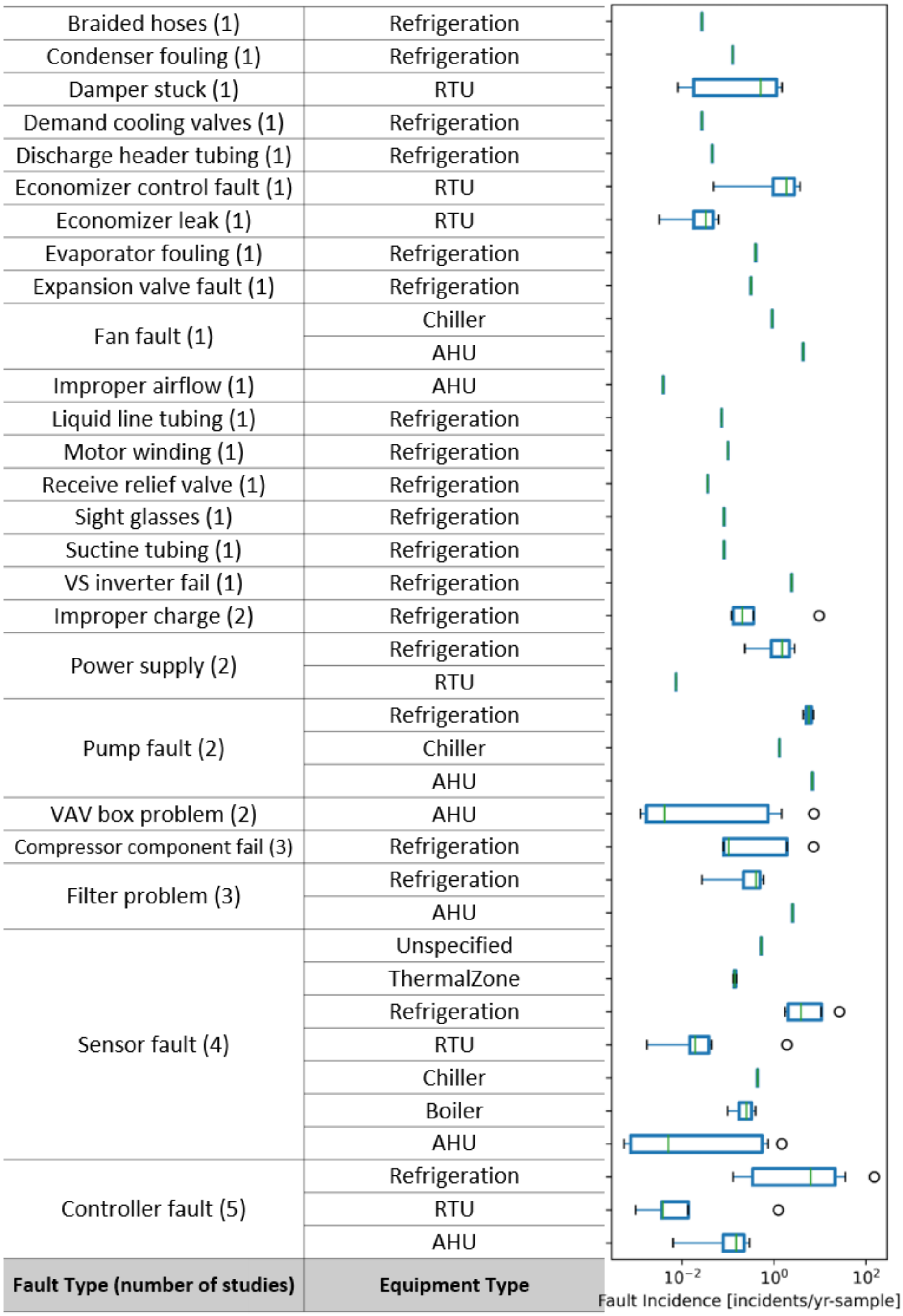


Figure 6. Reported range of fault prevalence for faults associated with each equipment including number of studies quantified each fault



Fault Type (number of studies)	Equipment Type
Braided hoses (1)	Refrigeration
Condenser fouling (1)	Refrigeration
Damper stuck (1)	RTU
Demand cooling valves (1)	Refrigeration
Discharge header tubing (1)	Refrigeration
Economizer control fault (1)	RTU
Economizer leak (1)	RTU
Evaporator fouling (1)	Refrigeration
Expansion valve fault (1)	Refrigeration
Fan fault (1)	Chiller
	AHU
Improper airflow (1)	AHU
Liquid line tubing (1)	Refrigeration
Motor winding (1)	Refrigeration
Receive relief valve (1)	Refrigeration
Sight glasses (1)	Refrigeration
Suctine tubing (1)	Refrigeration
VS inverter fail (1)	Refrigeration
Improper charge (2)	Refrigeration
Power supply (2)	Refrigeration
	RTU
Pump fault (2)	Refrigeration
	Chiller
	AHU
VAV box problem (2)	AHU
Compressor component fail (3)	Refrigeration
Filter problem (3)	Refrigeration
	AHU
Sensor fault (4)	Unspecified
	ThermalZone
	Refrigeration
	RTU
	Chiller
	Boiler
	AHU
Controller fault (5)	Refrigeration
	RTU
	AHU

Figure 7. Reported range of fault incidence for faults associated with each equipment including number of studies quantified each fault

3. Interview with subject matter experts and stakeholders

To complement and supplement the literature review, we conducted interviews with subject matter experts from academia and industry. Specifically, the interviews were designed to provide insights into three primary topics: (1) the value proposition and needs for continued research in fault prevalence; (2) the current state of knowledge regarding fault prevalence in commercial buildings; and (3) addressing current knowledge gaps.

