

Development of a Simulated Air Handling Unit Fault Dataset for FDD Tools: Lessons Learned and Considerations for FDD Developers

*Armando Casillas, Guanqing Lin, and Jessica Granderson, Lawrence Berkeley National Laboratory,
Sen Huang and Yan Chen, Pacific Northwest National Laboratory*

ABSTRACT

As energy management and information systems (e.g., automated fault detection and diagnostics [AFDD] tools) become more prevalent in the commercial building stock, it is important to determine the effectiveness of these technologies by benchmarking their performance. The authors have been working to develop the largest publicly available dataset of HVAC fault data for performance benchmarking applications, covering the most common HVAC systems and designs including chiller plants, rooftop packaged units, dual duct air handling units and single duct air handling units. This study covers the development, modeling, and validation of a synthetic fault dataset for a single duct air handling unit (AHU), one of the most common HVAC configurations found in the commercial building stock. Despite this being a common system, real-world time series data are scarce and usually do not span a wide range of weather conditions. Due to this limitation, a detailed AHU model was employed to carry out annual simulations of numerous common sensor and mechanical faults, which were then validated by comparing their effects on system performance to expected symptoms. We summarize the nature of each fault and their impacts under different weather and operation conditions. Finally, we highlight considerations for FDD developers that may want to use this dataset to assess their algorithms' performance and their improvement over time.

Introduction

As building data becomes more readily available, and as the budding field of data science and analytics comes to buildings, fault detection and diagnostics (FDD) is of increasing relevance to the research and product development communities. A primary method of improving building controls and operational efficiency is through algorithms developed to perform FDD, which use building data to identify the presence of faults and potentially isolate root causes. Building owners and operators have already leveraged the benefits of FDD technology, with an estimated median whole-building portfolio savings of 9% (Kramer et al. 2019).

FDD development for the single duct hydronic air handling units (SDAHU) system, for example, are presented in a number of studies from the start of the millenium (House, Vaezi-Nejad, and Whitcomb 2001) (Bushby et al. 2001). Since then, a diversity of techniques have been developed for FDD in AHU systems (Yo, Woradechjumroen, and Yu 2014). In less sophisticated but more common methods, studies take a knowledge-based approach, using rules to determine faults in the HVAC system (Bushby and Schein 2006). Analytical-based physical

model FDD such as Wu and Sun's study (2012), determines accurate predictions in zone temperature based on AHU energy balances to detect when zone temperature deviates from predicted values. More access to building and HVAC system data has spawned data driven approaches to create neural networks as a fault diagnosis method as seen in Liao et. al's study for example (2021).

A persistent challenge, however, has been the lack of common datasets and test methods to benchmark the performance accuracy of FDD methods, and gauge improvement of these tools over time. Granderson (2018) most recently developed a test and benchmarking framework for FDD algorithm performance, demonstrating a growing need for HVAC fault datasets that can be used to further determine the accuracy and effectiveness of FDD algorithms. HVAC performance datasets have been developed before in the form of ASHRAE's RP1312 fault dataset. ASHRAE Project RP-1312 data (Li et al. 2010a, Li et al. 2010b) is the resulting dataset from a series of experiments that were performed on two multi-zone VAV AHUs (AHU-A and AHU-B) with the same configuration running simultaneously. AHU-A always ran under normal conditions while AHU-B simulated different fault conditions. Seventeen faults were tested, for example, outside air damper stuck/leakage, cooling coil valve stuck, heating coil valve leakage, AHU duct leaking, control unstable, outside air temperature bias, etc. Each fault was tested at multiple fault intensity levels in three seasons - spring, summer, and winter. At each level of each fault, the experiment lasted for one day and the operational data of 160 variables were collected at 1-min intervals. This dataset has been leveraged by a number of FDD studies (Yan et al 2016) (Zhong et al 2019) (Montazeri and Kargar 2020) (Yun, Hong, and Seo 2021). Further work was initiated to fill this gap with the introduction of an open sourced dataset for FDD evaluation purposes (Granderson et al. 2020), which introduced a first of its kind public dataset with ground-truth data on the presence and absence of faults for multiple HVAC systems, including a simulated SDAHU system.

This paper will more specifically dive into the expansion of a SDAHU fault dataset, which is considered one of the most typical HVAC system designs in commercial buildings. The data set consists of high resolution, simulated time series HVAC operational data (e.g. temperatures, pressures, control signals, component status, etc.) under a diversity of operating and weather conditions, combined with information on the presence and absence of faults and their associated intensity. Furthermore, the paper applies our previously established data validation and ground truth assessment protocol for the successful development of the SDAHU FDD test dataset (Casillas et al. 2020).

Methods

The overview of the SDAHU model, including the system configuration, controls specification and co simulation framework will be detailed in this section. Furthermore we will provide detail on the modeled HVAC faults, the nature of these faults and expected system performance and behavior as well as our method of imposition.

Model Overview:

The SDAHU model was developed in the Modelica language by developers at PNNL, based on model components available in open-source Modelica libraries such as the Modelica Buildings, IBPSA libraries. Modelica is an equation-based, object-oriented modeling language for complex dynamic systems. In order to capture the building's thermal response a reference commercial building model from EnergyPlus (Deru et al. 2011) was integrated. The data exchange between the EnergyPlus input data file (IDF) model and the Modelica system model, as pictured in Figure 1, was handled by a co-simulation framework, exporting the IDF file as a functional mockup unit, analogous to the methods in Huang et. al (2021). In addition to calculating the thermal loads of the space, the IDF file also stores pertinent weather information that is fed into the modelica model, which allows for annual modeling of a building based on a historical weather data set. For this study's purposes, the climate data modeled was that of Chicago, Il. The final result of this co-simulation process is a .mat result file that contains time-series building performance data.

