DEVELOPMENT OF NEW WHOLE BUILDING FAULT DETECTION AND DIAGNOSIS TECHNIQUES FOR COMMISSIONING PERSISTENCE

A Dissertation

by

GUANJING LIN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	David E. Claridge
Committee Members,	Charles H. Culp
	Dennis L. O'Neal
	Timothy J. Jacobs
Head of Department,	Jerald A. Caton

December 2012

Major Subject: Mechanical Engineering

ABSTRACT

Commercial building owners spent $167 billion for energy in 2006. Building commissioning services have proven to be successful in saving building energy consumption. However, the optimal energy performance obtained by commissioning may subsequently degrade. The persistence of savings is of significant interest. For commissioning persistence, two statistical approaches, Days Exceeding Threshold-Date (DET-Date) method and Days Exceeding Threshold-Outside Air Temperature (DET-Toa) method, are developed to detect abnormal whole building energy consumption, and two approaches called Cosine Similarity method and Euclidean Distance Similarity method are developed to isolate the possible fault reasons. The effectiveness of these approaches is demonstrated and compared through tests in simulation and real buildings. The impacts of the factors including calibrated simulation model accuracy, fault severity, the time of fault occurrence, reference control change magnitude setting, and fault period length are addressed in the sensitivity study. The study shows that the DET-Toa method and the Cosine Similarity method are superior and more useful for the whole building fault detection and diagnosis.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Claridge for his great guidance, support and encouragement through my Ph.D. study. Also I want to thank Dr. Charles Culp, Dr. Dennis O'Neal, and Dr. Timothy Jacobs, my dissertation committee, for their valuable suggestions on my dissertation and presentation. I appreciate my colleagues and friends at ESL: Hiroko Masuda and Dr. Lei Wang for their helpful discussions and suggestions.

Finally, special thanks to my mother, father, husband and my little boy Victor for their love and support.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

Page

Page

Page

Page

LIST OF TABLES

Page

Page

Page

Page

Page

1. INTRODUCTION

This section describes the background and objective of the work presented in this dissertation as well as gives a brief description of the contents of the sections to follow.

1.1 Background

Commercial building owners spent $167 billion for energy in 2006 and this cost is expected to rise to $227 billion by 2030(EIA 2009). One-tenth of energy savings in the commercial buildings sector means the saving of over $15 billion per year. According to the U.S. Department of Energy (DOE), HVAC systems consume 40%-60% of the energy used in U.S. residential and commercial buildings. Building commissioning services, which either ensure that building HVAC systems are installed and operated to provide the performance envisioned by the designer or identify and implement optimal operating strategies for buildings as they are currently being used, have proven to be successful in saving building energy consumption. A broad study of 224 new and existing commercial buildings in 21 states across the country, commissioned by 18 different commissioning service providers, netted a median savings of 15% of whole building energy use (Mills et al.2005). The Energy Systems Laboratory at Texas A&M University (TAMU) started Continuous Commissioning® (CC®) in 1992. During the last 20 years, the CC® process has produced average energy savings of about 20% without significant capital investment in over 300 large buildings in which it has been implemented (Claridge et al. 2004).

Though commissioning services are effective in reducing building energy consumption, the optimal energy performance obtained by commissioning may

subsequently degrade. Faults in an HVAC system can increase HVAC energy consumption by 30% (Brambley et al. 1998). Several researchers have indicated this problem in their papers. A study of the persistence of savings in ten university buildings over a period of 12 years found that on average $1000/year of savings in heating or cooling would decrease to $750/year in three to five years if there was no follow-up effort to maintain the savings (Toole and Claridge, 2011). The largest consumption increase was due to significant component failures and/or control changes that did not compromise comfort but caused large changes in consumption (Turner et al. 2001, Claridge et al. 2004). These major increases were not identified until two years had passed, and hundreds of thousands of dollars in excess energy costs had already occurred. Peterson (2005) analyzed the energy performance after commissioning of three buildings in Oregon and found that 89% of the electric savings but none of the gas savings persisted four years after commissioning due to numerous control overrides which had been made at the zone or box level.