We conducted interviews with subject matter experts from academia and industry to complement the literature review and to address key needs of the FDD community. Findings from the interviews are presented in this section regarding the value proposition of further study on fault prevalence, knowledge gaps in the understanding of fault prevalence, and perspectives on the design and scope of future studies concerning fault prevalence. Overall, the findings indicate that there is strong value in further study for both the research community and industry, that there remain significant gaps in what we know about fault prevalence in commercial buildings, and that we need common approaches in order to efficiently collect fault information and analyze fault prevalence.

3.1 Methodology

The interviews were conducted using a questionnaire to guide the discussion. Reflecting the primary topics of focus, the questionnaire was divided into three corresponding sections, covering a total of 24 questions. Seven questions pertained to value proposition, four to the state of knowledge and 13 to recommendations for addressing identified gaps. For most of the interview questions, interviewees were asked to elaborate on their responses. The questionnaire can be found in the Supplementary Material.

The questions took on three forms. The most common form was *ranked scale,* in which experts were asked to evaluate a set of options on a numeric scale, with the highest number representing the most important or best option. For example, interviewees were asked, “On a scale of 1 to 5, where 1 is least important and 5 is most important, how important is understanding fault prevalence compared with other topics of study in the field of FDD?” Ranked-scale questions were presented with both even and odd numbers of options to intentionally enable or disallow, respectively, a neutral response. The second form of question was *structured choice*, in which an unranked set of options was presented for selection. For example, one question asked, “What is the current state of understanding of fault prevalence in commercial buildings? Mark all that apply.” The options included “weak anecdotal,” “strong anecdotal,” “limited empirical,” and “substantial empirical.” Ranked-scale and structured choice questions were often followed with a third form, a *free response* question, to invite explanation or further comment—for example, “Please talk us through your responses.”

To minimize the possible bias from different interviewee types, a wide cross section of subject matter experts was targeted. The cohort of participants included 25 individuals; this comprised researchers (7 individuals), FDD providers (5 individuals), FDD end users (4 individuals), efficiency service providers[4] (7 individuals), and utilities (2 individuals). These 25 individuals engaged in 24 interview sessions (one session included two individuals). Additionally, a pilot interview was conducted (before the 24 interview sessions) to trial run the questions to avoid misleading or unclear questions as well as to confirm the ability to cover all topics in a one-hour discussion. The wording of the questions was refined based on the trial, and the full set of interviews took place in February and March 2019. The full set of responses was synthesized into key findings, as presented in Section 3.2 to 3.4. And interview questions were sent in advance of each session so that interviewees could review them if desired.

3.2 Value proposition

We asked the subject matter experts two multipart questions to understand the value proposition associated with studying fault prevalence. The first question is: "On a scale of 1 to 5, where 1 is less important and 5 is the more important, how important is understanding fault prevalence compared with other topics of study in the field of FDD?" The median value for each respondent type, and across all respondents, is shown in Figure 8. Overall, the median ranking was 4, indicating a high value of the fault prevalence topic. Researchers and FDD technology providers saw the topic as most valuable, whereas end users of FDD, efficiency service providers, and utility ranked it as relatively less of a priority. Interestingly, one researcher noted: "This is important because there is really nothing out there. Justifying investment is difficult without having good data."

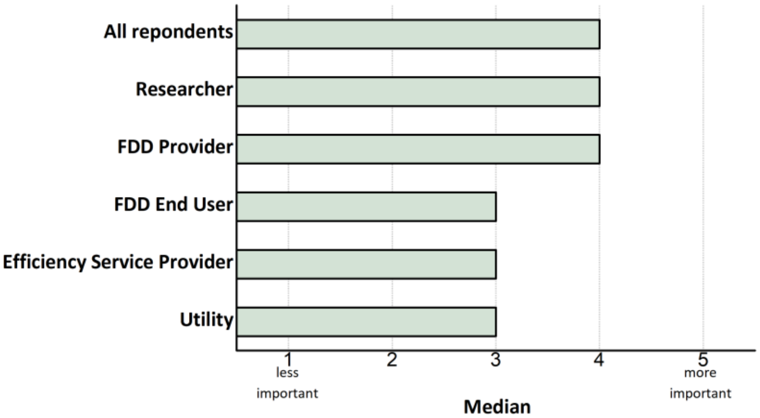


Figure 8. Median importance ranking of fault prevalence study by respondent group

The second multipart question presented respondents with a list of potential elements of what a refined understanding of fault prevalence might enable, including the ability to:

- Develop improved FDD algorithms for the most important faults;

---

[4] The term "efficiency service provider" used in this study includes companies that provides HVAC engineering, retro-commissioning (RCx), and efficiency implementation.