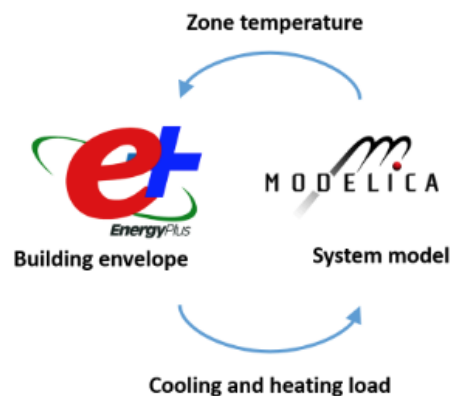


Figure 1. The data exchanged by E+ and Modelica are shown in the figure above

The major components of the modeled SDAHU, as shown in Figure 2, are supply air fan with a variable frequency drive (VFD), return fan with a VFD, cooling coil, cooling control valves, outdoor air (OA) and return air (RA) dampers. The control specifications of the AHU are shown in Table 1. The AHU's baseline control sequence is applied from engineering standard best practices (e.g. ASHRAE 90.1) and are detailed below in Table1. These control parameters and sequences are programmed in the modelica language with control and logic components.

The control loops are mostly concerned with three different components:

- Fan speed control determined by occupancy state and static pressure setpoints
- Cooling coil valve position determined by occupancy state and supply air temperature setpoint
- Damper positions determined by occupancy state, outdoor air temperature and mixed air temperature setpoint

