Active Actuator Fault Detection and Diagnostics in HVAC Systems

James Weimer*, Seyed Alireza Ahmadi*, José Araujo*, Francesca Madia Mele†, Dario Papale†, Iman Shames‡, Henrik Sandberg*, Karl Henrik Johansson*

*ACCESS Linnaeus Centre
KTH Royal Institute of Technology
Stockholm, Sweden
weimerj@kth.se,
saahmadi@kth.se, araujo@kth.se,
hsan@kth.se, kallej@kth.se

†Department of Electrical Engineering
Universitá Degli Studi di
Napoli Federico II
Napoli, Italy
francesc.mele@studenti.unina.it
da.papale@studenti.unina.it

‡Department of Electrical
and Electronic Engineering
The University of Melbourne
Parkville, Australia
iman.shames@unimelb.edu.au

Abstract

This paper introduces a new method for performing actuator fault detection and diagnostics (FDD) in heating ventilation and air conditioning (HVAC) systems. The proposed actuator FDD strategy, for testing whether an actuator is stuck in a single position, uses a two-tier approach that includes a dynamic model-based detector and a fast-deciding steady-state detector. The model-based detector is formulated to provide detection performance that asymptotically bounds both the probability of miss and probability of false alarm. To provide a quick confirmation the actuator is working, the steady-state detector utilizes a goodness-of-fit detection strategy to decide if the measurements could be described by an actuator failure. An architecture is introduced that requires multiple steady-state detection experiments to decide that the measurements could be explained by an actuator failure before performing model-based detection. An experimental test bed using a the KTH Royal Institute of Technology campus HVAC system is described and used to evaluate the steady-state and model-based detectors. The experimental test bed is utilized to identify a building dynamics model, that is employed through monte carlo analysis, to characterize the detection performance of both the model-based detector and the steady-state detector.

Categories and Subject Descriptors

G.3 [**Mathematics of Computing**]: Probability and Statistics

General Terms

Security, Experimentation, Design

Buildsys'12, November 6, 2012, Toronto, ON, Canada.

Keywords

Fault Detection, HVAC, Test Beds

1 Introduction

Studies indicate that residential, office and commercial buildings account for nearly 40% and 47% of the U.S. and U.K. energy consumption [26, 25], respectively. Heating, ventilation and cooling (HVAC) are known to be the largest contributor, accounting for 43% of U.S. residential energy consumption. The design of energy-efficient HVAC systems has therefore become a worldwide research priority. Recently, several researchers have studied how to improve the control of HVAC system by deploying more embedded sensors to monitor temperature, humidity, and CO_2 levels [15], using information about occupant behavior [16, 9, 2], and improving the modeling and control approaches [19, 17, 23, 18, 22, 21, 3].

Modern HVAC systems contain an increasing number of sensors and remotely controlled actuators. While the inclusion of these smart devices enable low-cost and environmentally friendly building energy management, undetected sensor and actuator failures can result in poor air quality management and result in societal rejection of these normally efficient systems. Moreover, HVAC Fault Detection and Diagnostic (FDD) schemes which result in unpredictable or sporadic false alarm rates can deter building managers from investigating potential failures. For these reasons, technological development of FDD schemes tailored for HVAC systems is paramount and has received much research interest in the recent years [14, 10, 8, 12].

The importance of FDD in complex technical systems, such as aircraft or power plants, has resulted in a large amount of research and development effort with a variety of solutions. However, its study with application to HVAC systems has only started in the late 1980s, with a particular interest in identifying low-cost, timely, and accurate methods for detecting actuator faults. A thorough review of approaches to HVAC actuator fault detection, diagnostics, and prognostics prior to 2006 is provided in [14, 13]. In general, approaches to HVAC actuator fault detection can be classified as either hardware-based or software-based solutions [14]. The hardware-based solutions introduce additional smart components strictly for the purposes of actuator fault detection and can provide accurate detection capabilities; however, hardware solutions are far more expensive to both deploy and maintain than software-based approaches, and are much more difficult to reconfigure with the introduction of additional smart-actuator devices [13]. Moreover, the inclusion of additional hardware has the added drawback of further increasing the complexity of the HVAC system itself. Software-based actuator FDD approaches are attractive in theory, but suffer from either a reliance on unknown (and difficult to learn) physical models or system-specific detector design specifications [10, 14, 13, 12].

In this paper, we propose a novel distributed software-based active actuator FDD for HVAC systems. The primary difference between the proposed approach, and the other model-based software approaches described in [14, 13] is the proposed approach utilizes a distributed two-tiered detection approach containing a distributed quantitative model-based approach and a distributed qualitative model-based approach to provide quick inference when an actuator is working and provide accurate detection when an actuator has failed. Additionally, the quantitative model-based approach (refereed to henceforth as the model-based approach) does not require full model knowledge as it only relies on the structure of a simplified thermodynamic model (and does not require model parameter knowledge). The introduction of the model-based detector that utilizes only the system structure is a significant difference from other model-based approaches [14]; moreover, the fact that the detector not only detects, but also isolates individual actuator failures is an added contribution. The qualitative model-based approach (refereed to henceforth as the steady-state approach) makes a qualitative decision whether the actuator is working based on a steady-state assumption utilizing a logical relation that when an actuator is working, it will cause a change in the air temperature. The proposed two-tier approach is evaluated in the fault detection of a cold water flow valve of the a real HVAC system at the KTH Royal Institute of Technology, Stockholm, Sweden using the available HVAC sensor measurements and actuation capabilities, where it is demonstrated that fast and accurate detection performance results.

