

Parameter-Invariant Actuator Fault Diagnostics in Cyber-Physical Systems with Application to Building Automation

James Weimer¹, José Araujo², Mani Amoozadeh², Seyed Alireza Ahmadi²,
Henrik Sandberg², and Karl Henrik Johansson²

¹ Department of Computer and Information Sciences
School of Engineering and Applied Sciences
University of Pennsylvania
Philadelphia, PA 19104, USA
weimerj@seas.upenn.edu

² ACCESS Linnaeus Centre, School of Electrical Engineering
KTH Royal Institute of Technology
Stockholm, Sweden 10044
{araujo, maniam, saahmadi, hsan, kalle}@kth.se

Abstract. This paper introduces a robust method for performing active 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 given position, is designed on using an invariant hypothesis testing approach and is an improvement of a previous strategy that employed an adaptive detection strategy. The parameter-invariant detector is formulated to provide a constant detection performance, invariant to unknown building parameters, and it is described how this approach can replace the adaptive detector in the previous work. A closed-loop experimental HVAC testbed at the KTH Royal Institute of Technology campus in Stockholm, Sweden is introduced and employed to evaluate the parameter-invariant detector.

Keywords: building automation, fault detection and diagnostics (FDD), invariant hypothesis testing, heating ventilation and air conditioning (HVAC) systems

1 Introduction

Heating, ventilation and cooling (HVAC) are known to be the largest consumer of energy in buildings, accounting for 43% of U.S. residential energy consumption. The design of energy-efficient HVAC systems has therefore become a worldwide research priority. In the U.S. and U.K., buildings consume nearly 40% and 47% of the national energy, respectively [28, 27]. Due to this high usage, there exists a high potential for energy consumption improvement, which has thrust HVAC system operations to the forefront of world-wide research agendas. Recently, several researchers have studied how to improve the control of HVAC systems by

deploying more embedded sensors to monitor temperature, humidity, and CO₂ levels [17], using information about occupant behavior [18, 11, 4], and improving the modeling and control approaches [21, 19, 25, 20, 24, 23, 5].

To achieve an increase in building efficiency requires an increase in the number of sensors and actuators deployed. While the inclusion of these smart devices enables low cost and environmentally friendly building energy management systems, undetected sensor and actuator failures can result in poor temperature and air quality management. Moreover, HVAC Fault Detection and Diagnostic (FDD) schemes which result in unpredictable or erratic performance 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 [16, 12, 10, 14].

The study of HVAC FDD systems has only been investigated since 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 [16, 15]. In general, approaches to HVAC actuator fault detection can be classified as either hardware-based or software-based solutions [16]. The hardware-based solutions introduce additional smart components strictly for the purposes of actuator fault detection and 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 [15]. 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 [12, 16, 15, 14].

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 [1], an HVAC FDD architecture is introduced that utilizes a fast-deciding steady-state detector and an adaptive model-based detector that are applied to a cooling vent fault detection. The primary difference between the previous approaches, and other model-based software approaches described in [16, 15] is the use of a 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 de-

tection when an actuator has failed. Logic indicates that in the event that an actuator is working, applying a significant change in the actuation input results in a measurable change in the temperature. Under this reasoning, the steady-state detector quickly identifies operational actuators, but it tends to result in a high probability of false alarm when the actuator is operational, but its effect on the temperature is dampened (i.e. a window is open such that the temperature is not significantly affected through actuation). To reduce the probability of false alarm in the HVAC FDD strategy, an adaptive model-based detector is employed based on a first-order building thermal dynamic model. The model-based detector utilizes a history of measurements to estimate the unknown model parameters, then employs the estimated model to detect faults. While the adaptive model-based detector was shown to have a significantly lower probability of false alarm in comparison to the steady-state detector, the resulting performance was highly variant depending on the environmental variables (windows open vs. closed, outside air temperature, etc.) which directly affected the underlying parameter estimation. This sporadic false alarm behavior is undesirable in HVAC FDD schemes as it leads to mistrust by building managers.

In this work, the primary contribution is a parameter-invariant detector (to replace the previously developed model-based detector [1]) that maintains a constant probability of miss across all operating conditions of the HVAC FDD strategy. Additionally, the parameter-invariant detector does not require full model knowledge as it only relies on the structure of a simplified thermodynamic model. The fact that the detector not only detects, but also isolates individual actuator failures is an added contribution. A second contribution of this work is the introduction and description of a closed-loop HVAC monitoring and control system that interfaces directly with the KTH Royal Institute of Technology HVAC system, located in Stockholm, Sweden. An evaluation of the system is provided using the experimental testbed that illustrates the performance of the parameter-invariant detector.

In terms of notation, we use plain lower case italic fonts to indicate scalars or functions with scalar range, bold lower case italic fonts to indicate vectors or functions with vector range, and plain upper case italic fonts to indicate matrices. We also use \otimes to denote Kronecker products, and $e_{i,j}$ to denote the elementary vector of dimension i consisting of all zeros with a single unit entry in the j -th position.