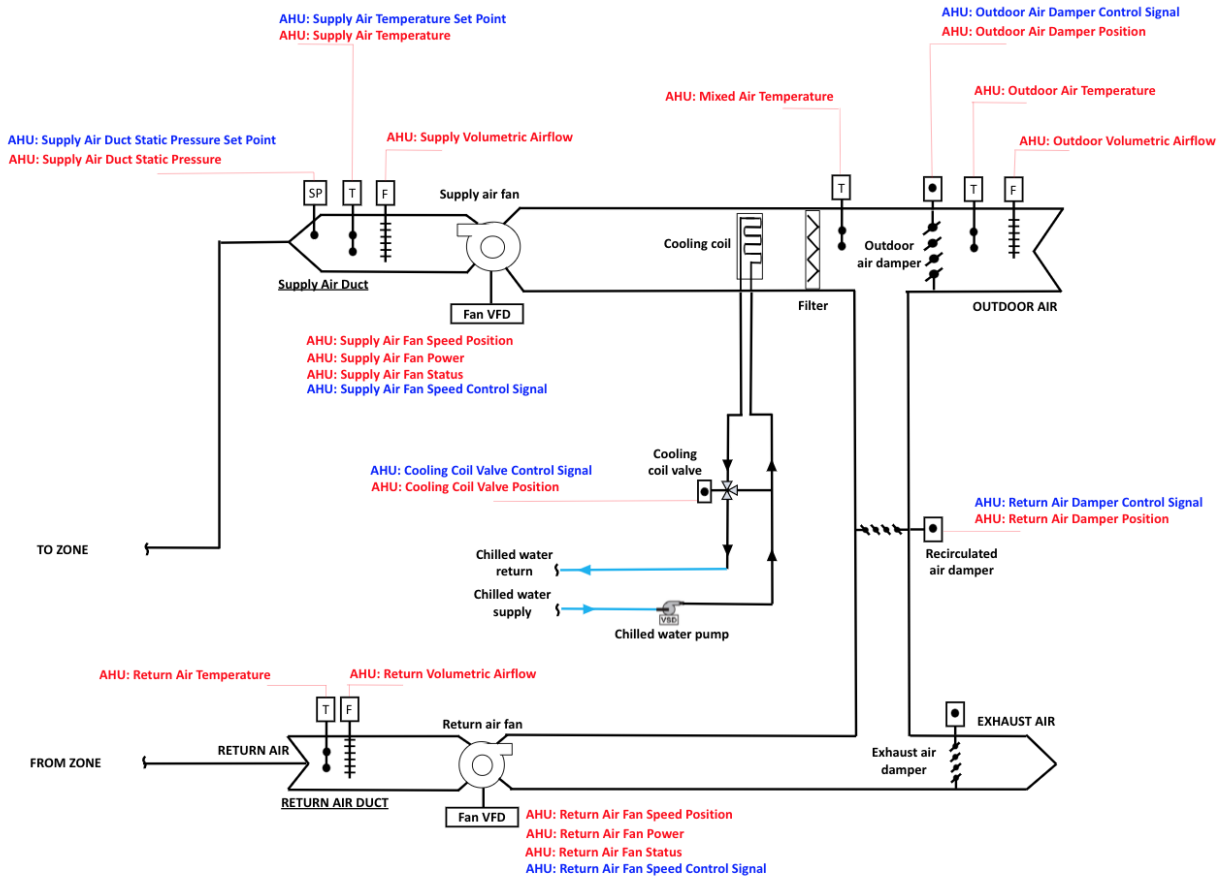


Figure 2. SDAHU diagram with all measurement points denoted

Table 1. Controls overview for SDAHU Model

Control/Operations Specification		Description	Specification
		Typical Building Control Baseline	Data Source/ References
System Operation Mode	Occupied Mode	Start the HVAC system 2 hr ahead of occupancy schedule: • Occupancy schedule (weekday 6:00-22:00) • Cooling set point (occupied): 26.7°C (75°F) • Heating set point (occupied): 21°C (70°F)	DOE commercial reference building (Deru, et al., 2011)
	Unoccupied Mode	Maintain the unoccupied heating and cooling setpoint: • Cooling set point (unoccupied): 24°C (80°F) • Heating set point (unoccupied): 15.6°C (60°F)	
Air Handling Unit	Supply /Return fan control	Fixed static pressure, SPset: 169.8 Pa (0.68 in. w.g.)	Based on testing, air balancing analysis of given system to meet cooling design condition
		Fixed differential speed ratio (10% less) between supply air and return air fan.	Based on engineering practices
	Supply air temperature control	Fixed supply air temperature setpoint: 12.7°C (55°F)	Based on engineering practices
	Minimum outdoor air control	Fixed minimum OA damper position (10% open) during the occupied hour. Closed during the unoccupied hour.	Based on engineering practices
	Economizer	Fixed dry bulb temp threshold, OA damper is engaged from 1°C to 15.6°C (33°F to 60°F), otherwise at minimum position (10%). Damper modulates to hold mixed air temperature of 55F.	ASHRAE Guideline 36-2020 / ASHRAE 90.1-2016

Fault Modeling:

Table 2. Overview of HVAC fault modeled and imposition method

Fault	Method of Fault Imposition	Fault intensities covered
Supply, Outdoor Air Temperature Sensor Bias	Add or subtract value from initial sensor reading	-4,-2,+2,+4 °C
OA Damper, Cooling Coil Valve Stuck	Automated override of OA damper position to indicate that OA damper is stuck. Automates to override of coil valve position to indicate that cooling coil valve is stuck.	10%, 25%,50%,75%, Fully open (100%)
Cooling Coil Valve Leak	Adjusted the minimum coil valve position value when control signal is zero	10%, 25%,40%,50%

3 different components were targeted for fault modeling in the SDAHU model: the outdoor air damper, the cooling coil and the temperature sensors.

The **outdoor air damper stuck** fault is a mechanical fault by nature and will directly affect the AHU's ability to take advantage of outdoor air to maintain supply temperatures while minimizing cooling energy as well its ability to maintain effective supply temperature control. During instances in which the OA damper is stuck above minimum position and supply air is cooler than desired setpoint, excess outdoor air may cause the cooling energy to be minimized while dramatically reducing the supply air temperature of the AHU. The case in which warmer temperatures are seen outdoors, the excess outdoor air will cause more cooling energy to be used, driving the control signal of the OA damper to minimum while maximizing the cooling coil control signal. Higher than normal supply air temperatures may occur.

A **stuck cooling coil valve** directly affects the AHU's ability to maintain effective supply air temperature control. During instances in which the cooling coil valve is stuck closed or at a position that is lower than needed, the supply air will be warmer than desired, driving the control signal to 100% due to the inability of the system to maintain cool enough air to the zone level. This will cause higher than normal supply temperatures, higher than normal return air temperatures, and lower overall cooling energy consumed, with higher energy consumed in the fan, as the zone demand increases with less cooling available.. During instances in which the valve is stuck open or higher than needed, the cooling coil will be providing too much cooling. This will result in a supply temperature colder than the setpoint and the control signal will eventually be driven to zero due to the inability of the system to maintain supply air temperature

set point. This will ultimately lead to lower than desired supply and return air temperatures and higher overall cooling energy consumed.

A **leaking cooling coil valve** affects the AHU's ability to fully close the cooling coil valve. During instances in which the control signal is driven to a level below the leakage level or to 0, the ground truth position of the valve will bottom out at the leakage level. This will cause lower than normal supply temperatures during these instances, and higher overall cooling energy consumed. During instances in which the leakage level is higher than the control signal, the fault will behave more like a stuck valve fault.

A **temperature sensor bias fault in the outdoor temperature sensor** would cause an adverse effect on supply temperature control, mainly the modulation of the outdoor air damper according to the economizer control sequence. As the bias becomes more positive (4°C), the seemingly higher outdoor air temperature would result in less activity in the economizer control signal, resulting in higher overall cooling energy consumption. As the bias becomes more negative (-4°C), the seemingly lower outdoor air temperature would result in a more active control signal for the economizer, resulting in lower overall cooling energy consumption by the cooling coils.

A **temperature sensor bias fault in the supply temperature sensor** would cause an adverse effect on supply temperature control, mainly the modulation of the cooling coil valve to meet setpoint. As the bias becomes more positive (4°C), the seemingly higher supply temperature would result in higher control signal for added cooling, resulting in higher overall cooling energy consumption, cooler rooms (lower return air temperatures) and possible impact on thermal comfort. As the bias becomes more negative (-4°C), the seemingly lower supply temperature would result in lower control signal for reduced cooling, resulting in lower overall cooling energy consumption, cooler rooms (lower return air temperatures).

The faults are all implemented by modifying or overriding the baseline control logic of the model. For example, the outdoor air damper stuck fault is implemented by overriding the position of the damper component. The fault imposition methods are summarized in the table below. As an example, for each intensity of the OA damper stuck fault, the fault is imposed by overriding the position of the modeled damper to the predetermined value. The scaled dataset creation is carried out with a parametric simulation Modelica script. This allows for the intensity of each fault to be modeled based on a single value that is passed as a parameter into the fault model component such as "TwoWayValveStuck" for both the cooling coil valve and OA damper.