- Develop improved metrics to assess FDD algorithm performance;
- Quantify impact estimates to support the business case for adoption of FDD technology and processes;
- Prioritize corrective action;
- Target effective use of operations and maintenance labor; and
- Prioritize monitoring instrumentation.

Experts were asked to indicate the value of each element on a scale of 1–4. One set of rankings was provided considering the needs of the research community, and another set was provided considering the needs of industry. Shown in Figure 9, the results indicate that there is strong multidimensional value to both the research and industry communities. The two elements experts deemed most important for the research community were the ability to develop improved algorithms and the ability to develop improved metrics to assess algorithm performance. For industry, the top two were the ability to quantify impact to support the business case for FDD adoption and to target effective use of operations and maintenance labor.

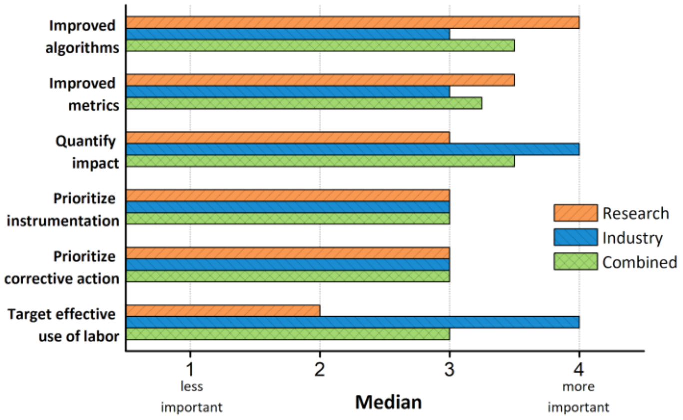


Figure 9. Median importance rankings for value elements

3.3 Knowledge gaps

Two interview questions focused on surfacing knowledge gaps with respect to the prevalence of faults in commercial buildings. In the first, experts were asked, "What is the current state of understanding of fault prevalence in commercial buildings? Mark all that apply [weak anecdotal; strong anecdotal; limited empirical; substantial empirical]." Anecdotally referred to information gleaned from observations of commercial building operations, whereas empirically referred to published studies. Figure 10 shows that experts confirmed that there is a weak empirical understanding of fault prevalence; they were evenly split as to whether the anecdotal understanding is weak or strong. This is consistent with the literature review findings—only 26 studies addressed fault occurrence metrics *at all*, and across these 26 studies, diverse metrics, definitions, and methodologies were used, preventing synthesis for rigorous conclusions.

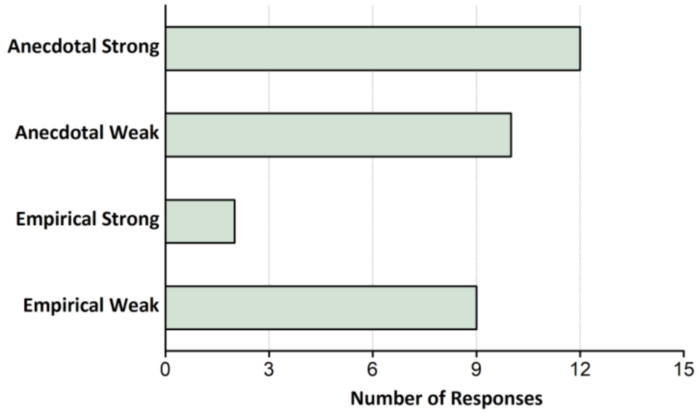


Figure 10. Number of responses on state of fault prevalence information[5]

The second question presented experts with a list of specific potential gaps. These included:

- Fault prevalence for specific equipment types (e.g., fans or valves);
- Fault prevalence for specific system types (e.g., AHUs or RTUs);
- Fault prevalence for specific faults;
- Conditions associated with fault occurrence and intermittency (e.g., operational mode, seasonality);
- Fault prevalence based on commercial building type;
- Fault prevalence based on climate zone;
- Economic impact of specific faults; and
- Other, please specify.

Experts were then asked to "indicate your thoughts on the importance of addressing each gap." These gap areas were all rated at a median value of 2 or higher on a scale of 1 to 3, as shown in Figure 11. Overall, fault prevalence for specific system types, conditions associated with fault occurrence and intermittency, and economic impacts of specific faults were ranked highest in importance. Experts typically used this question to talk about the importance of focusing on specific faults rather than focusing on faults tied to building or system types. This may be a case where respondents had difficulty tying their answers specifically to fault prevalence as defined in this study, and therefore, the results should be viewed cautiously.

---

[5] Some respondents selected more than one option.