In the following section, modeling of building thermodynamics is discussed and a first-order model is introduced. Section 3 presents a novel actuator FDD scheme consisting of a fast inferring steady-state detector and a robust model-based detector. An experimental test bed is presented in section 4 and is employed to evaluate the actuator FDD approach. The concluding section provides discussion and insight into future work.

2 Modeling Building Dynamics

Modeling buildings accurately enables us to implement advanced control strategies to improve the performance of building HVAC systems and to ensure that such systems are operating free of faults. However, detailed models are often hard to determine due to the inherently complex nature of building thermodynamics. For example, exact temperature distribution in a body mass is characterized by complex partial differential equations that are seldom easy to solve [4]. This is mainly due to two reasons. First, to solve them requires a significant amount of information regarding the building construction and insulating materials, location, weather forecast, time of year, occupancy, usage, equipment contained in the room, etc. Additionally, as PDEs are generally solved via numerical solvers one faces the computational problems that inherently arises from using such solvers. While it is worth noting that there are computer aided building modeling tools that result in complex models, e.g. EnergyPlus [6], these tools require a vast amount of measurements and computation to implement these models, which are only accurate for the learned building scenarios.

Having the aforementioned arguments in mind, in this paper, and similar to [24, 11, 7, 1], we use the first principle of the thermal dynamics modeling to obtain a simple, albeit nontrivial, model for describing temperature evolution in buildings.

In principle, the evolution of the temperature in a building can be described as a combination of the effects of the adjacent air masses with the thermal generation capabilities of the room. A generic thermodynamic model of this interaction is provided through a spatial discretization of the heat equation as

$$m_j \dot{x}_j(t) = \varepsilon_j(t) + \sum_{i \in \mathbb{N}_j} \alpha_{ij}(t) \left(x_i(t) - x_j(t) \right) + \beta_j(t) d_j(t) \tag{1}$$

where air masses in a building are modeled by a graph $G(V,E)$ where $j \in V$ is an air mass and $(j,i) \in E$ if air masses j and i are thermally coupled. Additionally, $x_j(t) \in \mathbb{R}$ is the temperature of the air mass j, m_j is the thermal capacitance of air mass j, $\alpha_{ij} \in \mathbb{R}$ is the thermal coupling (heat transmission) coefficient between air masses i and j, $\varepsilon_j(t)$ is the thermal generation present in air mass j (such as humans, computers, etc.), $\beta_j(t) \in \mathbb{R}$ is the thermal coupling between the air-mass and actuator, and $d_j(t) \in \mathbb{R}$ is the actuator input.

The difficulty in employing the building thermodynamical model is in the requirement that the building thermal coefficients be known. These parameters are notorious for changing with the time of year, the opening and closing of doors and windows, the humidity, etc. For example, the value of α_{ij} changes when a window or door is opened or closed, additionally, it changes with the airflow between adjacent air masses, a breeze between two rooms causes the temperatures in the respective rooms to converge faster (and vice versa). Thus making it very difficult to obtain accurate descriptions for their values. Moreover, the complexity involved in modeling the building behavior grows exponentially with the number of parameters which are known to vary.

While the model does capture the effects that the actuator input has on the temperature of the room, identifying the parameters necessary requires significant learning and may not be attractive or tractable for building environments with volatile changes in occupancy and air movement (such as academic building environments). In this paper we consider those methods that do not make an assumption on knowing the parameters values.

3 Actuator Fault Detection and Diagnostics

Modern building energy management systems require accurate HVAC control to minimize energy usage while maintaining an acceptable level of comfort for the building occupants. Thus, actuator fault detection is necessary to ensure proper building operation as HVAC systems are subject to various aging and operation errors which can lead to hardware malfunction. A common failure in HVAC systems occurs when the actuator "sticks" and no longer changes its set point, despite controller requests. This type of actuator failure can occur in any position. For example, a valve can be stuck fully open, fully shut, or at any intermediate setting. Additionally, being able to isolate specific actuator failures is paramount to performing timely maintenance.

In this section, we introduce a distributed HVAC actuator FDD strategy, tailored for detecting and isolating whether actuators are stuck in an unknown position. This strategy is designed as a two-tier approach consisting of a fast-deciding steady-state detector coupled with a dynamic model-based detector. The following subsections discuss the model-based detector, steady-state detector, and FDD system architecture, respectively.