Results

The SDAHU model was simulated under baseline, fault free conditions for one full calendar year and further simulated with each of the faults under different severity levels. The ability to conduct annual fault simulation is one of the most valuable contributions, since this allows us to observe the fault's impact on system behavior and performance across the full range of weather conditions. The difference in behavior across seasons will be covered in this section. First, are

details on the observed behavior of the SDAHU model under fault free conditions for two different seasons (Spring and Summer). In the subsequent section, the behavior under a sample fault case will then be analyzed in comparison to the previous baseline case. The outdoor air damper and cooling coil control are the focus of the analysis below. The fan speed will remain mostly constant across the annual dataset so it is not highlighted.

Baseline Operation:

Spring:

The first season analyzed is Spring in which we expect to see milder outdoor air temperatures. This equates to maximum activity for the economizer and minimum cooling coil use. The OA damper control sequence can be seen in Figure 3 as being activated in the range of 3°C to 12.7°C (37.5°F to 55°F). The supply temperature setpoint is set at 55°F, so the damper modulating to 100% is expected. Because the simulated range of this sample day never reaches the min or max thresholds for economizer mode (33°F, 60°F), we never see the minimum position of the damper OF 10%. The cooling coil is modulated based on the supply air temperature setpoint and works in conjunction with the OA damper in milder conditions. In Figure 4 we can see the OA damper is modulated to meet SAT setpoint until the OA temp reaches its max threshold of 60°F. The OA damper then is commanded to a minimum value of 10% while the cooling coil command signal is ramped up to meet the supply air temperature setpoint at this instance. As seen by the supply air temperature plot, the setpoint of 55°F is always met.

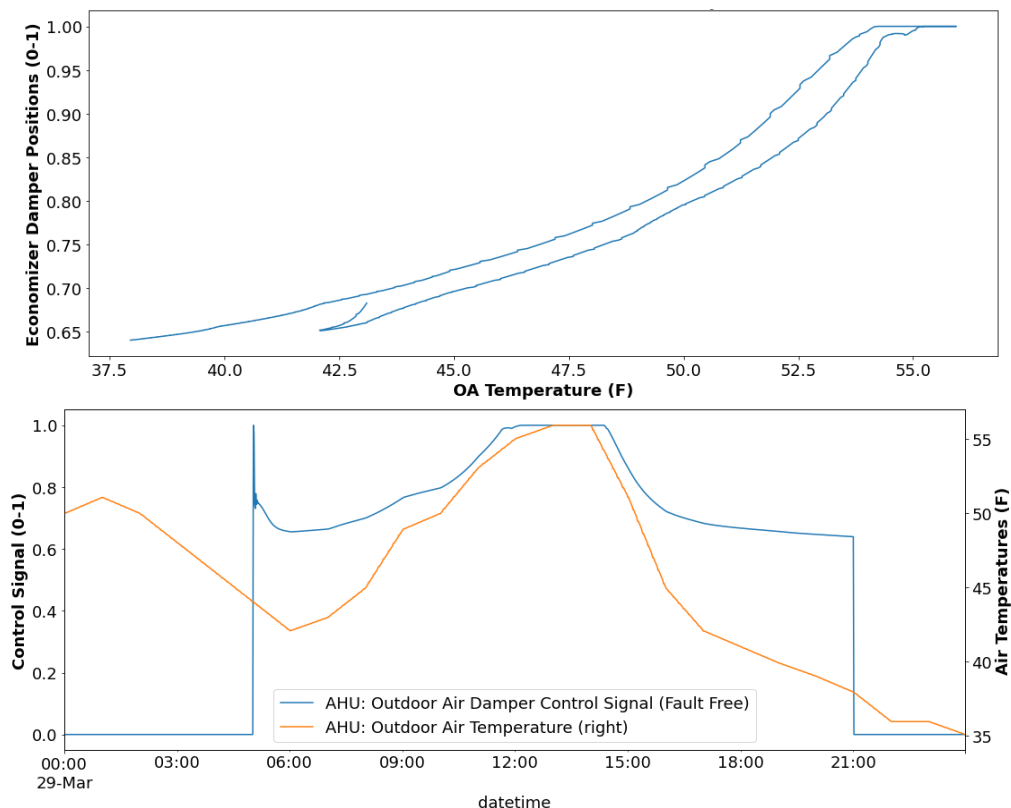


Figure 3. Validation of Economizer control sequence

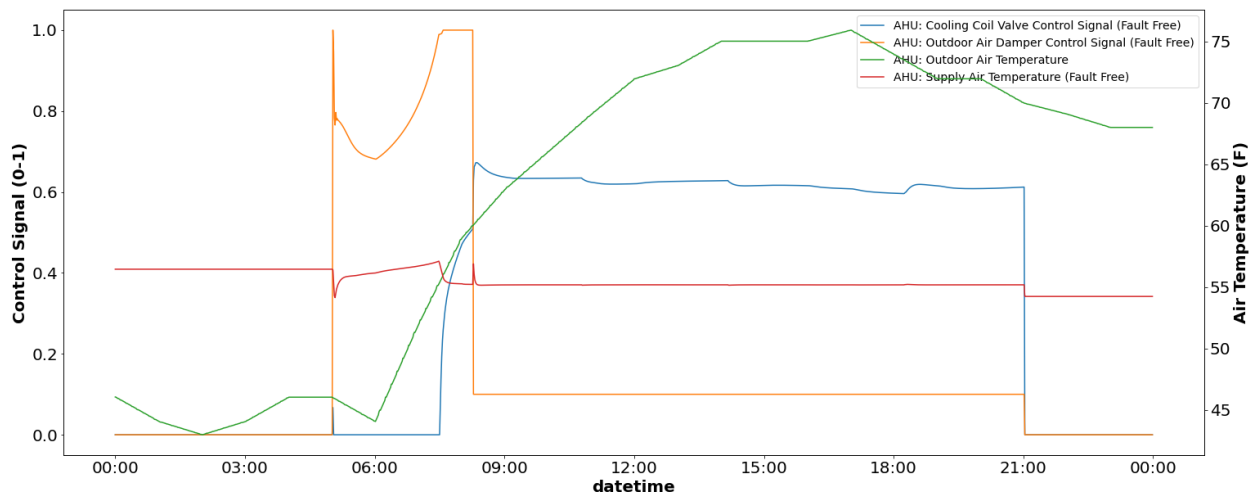


Figure 4. Spring Operation of OA damper and cooling coil working together to maintain 55F supply temperature