In the following section, we motivate and formulate the actuator FDD problem for building automation. Section 3 introduces the parameter-invariant detector and a novel diagnostic input design. The KTH Royal Institute of Technology HVAC test bed is described in section 4 and an evaluation of the parameter-invariant detector is provided in section 5. The concluding section provides discussion and insight into future work.

2 Problem Formulation

In this section, we formulate a distributed actuator fault detection problem for HVAC systems. While precise thermal modeling of buildings is an ongoing science [6, 8], it has repeatedly been experimentally demonstrated that a first principle of the thermal dynamics model is accurate for zone-level temperature evolution in buildings [26, 13, 9, 3]. We consider a building with M interconnected temperature zones for which there exists an underlying interconnection graph, $\mathcal{G}(\mathcal{V}, \mathcal{E})$, between the M zones, where $\mathcal{V} := \{1, \dots, M\}$ is the vertex set, with $i \in \mathcal{V}$ corresponding to zone i , and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set of the graph. The undirected edge $\{i, j\}$ is incident on vertices i and j if zones i and j interact. The neighborhood of zone i , \mathcal{N}_i , is defined as

$$\mathcal{N}_i := \{j \in \mathcal{V} \mid \{i, j\} \in \mathcal{E}\} \quad (1)$$

A generic thermodynamic model of the zone interactions is provided through a spatial and temporal discretization of the first-order heat equation as

$$\begin{aligned} x_j(k+1) &= x_j(k) + m_j \sum_{i \in \mathcal{N}_j} a_{ji} (x_i(k) - x_j(k)) + b_j d_j(k) + w_j(k) \\ y_j(k) &= x_j(k) + v_j(k) \end{aligned} \quad (2)$$

where:

- $k = 0, \dots, T$ is the time index (T even for notational simplicity³);
- $j = 1, \dots, M$ is the zone index;
- the temperature $x_j(k)$'s, measurements $y_j(k)$'s and actuator inputs $d_j(k)$'s are scalar;
- m_j is the volume of air contained in zone j ;
- $a_{ji} = a_{ij} \in \mathbb{R}$ and $b_j \in \mathbb{R}$ denote respectively the gains between $x_i(k)$ and $x_j(k+1)$, and between $d_j(k)$ and $x_j(k+1)$;
- $w_j(k), v_j(k) \in \mathbb{R}$ are uncorrelated i.i.d. Gaussian process noise and measurement noise with moments

$$\mathbb{E}[w_j(k)] = \chi_{j,w} \quad \mathbb{E}[v_j(k)] = \chi_{j,v} = 0,$$

$$\mathbb{E}[(w_j(k) - \chi_{j,w})^2] = \sigma_{j,w}^2 \quad \mathbb{E}[(v_j(k) - \chi_{j,v})^2] = \sigma_{j,v}^2.$$

We note that in the event of a zone containing multiple actuators and a single sensor, actuator FDD can be performed by allowing only a single actuator to vary. In doing this, the stationary actuators are effectively lumped into the process noise mean since it is a constant input into the thermal model. To compact

³ For ease of notation and without loss of generality we assume that the available measurements are over a given period whose length is fixed *ex ante*.

the notation we let, for $j = 1, \dots, M$,

$$\begin{aligned} A &:= [\alpha_{ij}] \\ \alpha_{ij} &:= \begin{cases} 1 - m_j \sum_{n \in \mathcal{N}_j} a_{nj} & \text{if } i = j \\ m_j a_{ij} & \text{if } i \in \mathcal{N}_j, \quad i \neq j \\ 0 & \text{otherwise} \end{cases} \\ B &:= \text{diag}[b_1, \dots, b_M] \\ \mathbf{y}_j &:= [y_j(0), \dots, y_j(T)]^\top \\ \mathbf{d}_j &:= [d_j(0), \dots, d_j(T)]^\top. \end{aligned}$$

Additionally, we consider the following quantities, assuming $\mathcal{N}_j = \{i_1, \dots, i_J\}$ is the sorted list of neighboring zones of zone j . Then

$$\begin{aligned} \bar{\alpha}_j &:= [\alpha_{i_1 j}, \dots, \alpha_{i_J j}]^\top, \\ \bar{y}_j(k) &:= [y_{i_1}(k), \dots, y_{i_J}(k)]^\top, \\ \bar{\mathbf{y}}_j &:= [\mathbf{y}_{i_1}^\top, \dots, \mathbf{y}_{i_J}^\top]^\top, \end{aligned}$$

where, $\bar{y}_j(k)$ is the set of the measurements of agent j and its neighbors (sorted lexicographically) at time k , while $\bar{\mathbf{y}}_j$ is the set of *all* the measurements of agent j and its neighbors (again sorted lexicographically).

Consider then a *specific* zone $\ell \in \{1, \dots, M\}$, containing an actuator. The structure of the actual actuator input \mathbf{d}_ℓ is assumed to be as follows:

- $\mathbf{u}_\ell := [u_\ell(0), \dots, u_\ell(T)]^\top$ is a *desired* and *known* actuation signal;
- $\theta_\ell \in \mathbb{R}$ is an unknown (but constant) input.