3.1 Model-Based Detector

A discrete-time model for the room temperature measurements can be generated from the system in (1) using a forward Euler approximation [20] as

$$\begin{aligned} x_{k+1,j} &= x_{k,j} + \sum_{i\in\mathbb{N}_j} a_{k,ij}\left(x_{k,i}-x_{k,j}\right) + b_{k,j}d_{k,j} + e_{k,j} \\ y_{k,j} &= x_{k,j} + v_{k,j}, \end{aligned} \tag{2}$$

where $x_{k,j}$ is the temperature of air-mass j, $y_{k,j}$ is the k-th temperature measurement of room j, $v_k \in \mathbb{R}$ is a zero-mean i.i.d. Gaussian measurement noise with variance σ_j, and the discrete-time lumped-parameters for all $i,j \in \mathbb{I}$ are scaled versions of their continuous-time counterparts,

$$a_{k,ij} = \frac{T_s}{m_j}\alpha_{ij}(kT_s), \quad b_{k,j} = \frac{T_s}{m_j}\beta_j(kT_s), \quad e_{k,j} = \frac{T_s}{m_j}\varepsilon_j(kT_s),$$

assuming T_s is the discrete-time sampling period.

While it is known that the building model parameters are time varying, over short periods of time or when the building is not being utilized, the building parameters tend to remain constant. The model-based detector is designed to determine whether the actuator has failed (has become stuck) or is working properly (not stuck). To isolate the performance of the j-th actuator, the dynamics of the j-th room temperature is written as

$$x_{k+1,j} = x_{k,j} + \sum_{i\in\mathbb{N}_j} a_{k,ij}\left(y_{k,i}-y_{k,j}\right) + b_{k,j}d_{k,j} + e_{k,j} + n_{k,j} \tag{3}$$

where $d_{k,j}$ is the applied actuation input such that $d_{k,j} = u_{k,j}$ when the actuator is operating correctly and $d_{k,j} = d_{k-1,j}$ when it fails (i.e. the actuator sticks). In (3) $n_{k,j}$ is a zero-mean process noise with covariance,

$$cov(n_{k,j}) = \left(\sum_{i\in N_j}\hat{a}_{k,ij}\right)^2\sigma_j + \sum_{i\in N_j}\hat{a}^2_{k,ij}\sigma_i \tag{4}$$

When $\hat{a}_{k,ij} = a_{k,ij}$ the discrete dynamics (3) is statistically equivalent to the discrete-time model in (2). By assuming the parameters are constant,

$$a_{k+1,ij} = a_{k,ij} \quad \text{and} \quad b_{k+1,j} = b_{k,j}$$

The discrete-time dynamics for measurement of the j-th room can be written as

$$\begin{aligned} z_{k+1,j} &= A(d_{k,j})z_{k,j} + w_{k,j} \\ y_{k,j} &= Cz_{k,j} + v_{k,j} \end{aligned} \tag{5}$$

where

$$\begin{aligned} z_{k,j} &= \begin{bmatrix} x_{k,j} & \vec{a}^T_{k,j} & b_{k,j} & e_{k,j} \end{bmatrix}^T \\ A(d_{k,j}) &= \begin{bmatrix} 1 & \vec{y}_{k,ij} - \mathbf{1}y_{k,j} & d_{k,j} & 1 \\ 0 & I & 0 & 0 \\ 0 & \mathbf{0} & 1 & 0 \\ 0 & \mathbf{0} & 0 & 1 \end{bmatrix} \\ w_{k,j} &= \begin{bmatrix} n_{k,j} & 0 & 0 & 0 \end{bmatrix}^T \\ C &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{6}$$

and $\vec{a}_{k,j}$ and $\vec{y}_{k,j}$ are the vectors of the neighboring measurements, $\{y_{k,i} | i \in \mathbb{N}_j\}$, and their respective thermal coupling parameters, $\{a_{k,ij} | i \in \mathbb{N}_j\}$, for the j-th air mass. The measurements have a Gaussian distribution, written as

$$f_j(y_k) = \frac{1}{\sqrt{2\pi\left(C\Sigma_{k,j}C^T+\sigma_j\right)}}\exp\left\{-\frac{1}{2}\frac{\left(y_k - Cm_{k,j}\right)^2}{\left(C\Sigma_{k,j}C^T+\sigma_j\right)}\right\} \tag{7}$$

where

$$\begin{aligned} m_{k+1,j} &= \left(A(d_{k,j}) - K_{k,j}C\right)m_{k,j} + K_{k,j}y_{k,j} \\ \Sigma_{k+1,j} &= \left(A(d_{k,j}) - K_{k,j}C\right)\Sigma_{k,j}A^T(d_{k,j}) + W \\ K_{k,j} &= A(d_{k,j})\Sigma_{k,j}C^T\left(C\Sigma_{k,j}C^T+\sigma_j\right)^{-1} \end{aligned} \tag{8}$$

are the mean and covariance of $z_{k,j}$ and the observer gain.

A test, $\phi_j(y_k) \in \{H_0, H_1, H_{-1}\}$ is employed to denote the decision to that actuator j is working properly ($\phi_j(y_k) = H_0$), has failed ($\phi_j(y_k) = H_1$), or there is not enough information to make a decision ($\phi_j(y_k) = H_{-1}$). This decision is made using the sequential probability ratio test (SPRT) [27], as

$$\phi_j(y_k) = \begin{cases} H_0 & l_j(y_k) \geq \eta_1 \\ H_1 & l_j(y_k) \leq \eta_0 \\ H_{-1} & \text{otherwise} \end{cases} \tag{9}$$

where $l_j(y_k)$ is the log-likelihood ratio,

$$l_j(y_k) = l_j(y_{k-1}) + \ln\frac{f_j(y_k | d_{k,j} = 0)}{f_j(y_k | d_{k,j} = u_{k,j})} \tag{10}$$

and η_0 and η_1 are the test thresholds chosen using Wald's approximation [27] as

$$\eta_0 = \ln\frac{p_M}{1-p_{FA}} \quad \text{and} \quad \eta_1 = \ln\frac{1-p_M}{p_{FA}} \tag{11}$$

where p_{FA} and p_M denote the maximum probably of false alarm and the maximum probably of miss, respectively.