Summer:

Analyzing the SDAHU model's behavior during the Summer period is relatively more straightforward compared to the nuanced operation of the OA damper during shoulder season, when economizing is frequent. As seen in Figure 5, during summer operation, the OA temperatures range at values greater than the 60°F max threshold for economizing, so the damper is always set at 10%, while the cooling coil is modulated instead to meet the supply air temperature setpoint.

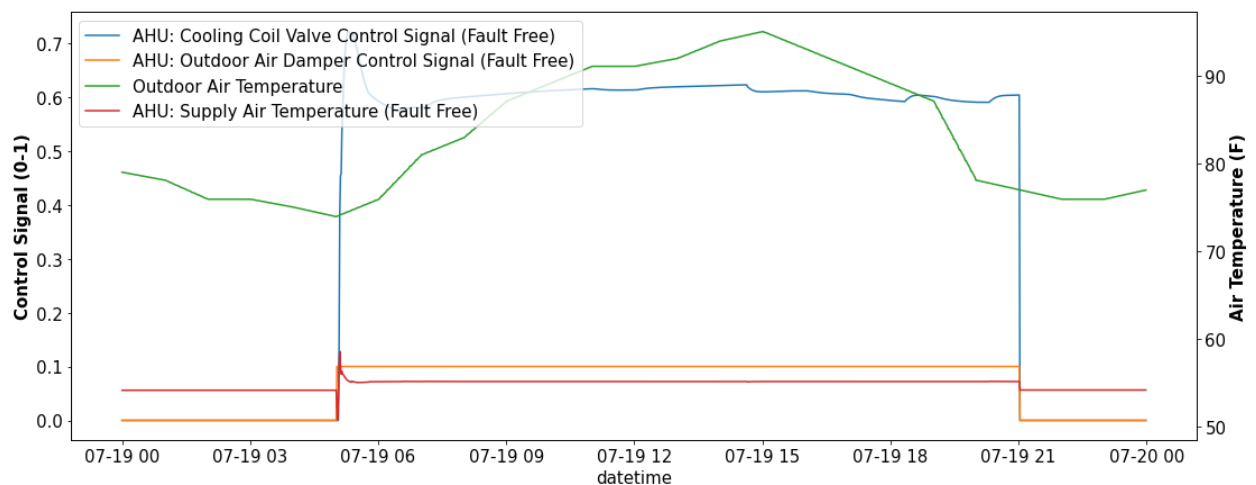


Figure 5. Summer Operation of OA damper at min position while cooling coil maintains 55F supply temperature

Faulty Operation - OA Damper Stuck at Minimum (10%):

Spring:

For the faulty operation example, the outdoor air damper stuck at minimum case is presented. As seen in Figure 6. The same day is plotted as in Figure 3, although now we see the faulty operation of the damper during this fault. The first thing to observe is that the control signal is at 100% throughout the entirety of the day. This is caused primarily by the feedback loop of the controller, which is calculating the difference between the mixed air temperature and the supply air setpoint of 55°F. As the temperature difference increases due to the stuck component, the control output is saturated at 100%. Meanwhile, the outdoor air damper ground-truth position is plotted at a constant value of 10%, which allows us to effectively validate the presence of our fault. This fault results in higher cooling coil activity and higher energy consumption due to the lost opportunity of economizing based on ideal weather conditions. This can be seen in the subsequent plot in Figure 6, in which the cooling coil signal is noticeable higher for the faulty case.