Then

$$\mathbf{d}_\ell = \theta_\ell \mathbf{1} + \mu_\ell \mathbf{u}_\ell \quad (3)$$

where the binary scalar $\mu_\ell \in \{0, 1\}$ is an unknown test parameter indicating whether the actuation signal is present ($\mu_\ell = 1$) or absent ($\mu_\ell = 0$).

We summarize the available information for detecting an actuator failure in zone ℓ as follows:

Assumption 1 Available information:

- the time-series measurements $\bar{\mathbf{y}}_\ell$
- the local desired actuation signal \mathbf{u}_ℓ ;
- the local zone air-volume weight m_ℓ ;
- when an actuator fails, its input to the system remains constant;
- the fact that the state dynamics are LTI-Gaussian, constant in time, and with $b_\ell \neq 0$.

For completeness, we summarize the unavailable information as:

Assumption 2 Unavailable information:

- all the time-series measurements except \vec{y}_ℓ
- all the local desired input signals except \mathbf{u}_ℓ ;
- all the local weights except m_ℓ ;
- the weights A and B ;
- the moments of the process and measurement noises $\chi_{j,w}, \sigma_{j,w}^2, \sigma_{j,v}^2$, for all $j = 1, \dots, M$;
- the actuation parameters θ_j and μ_j for all $j = 1, \dots, M$;
- the initial conditions $x_1(0), \dots, x_M(0)$;
- the input signals $\mathbf{d}_1, \dots, \mathbf{d}_M$.

We then assume the unknown μ_ℓ to be either 0 (actuator ℓ is at fault) or 1 (actuator ℓ not at fault) and pose the following binary hypothesis testing problem:

Assumption 3 Structure of the actuator fault μ_ℓ satisfies either one of the two following hypotheses:

H_0 (null hypothesis):	$\mu_\ell = 1$	(no fault)
H_1 (alternative hypothesis):	$\mu_\ell = 0$	(fault)

In words, both hypotheses assume the actual \mathbf{d}_ℓ to be unknown, since θ_ℓ is unknown, but with a fixed and known functional structure. H_0 additionally assumes the presence of the known actuation input \mathbf{u}_ℓ . Our aim is the following: develop a distributed test that considers a **specific** zone $\ell \in \{1, \dots, M\}$, and decides among the hypotheses H_0 vs. H_1 in Assumption 3 using only the information in Assumption 1 and, at the same time, being invariant to the unavailable information in Assumption 2. Thus, we state the following problem:

Problem 4

Find a test that detects whether zone ℓ has an actuator fault independently of whether a fault exists at any other zone $j \neq \ell$ (fault isolation) *and* minimizes the probability of false alarm while maintaining a constant probability of detection.

The following section presents a solution to the problem introduced in this section.

3 Parameter-Invariant Actuator FDD

In this section, we introduce a distributed HVAC actuator FDD strategy, tailored for detecting and isolating whether actuators are stuck in an unknown position such that a constant level of detection is maintained. To achieve this goal, the following subsections introduces a parameter-invariant detector and an actuator diagnostic input, respectively.

3.1 Parameter-Invariant Detector

In this section, we recall the test developed for distributed detection of inputs in networked systems in [2]. This test, in the context of the HVAC detection problem, is designed to minimize the probability of false alarm, subject to a constraint on the probability of detection. We state the primary result of [2], augmented for the actuation detection problem, in the following lemma:

Lemma 5 A maximally invariant statistic for Problem 4 is

$$T[\mathbf{z}_\ell] = \frac{\mathbf{z}_\ell^\top P_\ell \mathbf{z}_\ell}{\frac{1}{N_\ell - 1} \mathbf{z}_\ell^\top (I_{N_\ell} - P_\ell) \mathbf{z}_\ell} \quad (4)$$

with

$$\begin{aligned} \mathbf{z}_\ell &:= F_\ell Q \mathbf{y}_\ell \\ P_\ell &:= \frac{F_\ell Q \mathbf{u}_\ell \mathbf{u}_\ell^\top Q^\top F_\ell^\top}{\mathbf{u}_\ell^\top Q^\top F_\ell^\top F_\ell Q \mathbf{u}_\ell} \\ N_\ell &:= \frac{T}{2} - \|\mathcal{N}_\ell\|_0 - 1 \end{aligned} \quad (5)$$

and where the exploited quantities satisfy

$$\begin{aligned} F_\ell^\top F_\ell &= I_{\frac{T}{2}} - \vec{Y}_\ell (\vec{Y}_\ell^\top \vec{Y}_\ell)^{-1} \vec{Y}_\ell^\top \\ Q &= I_{\frac{T}{2}} \otimes [0 \ 1] \\ \vec{Y}_\ell &= \begin{bmatrix} \vec{y}_\ell^\top(0) & 1 \\ \vec{y}_\ell^\top(2) & 1 \\ \vec{y}_\ell^\top(4) & 1 \\ \vdots & \vdots \\ \vec{y}_\ell^\top(T) & 1 \end{bmatrix} \end{aligned} \quad (6)$$

Applying the maximally invariant statistic, and following the same reasoning as in [2], we write the test for detecting actuator failures as:

Corollary 6 A distributed test minimizing the probability of false alarm and providing a constant probability of missed detection of α for Problem 4 is

$$\phi_\ell(\mathbf{z}_\ell) = \begin{cases} H_0 & \text{if } T_\ell[\mathbf{z}_\ell] > \mathcal{F}_{1, N_\ell-1}^{-1}(\alpha) \\ H_1 & \text{otherwise} \end{cases} \quad (7)$$

where $\mathcal{F}_{n,m}^{-1}(\alpha)$ is the inverse central cumulative F -distribution of dimensions n and m .