To identify the actuator input for evaluating the detection problem, we utilize an information-theoretic approach and choose the actuator input to maximize the next step Kulbach-Liebner [5] divergence according to

$$u_k = \arg\max_{0 \leq u \leq 1} -\mathbb{E}\left[l_j(y_k)\right] \tag{12}$$

This approach is common in information theory as it results in the control sequence that maximizes the next step log-likelihood ratio. Since the log-likelihood is a convex function of the control sequence, it is maximized at the extreme points of the range of the control sequence. In an HVAC system this equates to either turning the HVAC actuator completely on or completely off. While this control input is advantageous for fault detection and diagnostics, it comes at a trade-off with the performance of the HVAC system since the control input does not correspond to the optimal building operation set-point.

$$u_k = \begin{cases} 1 & \text{if } \mathbb{E}\left[l_j(y_k|u_k=1)\right] > \mathbb{E}\left[l_j(y_k|u_k=0)\right] \\ 0 & \text{if } \mathbb{E}\left[l_j(y_k|u_k=1)\right] \leq \mathbb{E}\left[l_j(y_k|u_k=0)\right] \end{cases} \tag{13}$$

Unlike other model-based approaches to actuator FDD, this approach does not require occupancy data, a priori building parameter assumptions, or centralized computation since each room only requires measurements of the adjacent rooms (or outside air temperature) to test its actuator.

It will be shown in the experimental evaluation section that the model-based detector requires significant monitoring periods to accurately determine whether an actuator has failed. Since the model-based strategy requires that the building dynamics do not change during its monitoring period, it is likely that the model-based detector can only be performed at night. Moreover, fault detection schemes that require long monitoring periods may not be necessary to identify a working actuator. The following section introduces a steady-state detector capable of quickly identifying a working actuator.

3.2 Steady-State Detector

Logic indicates that in the event an actuator is working, applying a significant change in the actuation input will result in a change in the measured temperature. Using this reasoning, this section introduces a steady-state detector that can confirm an actuator is working based on a change in the measured temperature.

To mathematically describe the steady-state detector, we assume that the room (and its surrounding rooms) have reached a steady state condition,

$$x_{k+1,j} = x_{k,j} \tag{14}$$

where $x_j(t) \in \mathbb{R}$ is the temperature of room j. The premise of the steady-state detector is that if an actuator fails, the temperature in the room is not expected to change since no change in actuation will be occur. Thus, a goodness-of-fit test $\bar{\phi}_j(y_k) \in \{H_0, \neg H_0\}$ is employed to determine whether the measured temperatures are likely to be explained by a steady state model, which accurately describes several possible events including, but not limited to, an actuator failure. Another more common reason, besides an actuator failure, that the steady-state model may match the measured data is that the window is open and the outside air mixes quickly with the room air such that no change in the actuator can cause a significant change in the room temperature.

For this test, we assume that the control input at the start of the test is $\bar{u}$, and select the new control input, u_k, at time k to be

$$u_k = \begin{cases} 1 & \text{if } \bar{u} \leq 0.5 \\ 0 & \text{if } \bar{u} > 0.5 \end{cases} \tag{15}$$

this strategy will yield a control input that maximizes the change in the actuation, which improves the detector performance. For a Gaussian process, the goodness-of-fit detector makes a decision using a chi-squared test,

$$\phi_j(y_k) = \begin{cases} H_0 & \bar{l}_j(y_k) \geq \bar{\eta} \\ \neg H_0 & \bar{l}_j(y_k) \leq \bar{\eta} \end{cases} \tag{16}$$

where

$$\bar{l}_j(y_k) = \bar{l}_j(y_{k-1}) + \frac{\left(y_k - \bar{m}_{k,l}\right)^2}{\bar{\Sigma}_{k,l} + \sigma_j} \tag{17}$$

and

$$\begin{aligned} \bar{m}_{k+1,l} &= \left(1 - \bar{K}_{k,j}\right)\bar{m}_{k,l} + \bar{K}_{k,j} y_{k,j} \\ \bar{\Sigma}_{k+1,l} &= \left(1 - \bar{K}_{k,j}\right)\Sigma_{k,l} + \sigma_j \\ K_{k,j} &= \Sigma_{k,l}\left(\Sigma_{k,l} + \sigma_j\right)^{-1} \end{aligned} \tag{18}$$

and the test threshold

$$p_M = \int_0^{\bar{\eta}} \chi_N^2(\bar{l}_j(y_k)) \tag{19}$$

is determined such that the integral from 0 to $\bar{\eta}$ of the chi-squared distribution with N degrees of freedom equals the maximum probability of miss, p_M.