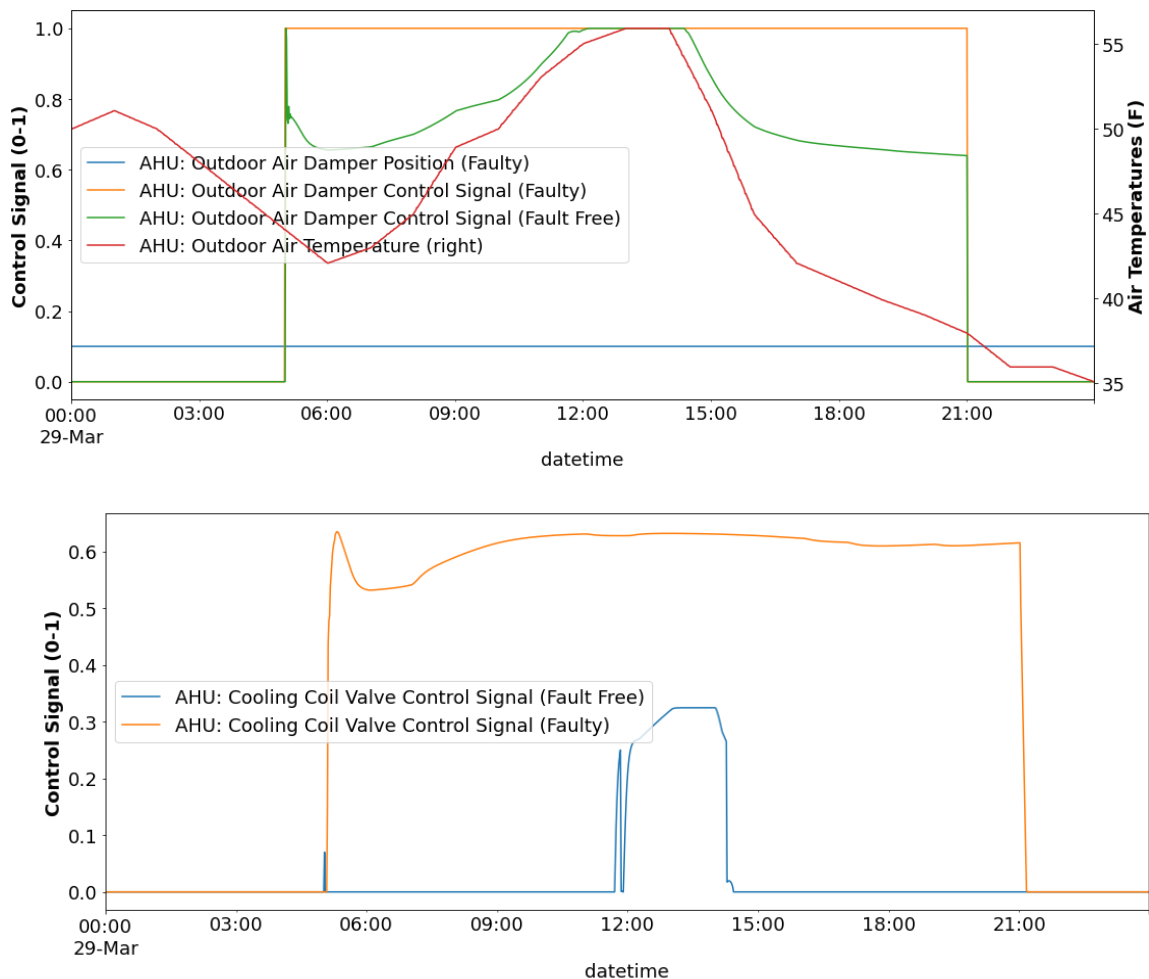


Figure 6. Spring Operation of a stuck OA damper. Cooling coil is more active in faulted case in order to maintain 55F supply temperature

Summer:

The presence and symptoms of each fault will not always be evident, based on the weather conditions and/or the operational state of the HVAC system. This is most evident in the Summer case for the OA Damper Stuck at Minimum case, shown in Figure 7, where the OA temperatures reach their maximum range, up to 95°F. This is well beyond the maximum threshold for economizing and as a result the damper is already at minimum position. The lack of OA damper modulation means the faulty and baseline cases are virtually indistinguishable from one another, as seen by the pair of Figures below, where the OA damper control signal overlaps at 10%, and the cooling coil control signal for both cases are also equivalent.