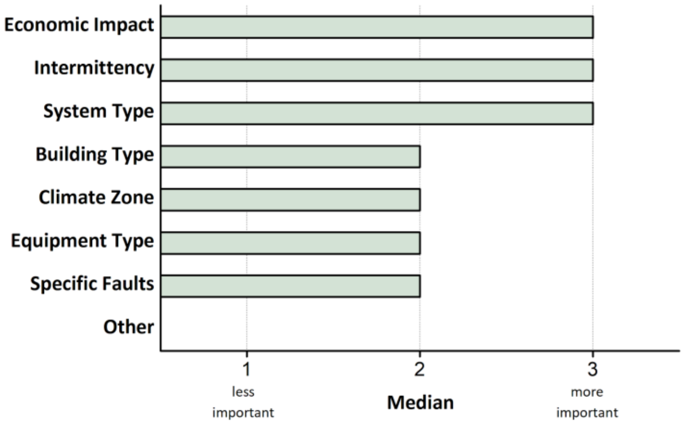


Figure 11. Median importance rankings of specific knowledge gaps concerning fault prevalence

3.4 Addressing knowledge gaps

This set of interview questions targeted subject matter experts' perspectives on tractable approaches to address existing knowledge gaps related to fault prevalence. To understand high-priority focus areas for expanding the current state of knowledge, experts were asked to rank seven parameters to identify which are important to capture a wide variety of. A ranking scale from 1 (least important) to 4 (most important) was used to evaluate the importance. Figure 12 illustrates ranking results for each parameter.

With median rankings of 4, experts felt that it is more important to span a large range of fault types and system types than it is to span a diversity of building types, diversity of equipment, climate zones, building age/condition, or ownership and management models. This illustrates a consensus that fault types (and prevalence) are fairly consistent across building types (assuming that they used the same system or equipment types) and across regions. As an example, many experts brought up the example of air handling units as an equipment type that is widely used across regions and issues with their economizers as a common fault.

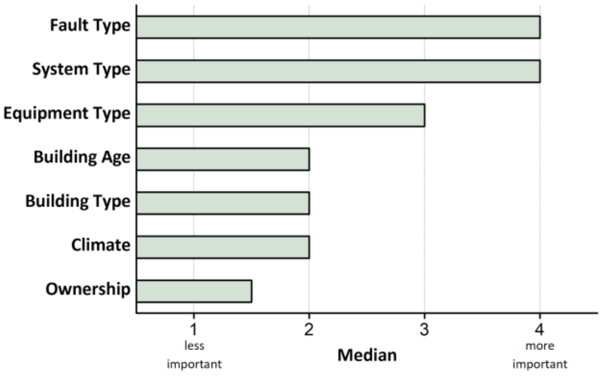


Figure 12. Median importance rankings of study sample focus area

When asked what system or fault types are most important to understand, experts generally referred to the most common systems and their components, such as variable air volume terminals, air handlers, chillers, boilers, and RTUs. As far as fault types are concerned, faults in economizers such as the damper stuck were often brought up as being problematic.

Additional questions focused on how data should be collected to address identified knowledge gaps, including actors to be engaged and sources to be mined. Figure 13 illustrates that experts see strong value in engaging FDD providers and vendors to obtain expanded sets of data on fault prevalence. On a scale of 1 (least value) to 3 (most value), the median ranking for FDD vendors was 3. The other options all received median rankings between 2 and 2.5, with owners and RCx service providers rated second. In general, experts expressed an opinion that working with the right owners was very valuable; however, they noted that building owners and operators vary widely in their ability to support data collection as well as in their understanding of the prevalence of faults.

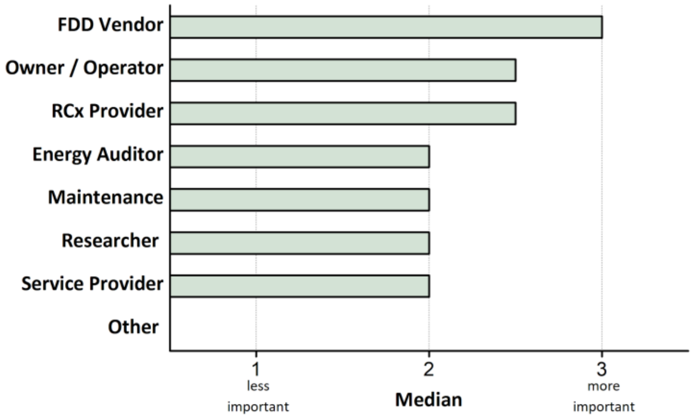


Figure 13. Median importance rankings of access to expanded data sets

Similarly, shown in Figure 14, experts ranked FDD software output as the most important source of data to support enhanced understanding of fault prevalence. All other sources received median values between 2 and 2.5 on a scale of 1 (least value) to 3 (most value). This is interesting, given that FDD software may or may not provide accurate outputs, whereas data sources such as audits, commissioning findings, and direct data surveys include professional inspection and a degree of validation/verification.