The result of this test is a decision to either accept that the actuator is working properly, or to concede that a decision as to wether the actuator is working can not be made at this time since the reason for not accepting the actuator is working can be described not only by an actuator failure, but also be the existence of non-steady state building conditions or a significant influence from an adjacent air mass (such as opening the windows of a room). For this reason, the steady-state detector can only infer that the actuator is working.

3.3 FDD Architecture

It is expected that faults are infrequent, thus it is preferred to have a test that can quickly identify whether the actuator is working, without having the added testing complexity required to determine if the actuator has failed. Figure 1 illustrates the flow chart for the actuation fault detection and diagnosis system. The system starts with the steady-state detector and decides either that the detector is working (H_0) or that no decision can be made about the state of the actuator ($\neg H_0$) using the quick steady-state detector. If after N^* attempts of running the steady-state detector a decision can not be made about the state of the actuator, then the model-based detector is evaluated until a decision that the actuator is working (H_0) or that a failure has occurred (H_1) results.

The selection of N^* can be made based upon how willing the building operator is to run the model-based detector, which requires significantly longer decision times and

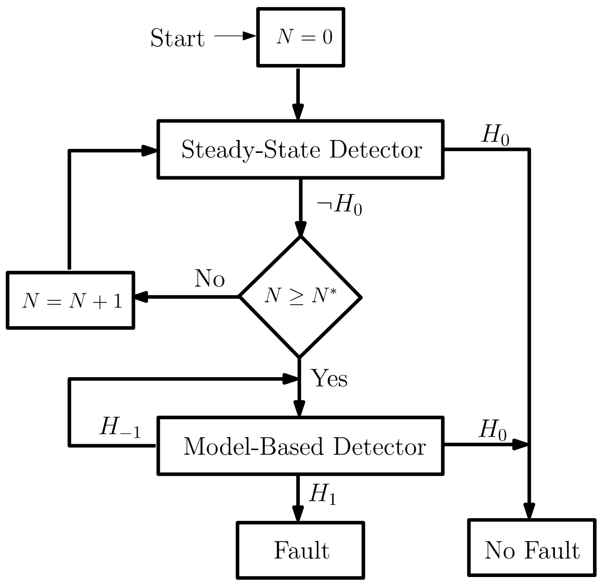

Figure 1. Fault detection and diagnostic architecture

Figure 2. KTH Royal Institute of Technology building SCADA system map

will result in reduced building efficiency. However, excessive testing with the steady-state detector can also result in reduced building energy management performance. The actuator fault detection and isolation strategy introduced in this section is evaluated experimentally in the following section.

4 Experimentation

To evaluate the actuator fault detector, an HVAC test bed was developed. This section both describes the experimental platform and presents the results of the actuator FDD evaluation in the following subsections.

4.1 Experimental Platform

The KTH Royal Institute of Technology main campus in Stockholm, Sweden consists of over 45 buildings which house roughly 559 laboratories, 2569 office rooms and 87 lecture halls. The campus buildings, as depicted in SCADA system map in Fig. 2, are equipped with a central SCADA system which connects all the HVAC units to a centralized campus monitoring and control site. The SCADA system provides detailed information on the consumption of energy for HVAC. Considering the size of the campus and the large

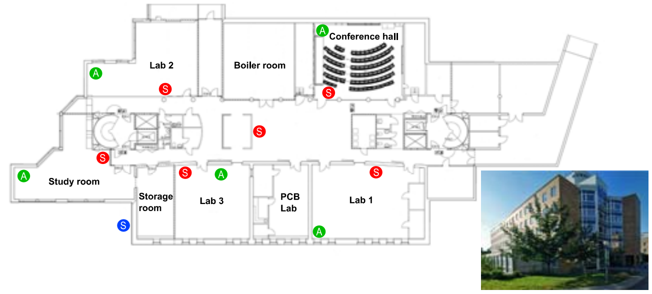

Figure 3. HVAC system deployment in the second floor of the Q-building

number of buildings, we have selected the Q-building (denoted by the black circle in Fig. 2) to evaluate the actuator FDD strategy. The Q-building is a multi-story building housing multiple academic departments, lecture rooms, and laboratories. The Q building is equipped with three separate ventilation units for fresh air supply and relies on a district-managed water supply for induction-based heating and cooling through radiators. The fresh air flow for high usage rooms, such as laboratories and lecture halls, is regulated on weekdays by demand controlled ventilation (DCV) from $08:00$ to $17:00$.

Interaction with the SCADA system is achieved through a web portal that links to an OPC client/server interface. The corresponding secure SCADA web application allows manual actuation of the individual Programmable Logic Controllers (PLCs) and storage of sensory data. As a preliminary deployment, the second floor has been selected to deploy the actuation FDD scheme. This floor houses four laboratories, one conference hall, one storage room and one study room as depicted in Fig. 3 In the figure, the red circles depict sensor locations, the green circles illustrate the actuator locations (which connect to the air conditioning vents) and the blue circle denote the external temperature sensor. The actuators in this system are the cooling valves for the air conditioning system. The cooling valves can actuate between a value of 0 and 1, which corresponds to a percentage of how far open the value is at the current time. The testing environment described in this subsection is utilized in the following subsection to evaluate the actuator FDD scheme.