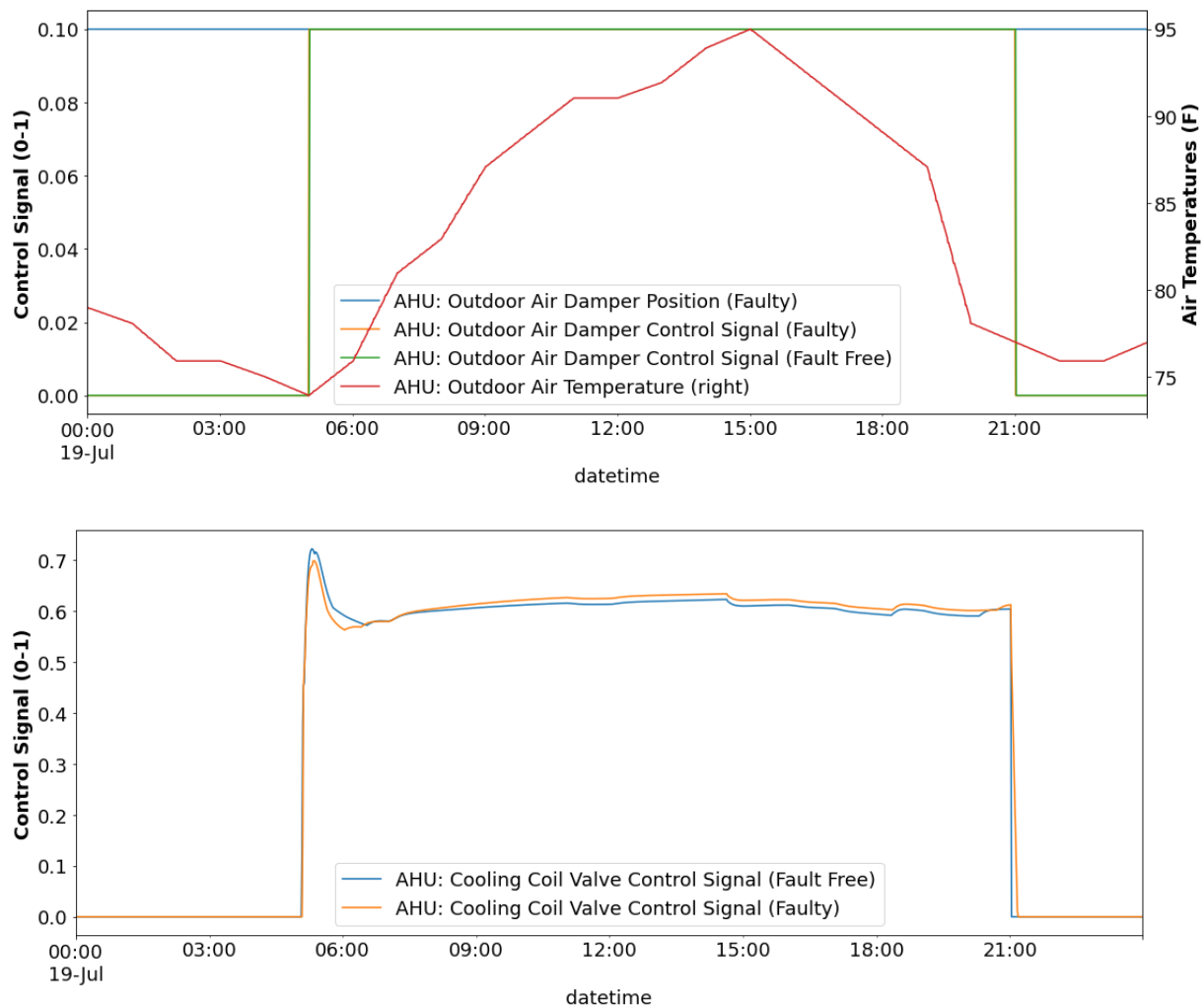


Figure 7. Summer Operation of a stuck OA damper, the ambient conditions during summer cause the damper to stay at minimum and therefore the symptoms are not prevalent in this season.

Discussion

The results presented in this study are only a small subset of a significant, large-scale effort. The process was a multi-year effort which presented challenges and subsequent lessons learned that we have been able to draft into a series of best practices. The aforementioned study detailing our data validation protocol (Casillas et al. 2020) lists the process of building the datasets by first conducting small scale simulations and validation before proceeding to full scale simulation. Individual control sequences are best tested in functional tests that can be executed in Modelica, but may lead to unintended behaviors as part of a full scale model with varying weather conditions and operational states. This is particularly effective in validating documented control sequences across different seasons. Dedicating some time to validating small scale simulation results can avoid wasted time and resources, given that scaled up annual simulations can take several days to complete and occupy anywhere between 1-10 GB of memory.

One such incident during the generation of this dataset was the discovery of competing PID loops from the OA damper and the cooling coil in simultaneous economizing and cooling situations. The issue is compounded when the simultaneous conditioning of these two systems have different control variables, with the economizer sequence, controlled to mixed air temperature and cooling coil valve controlled to the supply air temperature. When applying PI controllers to a process, there will always be static errors between these two values. The static error of the PI controller # 1 may trigger the action of the PI controller # 2, which is downstream. Best practice controls programming for AHU's recommends cooling coil valve control to be disabled in economizing mode and disabled until economizer reaches 100% as seen in Figure 8.

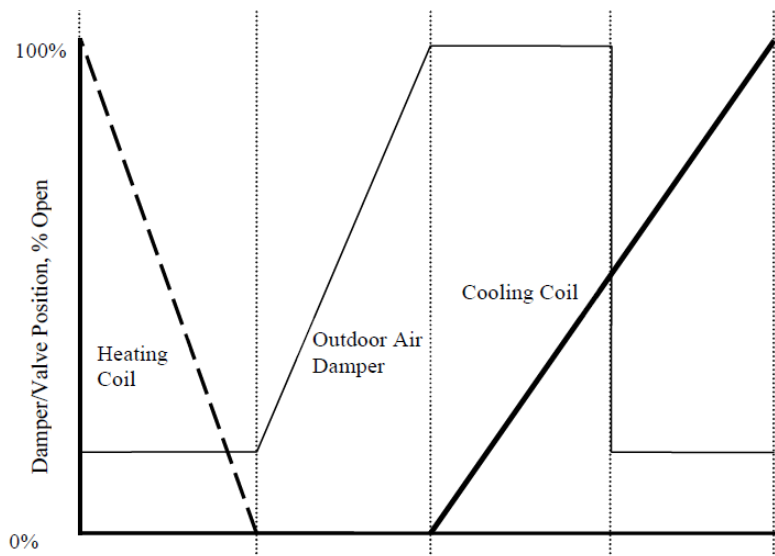


Figure 8. Traditional interaction between OA damper and heating/cooling coils

Another lesson learned in best practice has to do with time expression in the co simulation framework. Daylight savings options exist in EnergyPlus time reporting, which may affect the time schedules of the HVAC operation when modeled in the Modelica environment. To avoid this, it would be best to disable daylight savings options for IDF files.

Along with best practices for large scale simulations, we have compiled considerations for FDD developers that may want to use this dataset to assess their algorithms' performance. The dataset has undergone a validation protocol, which includes common naming convention, topology, and post processed variables. We have tried to reduce the number of measurements available to us via simulation to those that are usually found in a typical building management system. This removes variables like actuator position and wet bulb temperatures from the dataset list.

The common naming convention has been most recently paired with an attempt to align the dataset description and topology to that of the Brick Schema ontology (Balaji et al. 2016). This allows some algorithms to determine the topology of the HVAC system and the relationships more efficiently between parts, and measurement associated with each part. The Brick Schema tool set allows us to generate machine readable files that can be visualized and processed by more sophisticated algorithms with rule-based sets or clustering based on the topology of the system.

This dataset along with others from different data partners and contributors are well on their way to being part of an eventual publicly available dataset in which users will be able to download along with documentation of the system, fault cases and the aforementioned machine-readable files. The hope is for the FDD community to apply to their algorithm development similar to that of ASHRAE RP-1312 and further spur the growth of this valuable technology.