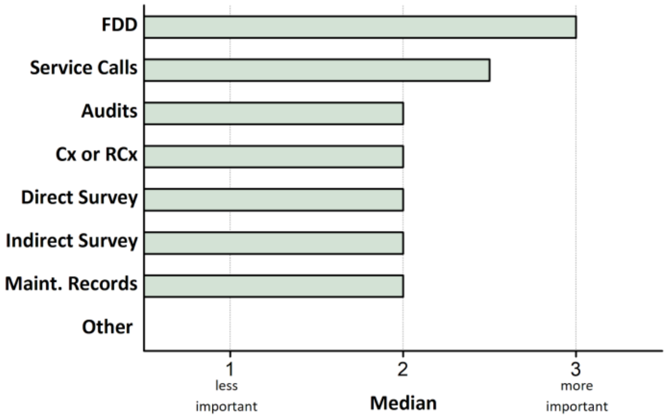


Figure 14. Median importance rankings of fault prevalence data sources

4. Discussion

The findings from the literature and interviews suggest that more work is needed to provide researchers and industry with a comprehensive understanding of the HVAC fault prevalence that exists in commercial buildings. Additionally, the current body of work consists of studies that are impossible to aggregate for high-level evidence-based statistics on fault prevalence. The literature review surfaced two primary gaps across the body of prior work:

1. **Outdated Studies:** The current body of work includes studies from as far back as 1989, and the majority of studies were done before 2010. As time passes, building technologies continue to advance, and original building equipment ages or could be replaced. The studies are out of date.
2. **Aggregation/Scalability:** It is impossible for future studies to build upon these previous studies because they lack consistent metrics, data might have been skewed, and only cover narrow portions of various sample spaces. This precludes aggregation across studies and scaling to estimate fault prevalence across other various regions, systems, equipment ages, or building types.

The experts interviewed described instances of faulty equipment leading to wasted energy. Although they have direct experience with faults in their work, understanding of prevalence tends to be unique to each individual or company. A comprehensive study using consistent metrics, methods, and classification of systems would provide valuable insights to FDD developers, technology providers, and end users of FDD tools. FDD providers would be able to refine the value proposition for technology adoption by better quantifying fault impacts, and developers could focus their efforts on the most common and impactful faults. Building operators who are end users of FDD tools could establish the business case for integrating fault identification and prioritization of corrective action into ongoing operations and maintenance processes. Finally, researchers could use the outcomes of a comprehensive fault prevalence study to identify future research needs and improve diagnostic algorithms and metrics. While these insights are indicative of the perspectives of the 25 experts interviewed, we do not claim that they are statistically representative of the views of all FDD stakeholders - that level of conclusiveness would require a much larger sample size and additional statistical analyses.

A comprehensive fault prevalence study is an ambitious undertaking; however, as suggested in the expert interview findings, emerging data from building-automation-system-integrated automated FDD software products allow investigation at a scale and breadth orders of magnitude larger than what has been possible to date. This opportunity mirrors the trends seen in other fields, such as digital epidemiology, where the explosion of digital information (both direct physiological information as well as associated communications and queries) can be used to understand health and disease at a scale, speed, and breadth that was not possible with the direct generation of data by health professionals (Salathé et al. 2012). A simple example of the scale of these new data sources is Fitbit's analysis of 149 billion hours of heart rate data from 10 million individuals showing relationships between resting heart rate (a proxy for health) and body fat, exercise, and sleep habits (Hodgkins 2018). Although these types of data sources hold promise, care is needed to understand biases and limitations in large complex heterogeneous data sets of mixed quality that have not been directly generated by investigators. Perhaps the most famous example showing the challenges and pitfalls of large-scale data analysis is the failings of the Google Flu Trends initiative where insufficient integration of traditional data sources and understanding of dependencies on intermediary data processing led to inaccuracies (Lazer et al. 2014).

Specific to handling the heterogeneity of field data capturing detected HVAC faults (in contrast to the building automation system data itself) is the domain-specific need for a taxonomy that provides a means to navigate disparate fault naming or labeling conventions, aggregation across different hierarchical levels, and relational mappings between condition-based fault conventions and behavior-based fault conventions. Condition-based faults define the presence of an improper or undesired physical *condition* in a system or piece of equipment (e.g., stuck valves, fouled coils). Behavior-based faults define improper or undesired *behavior* during the operation of a system or piece of equipment (e.g., simultaneous heating and cooling and short cycling). Typically, the faulty behavior is caused by some underlying faulty condition. Discussed more extensively in Frank et al. (2019), both conventions are used in commercial FDD offerings, and these must be resolved and unified for accurate prevalence counts.

Managing data quality and uncertainty could conceivably be achieved by complementing FDD software outputs with smaller samples of data from direct site-level inspections and maintenance, audit, or commissioning records. Although difficult to obtain at scale, these complementary data sources could provide a means of validating the outputs from the FDD software tools. Finally, any large-scale approach would benefit from the incorporation of strategies to understand the impact of selection bias on the results and the degree of representativeness of the findings.