4.2 Experimental Results

To evaluate the model-based detector performance, multiple experiments were performed utilizing laboratory two in Fig. 3. We assume that the air mass of laboratory two, whose temperature is measured by sensor one, interacts with the outdoor and corridor air masses, whose temperatures are measured by sensors seven and six, respectively. The actuator corresponding to laboratory two is denoted in Fig. 3 as actuator one. The temperature measurements are gathered every five minutes. Laboratory two has two windows which can be manually opened and closed to change how the air mass in laboratory two interacts with the outdoor air mass.

During the experiments, we wish to detect an actuator failure in the air conditioning system, namely whether the cooling valve is stuck or not. To emulate an actuator failure, we simply do not apply the calculated control value to the

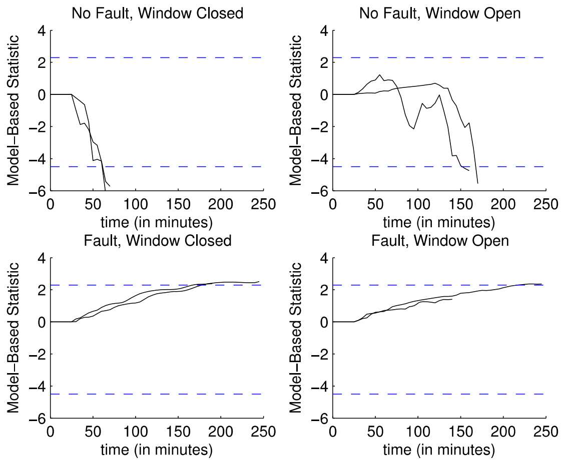

Figure 4. Experimental model-based detector results

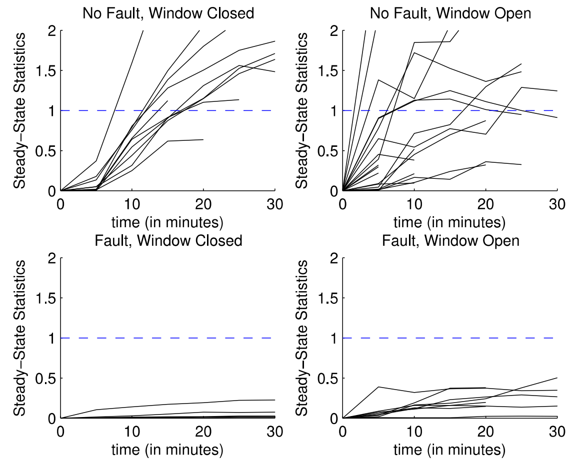

Figure 5. Experimental steady-state detector results

cooling valve actuator and merely leave it in its current state (as if it were stuck). During all of the experiments, the door to laboratory two was shut such that the only changes in the air mass interactions were caused by the opening and closing of the windows. Additionally, the actuators in the adjacent rooms were allowed to operate normally, emulating a normal testing scenario where we wish to not only detect, but also isolate. Under these testing conditions, both the steady-state detector and the model-based detector were evaluated under four different cases, namely (a) windows closed and actuator working, (b) windows open and actuator working, (c) windows closed and actuator stuck, (d) windows open and actuator stuck.

The experimental results for evaluating the model-based detector are shown in Fig. 4, where in all four subplots the horizontal axis represents the time since the test began in minutes and the vertical axis indicates the log-likelihood statistic value for the tests performed under the respective testing scenario, as calculated by using (10). The two dashed lines in each subplot indicate the decision regions of the test, $\phi_1(y_k)$, where it is decided that an actuator failure occurs ($\phi_1(y_k) = H_1$) when the statistic is greater than the dashed line, and conversely, it is decided that the actuator is working properly ($\phi_1(y_k) = H_0$) when the log-likelihood statistic is less than the dashed line. When the statistic is between the two dashed lines, no decision can be made regarding the state of the actuator ($\phi_1(y_k) = H_0$). Each line in the subplots represents a single experiment, where all the experiments for each testing condition (fault vs. no fault and window closed vs. window open, as described earlier in this subsection) are plotted together to help illustrate trends. The results in Fig. 4 suggest that the model-based detector accurately identifies whether a fault occurs regardless of the state of the window, if given enough time. These results indicate that when a fault does not exist, it takes about 3 times longer to detect the failure when the window is open as opposed to closed. However, we observe that when a fault occurs, the performance of the model-based detector is relatively unaffected by the window state. From the results in Fig. 4, we observe that testing can require significant monitoring periods ranging on the scale of hours. Since the model-based test is based on the assumption that the room state does not change, it is likely that this test must be performed at night when it is unlikely that the windows, doors, and occupancy will change.

The experimental results for the steady-state detector are shown in Fig. 5, where similar to Fig. 4, the horizontal axis represents the time since the test began in minutes and the vertical axis indicates the chi-squared statistic value. The dashed line in each subplot indicates the decision regions of the test, where it is decided that an actuator has not failed ($\bar{\phi}_1(y_k) = H_0$) when the statistic is greater than the dashed line and conversely no decision is made ($\bar{\phi}_1(y_k) = \neg H_0$) if the statistic is less than the dashed line. In contrast to the model-based detector, we observe that the steady-state detector makes a much quicker decision, where a decision that the actuator has not failed is typically made within thirty minutes (if it occurs). When no fault occurs and the the windows are closed, the detector tends to quickly and accurately indicate that the actuator has not failed. However, when the window is open, the detector is sporadic in deciding that a fault has not occurred. This behavior motivates our testing assumption that no decision about the faultiness of the actuator can be made when the test statistic is less than the decision threshold because the resulting decision is likely to have a high probability of false alarm. When an actuator fault occurs, regardless of the window state, the steady-state detector tends to not indicate no fault occurs, which exhibits a low probability of miss.