We remark that test (7) can be performed in simultaneously across multiple zones and it is invariant to the non-local measurements. This comes with a price, namely, the test exploits half the measurements for testing (the other half are used to establish invariance). To maximize the performance across this reduced data set, the following subsection introduces an adaptive actuator diagnostic signal.

3.2 Diagnostic Input Design

The performance of the detector is significantly affected by the actuation input driving the test. In this subsection, and motivated by the performance of the adaptive model-based detector in [1], we design a diagnostic input that attempts to maximize the divergence of test.

To design the diagnostic actuation input, we observe that the discrete-time dynamics for measurement of the j -th zone can be written as

$$\begin{aligned} \mathbf{z}_j(k) &= \mathbf{G}_k \mathbf{z}_j(k) + \mathbf{n}_j(k) \\ y_j(k) &= \mathbf{C} \mathbf{z}_j(k) + v_j(k) \end{aligned} \quad (8)$$

where

$$\begin{aligned} \mathbf{z}_j(k) &= [x_j(k), m_j \boldsymbol{\alpha}_j^T, b_j, \chi_{j,w} + b_j \theta_j]^T \\ \mathbf{G}_k &= \begin{bmatrix} 1 & \mathbf{y}_j^T(k) - \mathbf{1}^T y_j(k) & \mu_j u_j(k) & 1 \\ 0 & \mathbf{I} & 0 & 0 \\ 0 & \mathbf{0}^T & 1 & 0 \\ 0 & \mathbf{0}^T & 0 & 1 \end{bmatrix} \\ \mathbf{n}_j(k) &= [w_j(k) + \sum_{i \in \mathbb{N}_j} \alpha_{ij} v_i(k), \mathbf{0}^T, 0, 0]^T \\ \mathbf{C} &= [1 \ \mathbf{0}^T \ 0 \ 0] \end{aligned} \quad (9)$$

As a heuristic, we assume (strictly for the purpose of designing an actuation input) that the true values of $\boldsymbol{\alpha}_{ij}$ are equal to the corresponding values provided through $z_j(k)$. Under this assumption, the measurements have a Gaussian distribution, parameterized by μ_j , written as

$$f_j(y_j(k)) = \frac{1}{\sqrt{2\pi (\mathbf{C} \boldsymbol{\Sigma}_{k,j} \mathbf{C}^T + \sigma_{j,v}^2)}} \exp \left\{ -\frac{1}{2} \frac{(y_j(k) - \mathbf{C} \mathbf{m}_{k,j})^2}{(\mathbf{C} \boldsymbol{\Sigma}_{k,j} \mathbf{C}^T + \sigma_{j,v}^2)} \right\} \quad (10)$$

where, assuming $\Sigma_{j,n} = \mathbb{E} [n_{k,j} n_{k,j}^\top]$

$$\begin{aligned} \mathbf{m}_{k+1,j} &= (\mathbf{G}(k) - \mathbf{K}_{k,j} \mathbf{C}) \mathbf{m}_{k,j} + \mathbf{K}_{k,j} y_j(k) \\ \Sigma_{k+1,j} &= (\mathbf{G}(k) - \mathbf{K}_{k,j} \mathbf{C}) \Sigma_{k,j} \mathbf{G}^\top(k) + \Sigma_{j,n} \\ \mathbf{K}_{k,j} &= \mathbf{G}(k) \Sigma_{k,j} \mathbf{C}^\top (\mathbf{C} \Sigma_{k,j} \mathbf{C}^\top + \sigma_{j,v}^2)^{-1} \end{aligned} \quad (11)$$

are the mean and covariance of $\mathbf{z}_j(k)$ and the observer gain, 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 [7] divergence according to

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

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

$$l_j(k) = l_j(k-1) + \ln \frac{f_j(y_j(k) | \mu_j = 0)}{f_j(y_j(k) | \mu_j = 1)}. \quad (13)$$

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 as denoted as follows

$$u_k = \begin{cases} 1 & \text{if } \mathbb{E} [l_j(k+1 | u_j(k) = 1)] > \mathbb{E} [l_j(k+1 | u_j(k) = 0)] \\ 0 & \text{if } \mathbb{E} [l_j(k+1 | u_j(k) = 1)] \leq \mathbb{E} [l_j(k+1 | u_j(k) = 0)] \end{cases} \quad (14)$$

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.

It will be shown in the experimental evaluation section that the parameter-invariant detector requires significant monitoring periods to accurately determine whether an actuator has failed. Moreover, fault detection schemes that require long monitoring periods may not be necessary to identify a working actuator if the actuator has a significant effect on the temperature. For this reason, the parameter-invariant detector is best suited to replace the model-based adaptive detector in the previous work.