5. Conclusions

Operational faults in commercial building HVAC systems cause significant energy waste and negatively impact occupant comfort. Various industry stakeholders have an interest in better understanding the nature of these HVAC operational faults, including building owners, FDD software developers, retro-commissioning providers, and researchers. Industry stakeholders were interviewed to determine the HVAC fault data that would be of greatest value, and a comprehensive literature review was conducted to determine the extent to which those high-value data are currently available. Through the interviews and literature review, we discovered unmet needs for empirical data on the prevalence of HVAC faults at the desired level of granularity, consistency, and scale. Resolving HVAC fault data for variables such as building type, system type, and climate zone would be valuable for quantifying fault impact and for developing strategies to avoid or at least detect and rapidly address faults.

In order to address the data gaps identified in this paper, we recommend a comprehensive study on commercial buildings' HVAC fault prevalence. Based on the outcomes of the expert interviews and literature review, we recommend that such a fault prevalence study be based on a robust fault taxonomy and a variety of meaningful fault prevalence metrics. To the extent possible, the fault prevalence study should target the highest-value HVAC system types to maximize the statistical significance of the resulting data. The study should also select data sources that balance the need for data accuracy and volume/spread and validate that strategy through a pilot study.

The field of commercial building HVAC system data science and analytics has made significant progress in recent decades, even with the unmet data needs identified in this paper. Despite that progress, the majority of the U.S. commercial building stock still falls far short of meeting the aggressive carbon reduction goals being set by corporations, public-sector building owners, and states. Addressing HVAC data gaps through a comprehensive fault prevalence study that follows the recommendations in this paper would lay the foundation for the next step of improvement in commercial HVAC building science and analytics.

Acknowledgments

The authors thank Marina Sofos and Erika Gupta of the U.S. Department of Energy Building Technologies Office for their support of this work and all the subject matter experts who participated in the interviews for this study. In addition, authors also appreciated the support from Piljae Im, Mini Malhotra, and Jaewan Joe at ORNL, Guanjing Lin at LBNL, and Bennett Doherty at NREL for their participation in the early phase of the project.

This paper was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Building Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Lawrence Berkeley National Laboratory’s contribution to 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.

Pacific Northwest National Laboratory’s contribution to this work was supported under Contract No. DE-AC05-76RL01830.

References

ADM Associates, Inc. 2009. *Market Assessment and Field M&V Study for Comprehensive Packaged A/C Systems Program*. No. SCE0286.01. Southern California Edison Company. http://www.calmac.org/publications/CPACS_Assessment_Final_Report_7-24-09.pdf.

Ardehali, M.M., T.F. Smith, J.M. House, and C.J. Klaassen. 2003. "Building Energy Use and Control Problems: An Assessment of Case Studies." *ASHRAE Transactions*, *109*:111–121.

Behfar, A., D. Yuill, and Y. Yu. 2018. "Supermarket system characteristics and operating faults (RP-1615)." Science and Technology for the Built Environment, 24(10):1104-1113.

Braun, J.E. 2003. "Automated Fault Detection and Diagnostics for Vapor Compression Cooling Equipment." *Journal of Solar Energy Engineering*, *125*(3):266–274.

Breuker, M., T. Rossi, and J. Braun. 2000. "Smart Maintenance for Rooftop Units." *ASHRAE Journal*.

Breuker, M. S., and J. E. Braun. 1998. "Common Faults and Their Impacts for Rooftop Air Conditioners." *HVAC&R Research*, *4*(3):303–318.

Codes and Standards Enhancement (CASE). 2011. *Draft Measure Information Template – Light Commercial Unitary HVAC* (p. 186). California Statewide Utility Codes and Standards Program.

Comstock, M., and J.E. Braun. 1999. *Literature Review for Application of Fault Detection and Diagnostic Methods to Vapor Compression Cooling Equipment*. No. #4036-2. p. 86. ASHRAE Deliverable for Research Project 1043-RP Fault Detection and Diagnostic (FDD) Requirements and Evaluation Tools for Chillers.

Comstock, M., J. E. Braun, and E. A. Groll. 2002. "A survey of common faults in chillers." *ASHRAE Transactions*. 108(1):819.

Cowan, A. 2004. *Review of Recent Commercial Roof Top Unit Field Studies in the Pacific Northwest and California*. New Buildings Institute. Prepared for Northwest Power and Conservation Council and Regional Technical Forum. https://newbuildings.org/sites/default/files/NWPCC_SmallHVAC_Report_R3_.pdf.

Close, B. 2010. *HVAC energy efficiency maintenance study* (CALMAC Study ID SCE0293.01), Southern California Edison. http://www.calmac.org/publications/HVAC_EE_Maintenance_Final.pdf

Davis, R., D. Baylon, R. Hart, and E. Water. 2002. "Identifying Energy Savings Potential on Rooftop Commercial Units." In *Proc. ACEEE Summer Study on Energy Efficiency in Buildings*. https://aceee.org/files/proceedings/2002/data/papers/SS02_Panel3_Paper07.pdf.