The long-term persistence of commissioning energy savings hinges on the ability of the operator to track system performance and troubleshoot the systems (Friedman et al. 2003). Building energy consumption is the most attractive flag to the building operators since it directly links to the utility bill. An automated building energy performance analysis tool can help the building operator continuously monitor building energy consumption, alert operations personnel early after the onset of significant increases in consumption and assist them in identifying the problem.

Fault detection and diagnosis (FDD) is the core part of the automated building energy performance analysis tool. Fault detection is a process of determining if there are faults in the building. Fault diagnosis is a process of identifying the possible causes of the detected fault. Extensive research has been conducted during the past decade to develop different FDD technologies that are suitable for building HVAC systems (Katipamula and Brambley 2005a and 2005b, Friedman and Piette 2001). Computer models of the whole building or of systems and components form the basis of HVAC FDD in monitoring routine operations. They can quantitatively indicate the intended fault-free performance and hence offer a benchmark against which to compare the actual measured performance. Physical redundancy, expert system rules, statistical bands and user selected thresholds are commonly used to detect faults. Fuzzy logic scoring systems, artificial neural network (ANN) classifiers, rule-based classifiers, and other classifiers are usually used to diagnose faults.

There are two fundamental approaches to fault detection and diagnosis in buildings: a component level (bottom-up) approach and a whole building (top-down) approach (Seem 2007). The component level approach focuses on detecting faults in individual systems such as air-handling units (AHUs) (Norford et al. 2002, Comstock and Braun 1999), variable-air-volume (VAV) boxes (Xu and Haves 2002, Wang and Qin 2005), chillers (Grimmelius et al. 1995, Reddy et al. 2003), or boilers (Elyas et al. 2009, Borsjie, H.J. 1999). The whole-building approach specializes in abnormal behavior in high-level measurements such as the whole building cooling, heating or electrical consumption (Haberl and Claridge 1987, Haberl et al. 1989, Claridge et al. 1999, Yu and

Dolf 2003, Friedman and Piette 2001, Lee and Claridge 2007, Curtin 2007, Seem 2007, and Song et al. 2008).

Although there have been research efforts to develop FDD methodologies for building HVAC systems for over a decade, most studies focus on detecting and diagnosing faults at the system component level, and there have been limited investigations of FDD at the whole building level. FDD at the whole building level could report abnormal building energy consumption and narrow the list of possible fault causes to a few options. It needs much less data compared to FDD at the component level, and thus saves time and effort. Another benefit of the whole building FDD is that it would not replace the operators' work with a diagnostic tool. The operators would be happy to be a part of the FDD process. The literature review indicates several whole building fault detection techniques, but very few whole building fault diagnosis approaches were found in available literature. This indicates a need for new whole building FDD techniques for application in automated building energy performance analysis tools.

1.2 Research Objective

The goal of this research is to develop new whole building HVAC fault detection and fault diagnosis approaches for Continuous Commissioning® persistence, and demonstrate their effectiveness through tests in simulation and real buildings. This objective is achieved in the following five steps:

1. Develop practical whole building HVAC fault detection and diagnosis techniques that

a. are capable of identifying abnormal building energy consumption performance

b. are capable of showing the severity of abnormal performance

c. are capable of indicating possible cause(s) for the detected abnormal energy consumption

d. can be developed based only on energy consumption and weather data

e. are based on objective metrics

f. are affordable, simplified and efficient

g. are robust under varying HVAC systems

2. Design whole building fault detection and diagnosis programs using the proposed techniques.

3. Test the whole building fault detection and diagnosis techniques with simulated fault-free and faulty data.

4. Test the whole building fault detection and diagnosis techniques with faulty field data from real buildings.

5. Study the sensitivity of whole building fault detection and diagnosis approaches through simulation.

The investigated building energy consumption includes cooling and heating consumption. The proposed fault detection approach is intended to detect persisting small increases or decreases in the normal energy consumption which have a significant impact when allowed to continue for a period of weeks, months or sometimes years. The approach should focus on the faults that truly have a significant energy consumption impact. The agenda of the proposed fault diagnosis approach is to narrow the classifying faults into a subset of possibilities at the whole building level rather than specify the

fault to a certain component. The fault diagnosis approach will not replace the operator or technician but will enable them to find the real fault cause and finish the fault correction more quickly and efficiently. The sensitivity of the fault detection and diagnosis approaches will focus on the impact of fault severity, calibrated simulation model accuracy, the fault time of occurrence, and the fault period length.