4 Experimental Testbed

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 has an HVAC system managed by a centralized SCADA system. The SCADA map of the KTH campus is depicted

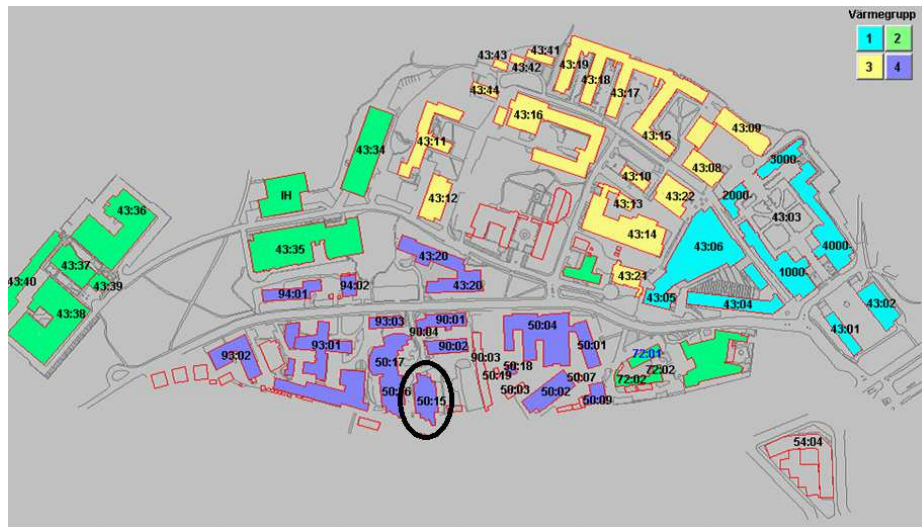


Fig. 1. KTH Royal Institute of Technology campus view from the SCADA system map.

in Fig. 1. Each building contains at least one Programming Logic Controller (PLC) units which measure and control the local HVAC system components. Each of the PLC units in the campus communicates to the SCADA system through an OPC client/server interface. The KTH HVAC test bed is currently deployed in the Q-building (denoted by the black circle in Fig. 1). The Q-building is a multi-story building housing the School of Electrical Engineering with multiple academic departments, lecture rooms, and laboratories. This 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. The fresh air supply takes place from 7 : 00 AM to 4 : 00 PM, and can be set manually by demand at other times of the day and weekends.

The HVAC testbed is currently comprised of the second floor of the Q-building and is depicted in Fig. 2. This floor houses four laboratories (rooms A225, A213, B230 and the PCB Lab), an office room (A235), the Q2 lecture hall, one storage room and a boiler room.

Each room in the testbed is considered to be a thermal zone and has a set of sensors and actuators that can be individually controlled. In the figure, the red circles depict sensor locations, the green circles illustrate the actuator locations and the blue circle denotes the external temperature sensor. The available sensors are temperature (GQ101) and CO₂ sensors (GT101)⁴. The actuators are the flow valve of the heating radiator (SV201), the flow valve for the air conditioning system (SV401), the air vent for fresh air flow at constant temperature of 21 °C

⁴ In this section we note, in parentheses, the PLC tags corresponding to each sensor and actuator such that they can be referenced in the downloaded KTH HVAC testbed data.

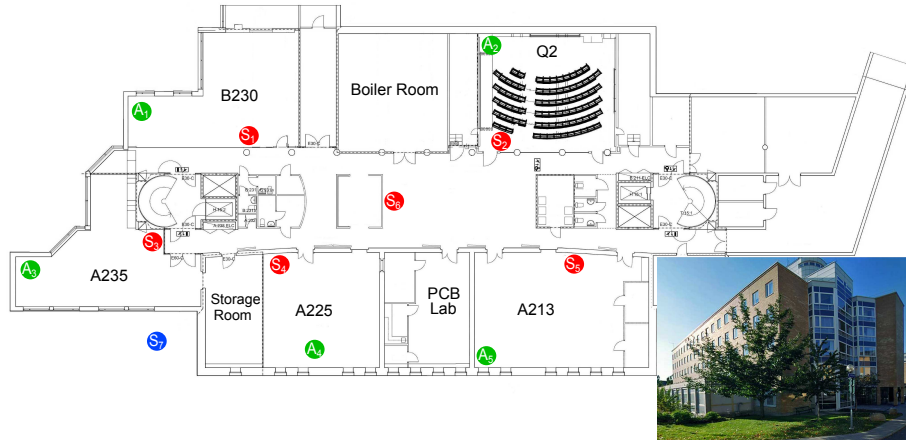


Fig. 2. KTH HVAC testbed at the second floor of the Q-building at KTH. Each of the five rooms considered contain sensors and actuators used for HVAC control. Additional sensors are located in the corridor and outside of the building.

(ST901) and the air vent for air exhaust to the corridor (ST902). Each actuator can be set between 0% and 100%. As an illustration, the HVAC system in room A225 is depicted in Fig. 3.