Davis, R., P. Francisco, M. Kennedy, D. Baylon, and B. Manclark. 2002. *Enhanced Operations & Maintenance Procedures for Small Packaged Rooftop HVAC Systems*. http://www.ecotope.com/wp/wp-content/uploads/2014/07/2002_002_EnhancedOperationsMaintenance.pdf

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

Djunaedy, E., K. van den Wymelenberg, B. Acker, and H. Thimmana. 2011. "Oversizing of

HVAC system: Signatures and penalties." *Energy and Buildings*, *43*(2):468–475. https://www.sciencedirect.com/science/article/pii/S0378778810003634.

Downey, T., and J. Proctor. 2002. "What Can 13,000 Air Conditioners Tell Us?" *ACEEE Summer Study Proceedings*, *1*:16. https://aceee.org/files/proceedings/2002/data/papers/SS02_Panel1_Paper05.pdf.

Emmerich, S. J., T. McDowell, and W. Anis. 2005. *Investigation of the impact of commercial building envelope airtightness on HVAC energy use.* No. NIST IR 7238. Gaithersburg, MD: National Institute of Standards and Technology. https://www.nist.gov/publications/investigation-impact-commercial-building-envelope-airtightness-hvac-energy-use-0.

Energy Information Administration. 2002. *Commercial Buildings Energy Consumption Survey*.

Energy Information Administration. 2017. "Energy Consumption Estimates by End-Use Sector, Ranked by State." https://www.eia.gov/state/seds/sep_sum/html/rank_use.html.

Energy Market Innovations, Inc. 2004. *Small Commercial HVAC Pilot Program.* No. #E04-135. Northwest Energy Efficiency Alliance.

Farahmand, F., C. Chappell, and H. Weitze. 2017. *Economizer Fault Detection and Diagnostics (FDD) for Built-Up Air Handlers – Final Report.* No. 2019- NR-MECH2- F. Codes and Standards Enhancement (CASE) Initiative 2019 California Building Energy Efficiency Standards.

Felts, D.R., and P. Bailey. 2000. "The State of Affairs—Packaged Cooling Equipment in California." In *Proceedings from ACEEE Summer Studies on Energy Efficiency in Buildings*.

Frank, S., G. Lin, X. Jin, R. Singla, A. Farthing, and J. Granderson. 2019. "A performance evaluation framework for building fault detection and diagnosis algorithms." *Energy and Buildings*, *192*:84–92. https://doi.org/10.1016/j.enbuild.2019.03.024.

Gage, C., and G. Troy. 1998. "Reducing Refrigerant Emissions from Supermarket Systems." *ASHRAE Journal*.

Goetzler, W., R. Shandross, J. Young, O. Petritchenko, D. Ringo, and S. McClive. 2017. *Energy Savings Potential and RD&D Opportunities for Commercial Building HVAC Systems* No. DOE/EE-1703. Navigant Consulting, Burlington, MA. https://www.energy.gov/sites/prod/files/2017/12/f46/bto-DOE-Comm-HVAC-Report-12-21-17.pdf.

Granderson, J., R. Singla, E. Mayhorn, P. Ehrlich, D. Vrabie, and S. Frank. 2017. *Characterization and Survey of Automated Fault Detection and Diagnostics Tools.* No. LBNL-2001075. Lawrence Berkeley National Laboratory. https://eta-publications.lbl.gov/sites/default/files/lbnl-2001075.pdf.

Gunay, H. B., W. Shen, and C. Yang. 2018. "Text-mining building maintenance work orders for component fault frequency." *Building Research & Information*, *47*(5):518–533. https://www.tandfonline.com/doi/abs/10.1080/09613218.2018.1459004

Heinemeier, K. 2012. *Rooftop HVAC Fault Detection and Diagnostics: Technology and Market Review / Energy and Demand Savings Estimates.* No. CEC PIER #500-08-049.

Hewett, M., D. Bohac, R. Landry, T. Dunsworth, S. Englander, and G. Peterson. 1992. "Measured Energy and Demand Impacts of Efficiency Tune-ups for Small Commercial

Cooling Systems." In *ACEEE Conference Proceedings*. https://aceee.org/files/proceedings/1992/data/papers/SS92_Panel3_Paper14.pdf.

Hodgkins, K. 2018, August 31. "Fitbit Data Mining Identifies Some Surprising Trends in Resting Heart Rate."

Hunt, M., K. Heinemeier, M. Hoeschele, and E. Weitzel. 2010. *HVAC Energy Efficiency Maintenance Study*. Davis Energy Group; UC Davis; WCEC.

Hyvärinen, J., and S. Karki (Eds.). 1996. *Building optimization and fault diagnosis source book*. Espoo: VTT Building Technology.

ICC, 2012. International Energy Conversion Code (IECC), International Code Council (ICC) Inc. Falls Church, VA

Jacobs, P. 2002. "Anticipating Economizer Failure During Small-Packaged System Installation." *HPAC Engineering*, *74*(9):4.

Jacobs, P. 2003. *Small HVAC Problems and Potential Savings Reports* (No. 500- 03- 082- A-25). California Energy Commission.