1.3 Outline of Contents

In this dissertation, Section 2 supplies a literature review about 1) HVAC FDD methods overview; 2) previous attempts at whole building fault detection and diagnosis; 3) the application of the Shewhart charts in the HVAC fault detection field; and 4) similarity measures and their application in FDD. Section 3 reviews the ABCAT methodology, and introduces the methodology of the proposed whole building fault detection and diagnosis approaches. Sections 4 and 5 record the tests of the proposed FDD approaches with simulated data. Section 6 describes tests of the proposed FDD approaches with field data. Sections 7 and 8 describe the process for a sensitivity study and give its results. Finally, a summary of current and future work is given in Section 9.

2. LITERATURE REVIEW

In the past few decades, researchers have been dedicated to the study of fault detection and diagnosis technology applied to HVAC systems. This literature review covers 1) HVAC FDD methods overview; 2) previous attempts at whole building fault detection and diagnosis; 3) the application of the Shewhart charts in the HVAC fault detection field; 4) similarity measures and their application in FDD; and 5) provides a summary.

2.1 HVAC FDD Methods Overview

The FDD technology was introduced into building HVAC systems in the 1970s, and systematic research began in the 1980s and continues to the present time. The International Energy Agency sponsored two international collaborative activities in this area: Annex 25 and Annex 34. The International Energy Agency Annex 25 source book (IEA 1996) describes typical faults in HVAC systems, the development of various methods for detecting and diagnosing HVAC equipment and control faults, and the test results with simulation and laboratory data. Annex 34 demonstrates the robustness and commercial feasibility of multiple FDD tools through tests in real buildings (IEA 2001). This overview discusses the general techniques applied in the HVAC FDD field. HVAC FDD at the whole building level is emphasized in Section 2.2.

There are essentially two types of HVAC faults, namely hard failures and performance degradations (IEA 1996). Hard failures are severe and abrupt faults. For example, a fan belt is broken. Performance degradations such as heat exchanger fouling

are mild faults which don't cause a failure but lead to deviation in operating performance from targeted performance.

The HVAC FDD process is commonly described using the following steps. First, the FDD system must collect and pre-process the data to extract the features reflecting the correct or normal performance of the building system or equipment. Second, the system evaluates the observed data with the extracted features. A fault is detected if the observed data exceeds its normal, predicted, or tolerable range. Then, the system analyzes the data that suggest a fault is present to develop a diagnosis and determines the cause of the fault (Rossi and Braun 1997, Friedman and Piette 2001). Sometimes the fault detection and diagnosis are combined in a single step. Residuals (differences between actual and modeled values) are typical extracted features that represent the normal/abnormal system behavior. Researchers build mathematical models to predict the system performance, and analyze the residuals between the measured data and the expected value. This process is called analytical redundancy and is widely used in HVAC FDD.

Computer models can usually be used to determine or predict the correct system performance. They broadly fall into two general categories: first principles-based models, which are based on the physical principles governing the building system behavior, and data-driven models that are built primarily on experience about the building system performance (Portland Energy Conservation Inc. 2003). The first principles-based models always incorporate quantitative or qualitative models of a building system to define how the system or the component should perform. Data-driven

models are primarily empirical. They rely heavily on training data sets – mostly historical measurements, which relate inputs to the corresponding outputs, to develop viable models of the relationship between the inputs and the outputs. Linear or multiple linear regression, logistic regression, and artificial neural networks (ANNs) are typical methods to build data-driven models. The first principles-based models tend to have greater accuracy than that of the data-driven models (Katipamula and Brambley 2005a), although this is not always true (Reddy et al. 2003).

Expert systems (Han et al 1999; House et al. 2001), statistical bands and user selected thresholds are usually used to detect faults. An expert system creates rules derived from expert knowledge in the form of if-then-else statements. To give an example of a rule, IF the control valve is closed AND the supply fan is running AND the temperature rising across the heating coil is greater than the combined uncertainty of the sensor readings THEN the operation is faulty (Haves and Khalsa 2000). Statistical bands define a threshold based on analysis using estimates of means or residuals' standard deviations. User selected thresholds allow users to choose and vary thresholds to satisfy their acceptable detection sensitivity and false alarm rate.