The HVAC testbed is developed in LabVIEW and is comprised of two separate components: the experimental application and a database/web server system. The database is responsible for logging the data from all HVAC components in real-time, which is publicly available through a web server (<http://hvac.ee.kth.se>). Additionally, the experimental application is developed by each user and interacts with the Data-logging and Supervisory Control (DSC) module in the HVAC Testbed Server, which connects to the PLC through and OPC client/server. This component allows for real-time sensing, computation and actuation. Even though the application is developed in LabVIEW, MATLAB code is integrated in the application through a Mathscript zone. An overview of the testbed architecture is shown in Fig. 4.

5 Experimental Results

The evaluation of the actuator fault detector was performed in the KTH HVAC testbed. To evaluate the parameter-invariant detector performance, multiple experiments were performed utilizing room A225 in Figs. 2 and 3 as the test room. The air mass of room A225 interacts with the outdoor and corridor air masses as well as the adjacent rooms, the PCB lab and the storage room (each representing a unique thermal zone). Since the PCB Lab and Storage room do not have sensors, we neglect their effect on room A225's temperature. The effect of ignoring the potential thermal contribution of these unobservable air masses is minimized by including room A235 and room A213 as *adjacent* rooms (or zones).

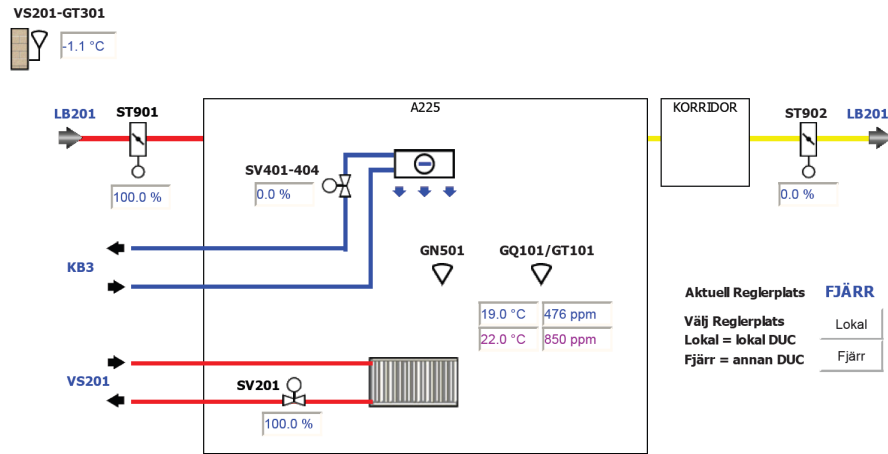


Fig. 3. The HVAC system components in room A225, the Automatic Control experimental lab. Various sensors and actuators are available allowing for the control of ventilation and heating.

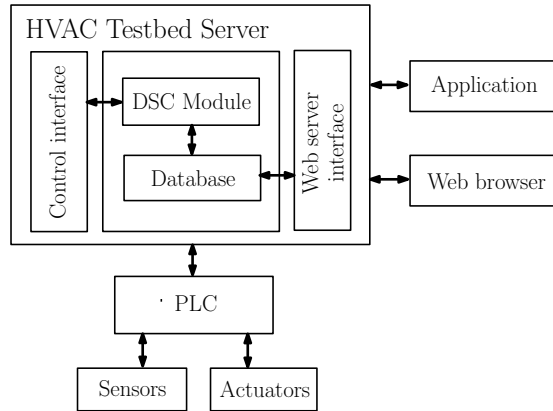


Fig. 4. The HVAC system architecture. Users are able to design experiments through a Labview application and remotely connect to the HVAC testbed. Additionally, through a web browser any user can download experimental data from the testbed database.

We note that room A225 has four exterior windows and one door connecting to the common corridor.

To evaluate the actuator fault detector, we attempt to detect an actuator failure in the air conditioning system, namely whether the fresh air vent, actuator ST901 in Fig. 3, is stuck in a given position. The system is set to measure the temperature at five minute intervals. In the following experiments, we neglect the tests using measurements gathered for less than 30 minutes (6 sampling

periods), as these tests yield irrelevant results since it requires at least 6 sets of measurements to calculate the test statistic for the parameter-invariant detector.

To emulate an actuator failure, we simply do not apply the control value given by the diagnostic input mechanism and leave the air vent closed (as opposed to physically breaking the actuator). The cooling actuator SV401, the radiator valve SV201, and the exhaust air vent ST902 were kept at constant values throughout each experiment. Additionally, the actuators in the adjacent rooms were allowed to operate normally, as in a normal operating scenario where we wish to not only detect the actuator failure, but do so in a distributed manner without PLC coordination. Under these testing conditions, the parameter-invariant detector was evaluated for the scenarios which exhibited erratic performance in [1]. Specifically, these scenarios are (1) detecting an actuator fault when the windows are open/closed and (2) detecting there is not an actuator fault when the windows are open. The reason for not evaluating the parameter-invariant detector when the actuator is working and the windows are closed is because the steady-state detector in [1] is very accurate under this scenario, thus the parameter-invariant detector is unlikely to be utilized.