Katipamula, S. 2017. *Improving Commercial Building Operations thru Building Re-tuning: Meta-Analysis*. Pacific Northwest National Laboratory. https://buildingretuning.pnnl.gov/documents/pnnl_sa_110686.pdf.

Katipamula, S., and M. R. Brambley. 2005a. "Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I." *HVAC&R Research*, *11*(1):3–25.

Katipamula, S., and M. R. Brambley. 2005b. "Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part II." *HVAC&R Research*, *11*(2):169–187.

Kim, W., and S. Katipamula. 2018. "A review of fault detection and diagnostics methods for building systems." *Science and Technology for the Built Environment*, *24*(1):3–21. https://www.tandfonline.com/doi/full/10.1080/23744731.2017.1318008.

Kramer, H., G. Lin, C. Curtin, E. Crowe, and J. Granderson. 2019. Building analytics and monitoring-based commissioning: industry practice, costs, and savings. *Energy Efficiency*.

Lazer, D., R. Kennedy, G. King, and A. Vespignani. 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science*, *343*(6176):1203–1205.

Li, Y., and Z. O'Neill. 2018. "A critical review of fault modeling of HVAC systems in buildings." Building Simulation, 11:953-975.

Lin, G., H. Kramer, and J. Granderson. 2020. "Building fault detection and diagnostics: Achieved savings, and methods to evaluate algorithm performance." *Building and Environment*, *168*:106505. https://www.sciencedirect.com/science/article/pii/S0360132319307176.

Liu, M., A. Athar, Y. Zhu, and D.E. Claridge. 1995. *Reduce Building Energy Consumption by Improving the Supply Air Temperature Schedule and Recommissioning the Terminal Boxes*. No. ESL-TR-95/01-04. Energy Management and Operations Division at the M D. Anderson Cancer Center.

Madani, H. 2014. "The Common and Costly Faults in Heat Pump Systems." *Energy Procedia*, 4.

Mills, E., 2011. "Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the United States." Energy Efficiency, 4(2), pp.145-173.

Mowris, R. 2006. *Evaluation Measurement And Verification Of Air Conditioner Quality Maintenance Measures.* No. CPUC Reference Numbers 1385-04, 1395-04 and 1437-04. Aloha Systems, Inc.

Mowris, R., A. Blankenship, E. Jones, and R. Mowris. 2004. "Field Measurements of Air Conditioners with and without TXVs." *ACEEE Summer Study Proceedings, 1*:16.

Neida, B. V., D. Manicria, and A. Tweed. 2001. "An Analysis of the Energy and Cost Savings Potential of Occupancy Sensors for Commercial Lighting Systems." *Journal of the Illuminating Engineering Society, 30*(2):111–125.

Qin, J., and S. Wang. 2005. "A fault detection and diagnosis strategy of VAV air-conditioning systems for improved energy and control performances." *Energy and Buildings, 37*(10):1035–1048.

Roth, K. W., D. Westphalen, P. Llana, and M. Feng. 2004. "The Energy Impact of Faults in U.S. Commercial Buildings." p. 9. Presented at the International Refrigeration and Air Conditioning Conference.

Roth, K. W., D. Westphalen, M. Y. Feng, P. Llana, and L. Quartararo. 2005. *Energy Impact of Commercial Building Controls and Performance Diagnostics: Market Characterization, Energy Impact of Building Faults and Energy Savings Potential.* No. D0180. TIAX.

Salathé, M., L. Bengtsson, T.J. Bodnar, D.D. Brewer, J.S. Brownstein, C. Buckee, E.M. Campbell, C. Cattuto, S. Khandelwal, P.L. Mabry, and A. Vespignani. 2012. "Digital Epidemiology. "*PLOS Computational Biology*, *8*(7):e1002616.

Sharma, A.B., L. Golubchik, and R. Govindan. 2010. "Sensor faults: Detection methods and prevalence in real-world datasets." ACM Transactions on Sensor Networks (TOSN), 6(3):1-39.

Shoukas, G., M. Bianchi, and M. Deru. 2020. Analysis of Fault Data Collected from Automated Fault Detection and Diagnostic Products for Packaged Rooftop Units. No. NREL/TP-5500-77077. National Renewable Energy Lab.(NREL), Golden, CO (United States).

Smart Energy Analytics Campaign. 2019. https://smart-energy-analytics.org/.

Stouppe, D.E., and T.Y.S. Lau. 1989. "Air Conditioning and Refrigeration Equipment Failures." *The National Engineer*, *93*(9):14–17.

Wen, J., and S. Li. 2011. *Tools for Evaluating Fault Detection and Diagnostic Methods for Air-Handling Units*. ASHRAE 1312-RP.

Yuill, D.P. and J.E. Braun. 2013. "Evaluating the performance of fault detection and diagnostics protocols applied to air-cooled unitary air-conditioning equipment." HVAC&R Research, 19(7):882-891.

Zhao, Y., J. Wen, F. Xiao, X. Yang, and S. Wang. 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.