Conclusion

This paper presents results and lessons learned from a multi-year modeling effort aimed at producing a diverse set of annual fault cases for the SDAHU system: one of the most common HVAC system configurations found in the building stock. Annual simulation of any system provides a full range of operating conditions across all seasons, a diversity that may not be available in limited experimental or field measured datasets. The results present baseline and faulted scenarios for Spring and Summer operation of the SDAHU model that demonstrate the value of assessing performance across multiple seasons. Lessons learned about the modeling process, including issues such as control loop programming in Modelica are discussed in order to provide some insight for FDD developers on how difficult it may be to accurately emulate a real system. Exciting new developments include the development of the largest publicly available FDD dataset, which will include the SDAHU dataset along with that of 6 other systems along with BRICK models for all systems.

Acknowledgement

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. We also recognize each of the fault detection and diagnostic tool developers who participated in this survey. We would also like to thank Erika Gupta and the Building Technologies Office as well as our data contributors.

References

- Bharathan Balaji, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, ... Kamin Whitehouse. 2016. Brick: Towards a Unified Metadata Schema For Buildings. In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys '16). Association for Computing Machinery, New York, NY, USA, 41–50. DOI://doi.org/10.1145/2993422.2993577
- Bushby, S. T., & Park, C. (2001). Using the Virtual Cybernetic Building Testbed and FDD Test Shell for FDD Tool Development NISTIR 6818 Using the Virtual Cybernetic Building Testbed and FDD Test Shell for FDD Tool Development.
- Casillas, A. Lin, G. Im, P. Granderson, J. Curation of Ground-Truth Validated Benchmarking Datasets for Fault Detection & Diagnostics Tools. Proceedings of the 2020 ACEEE Summer Study on Energy Efficiency in Buildings. (2020)
- Deru, M., Field, K., Studer, D., Benne, K., Griffith, B., Torcellini, P., ... Studer, D. (2011). U . S . Department of Energy Commercial Reference Building Models of the National Building Stock U . S . Department of Energy Commercial Reference Building Models of the National Building Stock. (February 2011).
- Granderson, J., Lin, G., Singla, R., Mayhorn, E., Ehrlich, P., & Vrabie, D. (2018). Commercial Fault Detection and Diagnostics Tools: What They Offer, How They Differ, and What’s Still Needed. Lawrence Berkeley National Laboratory Recent Work, 1–12. <https://doi.org/10.20357/B7V88H>
- Granderson, J., Lin, G., Harding, A., Im, P., & Chen, Y. (2020). Building fault detection data to aid diagnostic algorithm creation and performance testing. *Scientific Data*, 7(1), 1–14. doi.org/10.1038/s41597-020-0398-6
- House, J. M., Vaezi-Nejad, H., & Whitcomb, J. M. (2001). An Expert Rule Set for Fault Detection in Air-Handling Units. ASHRAE Winter Meeting CD, Technical and Symposium Papers, 1005–1018.
- Huang, S., Zuo, W., Vrabie, D., & Xu, R. (2021). Modelica-based system modeling for studying control-related faults in chiller plants and boiler plants serving large office buildings. *Journal of Building Engineering*, 44(May), 102654.doi.org/10.1016/j.job.2021.102654
- Kramer, H., Lin, G., Curtin, C., & Granderson, J. (2020). Proving the Business Case for Building Analytics. *Energy Technologies Area October* <https://doi.org/10.20357/B7G022>
- Li, S. and J. Wen, “Development and Validation of a Dynamic Air Handling Unit Model - Part I (RP 1312) ”, ASHRAE Transactions, Vol. 116, Pt. 1, pp. 45 - 56, 2010.

- Li, S., J. Wen, X. Zhou, and C. J. Klaassen, “Development and Validation of a Dynamic Air Handling Unit Model - Part II (RP 1312) ”, ASHRAE Transactions, Vol. 116, Pt 1, pp. 57 - 73, 2010.
- Liao, H., Cai, W., Cheng, F., Dubey, S., & Rajesh, P. B. (2021). An Online Data-Driven Fault Diagnosis Method for Air Handling Units by Rule and Convolutional Neural Networks.
- Montazeri, A., & Mohamad, S. (2020). Fault detection and diagnosis in air handling using data-driven methods. *Journal of Building Engineering*, 31(March), 101388. doi.org/10.1016/j.jobbe.2020.101388
- Schein, J., & Bushby, S. T. (2006). A Hierarchical Rule-Based Fault Detection and Diagnostic Method for HVAC Systems. *HVAC&R Research*, 12(1), 10–12.
- Wu, S., & Sun, J. (2012). A physics-based linear parametric model of room temperature in office buildings. *Building and Environment*, 50, 1–9. doi.org/10.1016/j.buildenv.2011.10.005
- Yan, R., Ma, Z., Zhao, Y., & Kokogiannakis, G. (2016). A decision tree based data-driven diagnostic strategy for air handling units. *Energy & Buildings*, 133, 37–45. doi.org/10.1016/j.enbuild.2016.09.039
- Yu, Y., Woradechjumnroen, D., & Yu, D. (2014). A review of fault detection and diagnosis methodologies on air-handling units. *Energy & Buildings*, 82, 550–562. doi.org/10.1016/j.enbuild.2014.06.042
- Yun, W., Hong, W., & Seo, H. (2021). A data-driven fault detection and diagnosis scheme for air handling units in building HVAC systems considering undefined states. *Journal of Building Engineering*, 35(December 2020), 102111. doi.org/10.1016/j.jobbe.2020.102111
- Zhong, C., Yan, K., Dai, Y., Jin, N., & Lou, B. (2019). Energy Efficiency Solutions for Buildings: Automated Fault Diagnosis of Air Handling Units Using Generative Adversarial Networks. *Energies*. 1–11. doi.org/10.3390/en12030527