To evaluate the parameter-invariant detector when an actuator is in fault and the windows are closed, 100 unique experiments were performed, each lasting three hours. For the parameter-invariant detector, we specify a probability of miss (probability of deciding there was no fault when there actually was a fault) of 0.10. The statistical results of these tests are shown in Fig. 5, where the upper subplot illustrates the average value of the test statistic (solid black line) versus the maximum and minimum value of the test statistic (dotted black lines) and the test decision threshold (dashed red line) for the tests, while the lower subplot illustrates the actual rate of miss by the black x's and the specified probability of miss of 0.10. When a test statistic is above the threshold, then the test decides there was no fault (which is incorrect in these experiments), and by specification should happen with a probability of 0.10, regardless of the test time. From Fig. 5, we observe that, the actual rate of missed detection varies between 0.06 to 0.16, which is very near the specified rate of 0.10, for all time. Averaging the miss rate over the three hours yields a rate of 0.101, which is nearly identical to the probability of miss specification. These results illustrate that when the windows are closed, the parameter-invariant detector accurately detects actuator faults as specified by the probability of miss.

Consistent with previous experiments, to evaluate the parameter-invariant detector when an actuator is in fault and the windows are open, 100 unique experiments were performed, each lasting three hours and employed the same specified probability of miss (0.10). The statistical results of these tests are shown in Fig. 6, where the plots follow the same structure as in Fig. 5. In Fig. 6, we observe very similar performance as in Fig. 5. This similarity is expected since by changing the state of the window (opening the window), we have merely changed the interaction between room A225 with the outside. Since the parameter-invariant detector is designed to be invariant to different thermal zone interactions, the performance (in terms of the probability of miss) should be unaffected. In com-

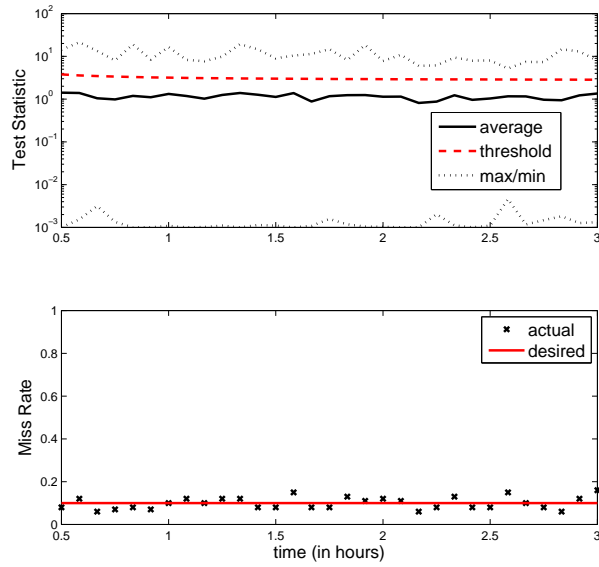


Fig. 5. Experimental parameter-invariant detector results when the actuator is at fault and the window is closed.

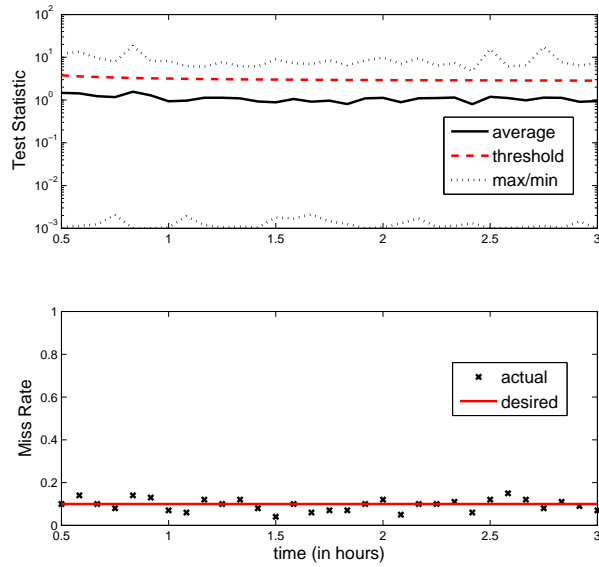


Fig. 6. Experimental parameter-invariant detector results when the actuator is at fault and the window is open.

parison to the performance of the model-based detector in [1], we observe that the parameter-invariant detector has nearly constant performance (in terms of probability of miss) with both the state of the window and with time, while the model-based detector exhibits varying performance with both the state of the window and with time. Having near-constant performance (which matches the specification) is preferred as it allows a building manager to reliably select the probability of missed detection of actuator fault.