Fault diagnosis is always thought of as a kind of pattern recognition. Features or signatures for how specific systems or system parameters behave when they have a fault are developed or predicted and stored in a database as reference patterns. Fuzzy logic scoring systems (Grimmelius et al.1995), ANN classifiers (Lee et al. 1996), rule-based classifiers (Haves and Khalsa 2000), and other classifiers are implemented to compare the symptoms of a current fault against the database of the reference fault signatures.

The one with the most similar fault pattern in the reference signature bank would be the cause of the observed fault.

2.2 Previous Attempts at Whole Building FDD

Whole building fault detection and diagnostics uses top-level information about building energy consumption (e.g., electricity, gas, chilled water) to assess whether or not a building and its systems operate efficiently(TIAX 2005). This approach can be expected to spot larger problems, e.g. problems which typically lead to an increase of 5% or more in energy use (Claridge et al. 1999). Representative previous studies about whole building FDD are discussed here.

Haberl and Claridge (1987) implemented an expert system for a university recreation center. Daily measured energy consumption was compared with the consumption predicted by a multivariate linear regression model developed from historical measurements. A standard error of the estimation above or below the predicted consumption was applied as the fault detection threshold. An expert system including several IF-THEN rules was used to determine the possible causes. A downside to this system is that it is not a generic system and can't be implemented in other buildings, and the consumption prediction techniques were statistically difficult mainly due to the identification of the predictor variables in the regression.

Haberl et al. (1989) introduced a similar system that used PRISM, the Princeton Scorekeeping Method, a three-parameter and steady–state model to predict the energy use. PRISM regresses the energy consumption against the heating or cooling degree days. A fault is identified when the absolute residual between the actual and the

predicted consumption exceeds a specified deviation (either 15% or 25% away from the predicted use was chosen in the example). A potential problem with this system is its applicability to commercial buildings, as the three-parameter model often does not apply to variable types of energy consumption seen in commercial buildings (Haberl and Cho 2004).

Claridge et al. (1999) discussed the applications of whole-building diagnostics which had proven to be practical for identifying and improving the operation of over 100 buildings. The paper summarizes whole-building diagnostic procedures and splits them into two major categories – examination of the time series data, and use of physical or empirical models in the analysis of whole-building data streams. The former approach could be used to diagnose shut-off opportunities and operating anomalies. The latter one could be used to diagnose a large variety of system problems which include: VAV behavior as CAV systems, simultaneous heating and cooling, excess supply air, excess outside air, sub-optimal cold/hot deck schedule, high duct static pressure, etc.

Yu and Dolf (2003) presented a Fuzzy Neural Networks (FNN) model to detect faults using the energy use of the whole building. The threshold for the fault detection was derived from twice the maximum multiplication of the moving mean value and variance in 24 hours in the fault-free process. The drawback for this threshold is that it is only suited for detecting failure faults and is not suitable for detecting degradation faults. Additionally, no diagnosis information is mentioned in the paper.

There are three commercially available tools with whole building FDD capability: Facility Explorer, Performance and Continuous Re-commissioning Analysis

Tool (PACRAT), and Whole Building Diagnostician (WBD) (TIAX 2005). Table 1 lists the developers of the tools, the building systems they serve, and some of the faults that they can detect from whole building energy consumption.

Table 1 Diagnostic tool inputs and detected faults (TIAX 2005)

Diagnostic Tool (Developer)	**Input From Following Systems**	**Fault Detected by Systems**
Facility Explorer (Johnson Controls)	Whole Building Energy	• High or low whole building energy consumption relative to prior baseline data for that building
PACRAT (Facility Dynamics Engineering)	Whole Building Energy	• Utility deviation from baseline • Unoccupied operation
	Other	• Economizer/Air handling unit • Central plant (chiller) • Distribution (hydraulic) • Zone control
Whole Building Diagnostician (Pacific Northwest National Laboratory)	Whole Building Energy	• Electric and gas consumption deviation from baseline
	Other	• Economizer/AHU function • Central plant and distribution

Seem (2007) introduced the method implemented in Facility Explorer and presented the field test results. Abnormal energy use in buildings is detected based on daily readings of energy consumption and peak consumption. Outlier identification using the "generalized extreme studentized deviate many-outlier procedure" is applied to determine the features of the particular day that are significantly different from the features for the same day type. All three detected problems in the field test are hard

failure faults. From the review of all known literature, it appears that Facility Explorer has no fault diagnosis information other than detecting that energy consumption is abnormal.