The parameters of the building system were estimated using experimentally gathered measurements and the grey box system identification toolbox in MATLAB. The purpose of estimating the parameters is two fold. Firstly, it serves to motivate the usage of only the underlying model structure in the model-based detector and not to employ a priori parameter values which may be incorrect. Secondly, it provides models over which to provide a preliminary statistical evaluation of the detector, which otherwise would require a prohibitively large number of experimental tests (each taking

Table 1. Estimated model parameters

window	a_{21}	a_{31}	b_1	c_1
closed	0.209	0.0310	-0.299	0.130
open	0.006	0.0003	-0.071	0.049

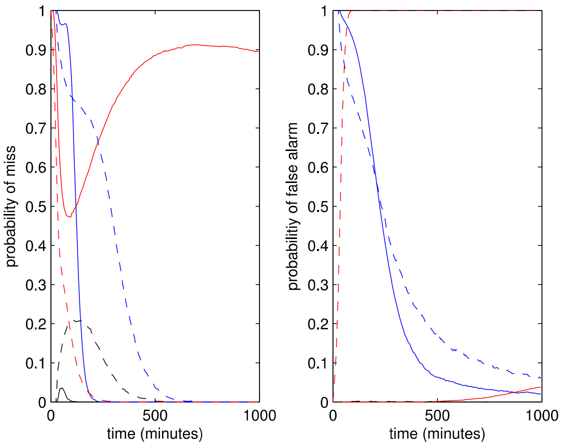

Figure 6. Simulated probability of False Alarm and probability of miss vs. time

multiple hours). The resulting parameter values are provided in Table 1 In the table, the top row lists all the model parameters of the experiments and the bottom two rows indicate their values. To calculate the parameter values, we used the data gathered for the model-based test when there was no fault when the windows were (a) open and (b) closed. Based upon these tests, we observe that there is a significant change in the magnitude of the parameters based upon whether the windows are open or closed. Additionally, we note that the parameter value for b is negative since turning the cooling valve on results in a decrease in temperature, which is consistent with what was expected.

4.3 Simulated Results

To investigate statistical performance of the actuator FDD scheme, $20,000$ monte carlo tests were performed using the nominal parameters in Table 1. The parameters were utilized to generate the temperature dynamics assuming the model in (1). Half the tests $(10,000)$ were performed assuming no fault, while the other half $(10,000$ tests) were performed assuming the actuator was stuck in a random position. The resulting false alarm rate and miss rate are provided in Fig. 6. Each subplot in the figure contains multiple lines indicating the rate of miss. The solid black and dashed black lines represent the rate of error for the model-based approach when the window is closed and open, respectively. Similarly, the solid red and dashed red lines represent the rate of error for the steady-state approach when the window is closed and open. Additionally, the solid blue and dashed blue lines indicate the rate of indecision for the model-based approach.

We observe that the false alarm rate is effectively zero for the model-based detector, while the miss rate has a maximum value of 0.4 when the window is close and 0.21 when the window is open. The error rates converge to zero as the testing time is increased. However, for the steady-state test, the false alarm rate increases slowly with the test time when the window is closed, and increases quickly to 1 when the window is open. This indicates that there is a significant difference in the false alarm rate of the steady-state test depending on whether the window is open or closed, which is consistent with the experimental results in the previous subsection.

The miss rate for the steady-state test has a different performance profile. When the door is closed, the miss rate of the steady-state test achieves a minimum value of 0.47 when the test time is 30 minutes and the window is closed, while the miss rate deceases to zero when the window is open (although this is not preferred since the corresponding false alarm rate when the window is open approaches 1 as time increases, indicating poor performance).

The rate of indecision (blue lines) of the model-based test tend to decrease as the test time increases. This indicates that extended monitoring periods will be able to make statistically significant decisions; however, this assumes that the state of the room does not change over the monitoring period. Thus there is a trade off between the length of monitoring time and the likelihood that the room does not change state and the model assumptions are satisfied. From these simulated results we observe that the steady-state detector is only useful for small monitoring times since extended monitoring periods yield high probability of false alarm. Conversely, the model-based test performance improves as the monitoring period is extended, where for small monitoring periods (under 1 hour), the probability of error is high. These results are consistent and motivate the proposed actuator fault detection and diagnostic scheme where the steady-state detector is used to quickly test if the detector is working. Upon multiple cases where the steady-state detector does not indicate the actuator is working, the model-based detector is utilized to determine if an actuator failure has occurred.

5 Discussion and Future Work

The proposed actuator FDD strategy presented in this paper uses a unique two-tier approach that includes a dynamic model-based detector and a fast-deciding steady-state detector for testing whether an actuator is stuck in a single position. The steady-state detector is shown to quickly confirm if an actuator is working, while the model-based approach can accurately decide whether an actuator fault has occurred regardless of the interaction between the surround air masses. An experimental test bed using a real HVAC system is described and used to evaluate the steady-state and model-based detectors, as well as to estimate the model parameters for a monte carlo analysis of the detection performance.