To evaluate the parameter-invariant detector when an actuator is working properly and the windows are open, 25 experiments were performed, each lasting three hours. For the parameter-invariant detector, and consistent with the previous experiments, we specify a probability of miss to be 0.10. The statistical results of these tests are shown in Fig. 7, where the upper subplot follows the same structure as in Figs. 5 and 6. The lower subplot illustrates the rate of false

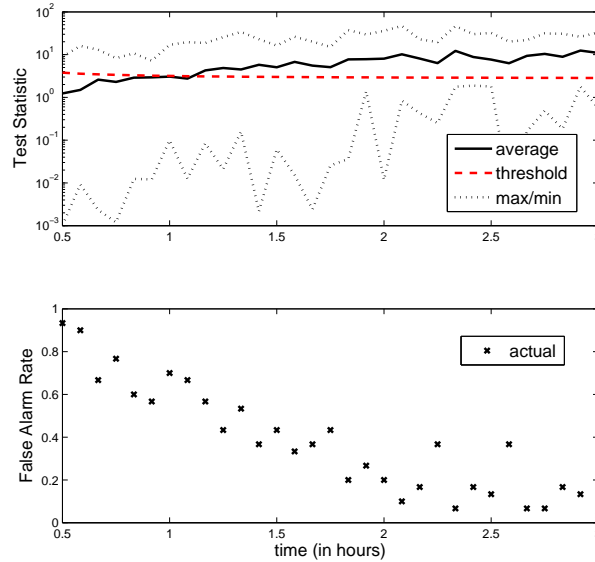


Fig. 7. Experimental parameter-invariant detector results when the actuator is working and the window is open.

alarm (deciding there is an actuator fault when in reality there is not a fault) versus time. In these experiments (and opposite of the previous experiments) a false alarm occurs when the test statistic is below the threshold. In Fig. 7, we immediately observe that the test statistic is (in general) increasing as the test time increases. This is desirable since the larger the test statistic, the more likely it is to claim no fault (which is true in these experiments). As time increases,

we observe from the lower subplot that the rate of miss is (in general) decreasing. The reason for non-monotonic performance is explained by the fact only 25 experiments were used to evaluate the parameter-invariant detector when the actuator is working and the windows are open. These results illustrate that performance (in terms of false alarm rate) can be improved (decreased) by allowing the parameter-invariant detector to run for longer time periods. For this room and configuration, a false alarm rate of 0.05 can be achieved by allowing the parameter-invariant detector to run for 3 hours. Depending on the zone (room) and its interactions with the adjacent zones, the probability of false alarm will vary. However, this variance is generally acceptable in practice so long as the probability of miss remains constant since it implies that by simply letting the test run longer will yield improved performance.

6 Discussion and Future Work

The parameter-invariant detector introduced in this work for building HVAC systems is based on a previously designed CFAR detector for networked systems, where the parameter-invariant detector with constant performance is designed to replace the model-based detector with unpredictably varying performance in the HVAC FDD scheme previously developed. The parameter-invariant detector is designed to maintain a constant probability of missing a fault, invariant to the unknown and time-varying building parameters. An experimental testbed using a real HVAC system is described that allows automatic sensing and actuation of several HVAC components. Future work includes a full evaluation of the two-tier fault detection strategy on the KTH HVAC testbed and extending the detection theory to handle the detection of faulty sensors and imminent actuator failure (i.e. detecting whether the actuator range of motion has decreased).

References

1. J. Weimer, A. Ahmadi, J. Araujo, F. Mele, D. Papale, I. Shames, H. Sandberg, K. H. Johansson, Active Actuator Fault Detection and Diagnostics in HVAC Systems, In *4th ACM Workshop On Embedded Sensing Systems For Energy-Efficiency In Buildings (BuildSys 2012)*
2. J. Weimer, D. Varagnolo, K. Johansson, Distributed Model-Invariant Detection of Unknown Inputs in Networked Systems In *2nd ACM International Conference on High Confidence Networked Systems (HiCoNS) 2013*.
3. J. Široký, F. Oldewurtel, J. Cigler, and S. Prívvara Experimental analysis of model predictive control for an energy efficient building heating system. *Applied Energy*, 88(9):3079 – 3087, 2011.
4. 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.
5. 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.

6. W. Chow. Application of computational fluid dynamics in building services engineering. *Building and Environment*, 31(5):425–436, 1996.
7. T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991.
8. 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.
9. 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.
10. N. Djuric and V. Novakovic. Review of possibilities and necessities for building lifetime commissioning. *Renewable and Sustainable Energy Reviews*, 13(2):486 – 492, 2009.
11. 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.
12. N. Fernandez, M. Brambley, S. Katipamula, H. Cho, J. Goddard, and L. Dinh. Self correcting HVAC controls project final report PNNL-19074. Technical report, Pacific Northwest National Laboratory, Richland, WA., 2009.
13. M. Gouda, S. Danaher, and C. Underwood. Building thermal model reduction using nonlinear constrained optimization. *Building and Environment*, 37(12):1255–1265, 2002.
14. L. Jagemar and D. Olsson. The EPBD and continuous commissioning. Technical report, CIT Energy Management AB, Goteborg, Sweden, Oct. 2007.
15. 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.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.

22. 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.
23. 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.
24. 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.
25. 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.
26. 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.
27. UK Department of Trade and Industry. *DTI, Energy Trends 2005*. Department of Trade and Industry, London, 2011.
28. U.S. Department of Energy. *Energy Efficiency Trends in Residential and Commercial Buildings*. U.S. Department of Energy, Oct. 2008.
29. A. Wald. *Sequential Analysis*. John Wiley & Sons, Inc., New York, 1947.