Future work includes extensions of the FDD strategy to include the detection of sensor failure as well as consider time varying actuator failures (such as detecting whether the actuator range of motion has decreased). Future experiments and evaluations are planned utilizing the entire HVAC test bed across multiple floors and buildings.

6 References

[1] J. Široký, F. Oldewurtel, J. Cigler, and S. Prívara Experimental analysis of model predictive control for an energy efficient building heating system. *Applied Energy*, 88(9):3079 – 3087, 2011.

[2] Y. Agarwal, B. Balaji, S. Dutta, R. Gupta, and T. Weng. Duty-cycling buildings aggressively: The next frontier in hvac control. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, pages 246 –257, April 2011.

[3] A. Aswani, N. Master, J. Taneja, D. Culler, and C. Tomlin. Reducing transient and steady state electricity consumption in hvac using learning-based model-predictive control. *Proceedings of the IEEE*, 100(1):240 –253, Jan. 2012.

[4] W. Chow. Application of computational fluid dynamics in building services engineering. *Building and Environment*, 31(5):425–436, 1996.

[5] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991.

[6] D. Crawley, L. Lawrie, F. Winkelmann, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte, et al. Energyplus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331, 2001.

[7] K. Deng, P. Barooah, P. Mehta, and S. Meyn. Building thermal model reduction via aggregation of states. In IEEE Proceedings of the *American Control Conference (ACC)*, pages 5118–5123., July, 2010.

[8] N. Djuric and V. Novakovic. Review of possibilities and necessities for building lifetime commissioning. *Renewable and Sustainable Energy Reviews*, 13(2):486 – 492, 2009.

[9] V. Erickson, M. Carreira-Perpinan, and A. Cerpa. Observe: Occupancy-based system for efficient reduction of hvac energy. In *10th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 258 –269, Apr. 2011.

[10] N. Fernandez, M. Brambley, S. Katipamula, H. Cho, J. Goddard, and L. D. b. Self correcting hvac controls project final report PNNL-19074. Technical report, Pacific Northwest National Laboratory, Richland, WA., 2009.

[11] M. Gouda, S. Danaher, and C. Underwood. Building thermal model reduction using nonlinear constrained optimization. *Building and Environment*, 37(12):1255–1265, 2002.

[12] L. Jagemar and D. Olsson. The EPBD and continuous commisioning. Technical report, CIT Energy Mangement AB, Goteborg, Sweden, Oct. 2007.

[13] S. Katipamula and M. R. Brambley. Methods for fault detection, diagnostics, and prognostics for building systems - a review, part ii. *HVAC&R Research*, 11(2):169-187, Apr. 2005.

[14] S. Katipamula and M. R. Brambley. Methods for fault detection, diagnostics, and prognostics for building systems - a review, part i. *HVAC&R Research*, 11(1):3-25, Jan. 2005.

[15] Y. Kim, T. Schmid, M. B. Srivastava, and Y. Wang. Challenges in resource monitoring for residential spaces. In *Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, BuildSys '09, pages 1–6, New York, NY, USA, 2009.

[16] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse. The smart thermostat: using occupancy sensors to save energy in homes. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, SenSys '10, pages 211–224, New York, NY, USA, 2010.

[17] J. Ma, J. Qin, T. Salsbury, and P. Xu. Demand reduction in building energy systems based on economic model predictive control. *Chemical Engineering Science*, 67(1):92 – 100, 2012.

[18] Y. Ma, A. Kelman, A. Daly, and F. Borrelli. Predictive control for energy efficient buildings with thermal storage: Modeling, stimulation, and experiments. *Control Systems, IEEE*, 32(1):44 –64, Feb. 2012.

[19] A. Marchiori and Q. Han. Distributed wireless control for building energy management In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, BuildSys '10, pages 37–42, New York, NY, USA, 2010.

[20] A. Melman. Geometry and convergence of euler's and halley's methods. In *SIAM Review*, volume 39, pages 728–735. Society for Industrial and Applied Mathematics, Dec. 1997.

[21] T. X. Nghiem, M. Behl, R. Mangharam, and G. J. Pappas. Scalable scheduling of building control systems for peak demand reduction. In *American Control Conference*, Jun. 2012.

[22] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45(0):15–27, 2012.

[23] H. A. I. S. Goyal and P. Barooah. Zone level control algorithms based on occupancy information for energy efficient buildings. In *American Control Conference*, Jun. 2012.

[24] J. Seem, S. Klein, W. Beckman, and J. Mitchell. Transfer functions for efficient calculation of multidimensional transient heat transfer. *Journal of heat transfer*, 111:5, 1989.

[25] UK Department of Trade and Industry. *DTI, Energy Trends 2005*. Department of Trade and Industry, London, 2011.

[26] U.S. Department of Energy. *Energy Efficiency Trends in Residential and Commercial Buildings*. U.S. Department of Energy, Oct. 2008.

[27] A. Wald. *Sequential Analysis*. John Wiley & Sons, Inc., New York, 1947.