The whole building energy analysis of PACRAT and WBD compares the actual consumption with the consumption predicted by the multiple-variable bin-based model to indicate whether the actual energy usage is abnormal. PACRAT creates bins by the hour of day, day of week and any third variable, typically ambient temperature (Friedman and Piette 2001). In the WBD, the bins are defined by ambient temperature, relative humidity and hour of week. The energy use in the baseline chosen from historical measurement is averaged in each bin. A daily energy consumption index (ECI), the ratio of actual energy use to expected energy use, is used to show deviations in the actual energy use from the expected use. The WBD alerts the building operator when the ECI values are statistically significantly greater than or less than 1.0 (Katipamula et al. 2003). Neither PACRAT nor WBD provide the potential causes for abnormal consumption, but they do estimate the cost impact of the abnormal consumption.

Song et al. (2008) developed a tool for fault detection and diagnosis of building air-conditioning systems based on indoor air temperature changes and energy consumption increase. A simulation was run to predict the energy consumption and the room temperature degradation from the normal values under different faults. A flow chart was created by organizing these deviated values. Then each fault can be estimated

by following the flow chart with measured deviated values as the starting point from the normal values.

Lee and Claridge (2007) examined the use of the ASHRAE simplified energy analysis procedure (SEAP) for fault detection at the whole-building level. A calibrated SEAP model is used for predicting the cooling and heating consumption during a post-commissioning period. Visual comparison with measured post-commissioning data is used to detect significant deviations from expected performance. The procedure was applied retrospectively to three years of measured consumption data as a test. Energy deviation is shown in monthly, daily percent and cumulative formats. Three significant operational changes were identified during the test period.

Curtin (2007) developed a prototype Automated Building Commissioning Analysis Tool (ABCAT) following the approach of Lee et al. (2007). A "Cumulative Cost Difference" plot is the primary fault detection metric used in the ABCAT prototype (Curtin 2007). It continuously computes and plots the algebraic sum of the daily differences between the measured and simulated consumption multiplied by a user specified energy cost. The positive values of cumulative difference on the plot indicate that the measured consumption exceeds the expected consumption and vice versa. Three significant energy consumption deviations were found with this plot in the online test. The "Cumulative Cost Difference" plot would directly show how the fault influences the cost, but it is a kind of visual fault detection approach and thus inherently depends heavily on subjective personal experience. A SDVAV w/Economizer Rules for Diagnostic Clarifier was proposed in the thesis for fault diagnosis.

2.3 Application of Shewhart Chart in the HVAC Fault Detection Field

Statistical process control (SPC) is the application of statistical methods to the monitoring and control of a process to ensure that it operates at its full potential to produce conforming products. SPC techniques can detect the shifts in process variables. The targeted variable(s) is compared to the upper and lower limits that define allowable operation. If the variable(s) falls outside these limits, the process is said to be out of the normal state and a fault is presumed to have occurred. The Shewhart chart is a fundamental SPC chart and has been shown to be effective in HVAC FDD applications. Traditionally, SPC charts and applications are based on the underlying assumption that the data being plotted are independent and normally distributed (Montgomery et al.1985).

A Shewhart chart (Montgomery et al.1985) is a two-dimensional plot of the evolution of the process over time. The horizontal dimension represents time, with samples displayed in chronological order, such that the earliest sample taken appears on the left and each newly acquired sample is plotted to the right. The vertical dimension represents the value of the sample. The threshold is usually chosen as three standard deviations of the sample values in normal operation (three-sigma limit). If a point is observed outside the $\pm 3\sigma$ control limit, we conclude that the process is out of control. Figure 1 is an example of a Shewhart chart. The sample mean is 10.05. Three-sigma is equal to 0.84. The upper and lower limits are 10.86 and 9.24 respectively.

The reason to choose three-sigma as the control limit is that for a normal distribution, the probability of encountering a point outside $\pm 3\sigma$ from the mean is 0.3%.

This is a rare event. Similarly, other events can be identified to be equally rare and can be used as flags for a fault. As an example, for a normal distribution, the probability of observing four points out of five in a row larger than ±1 σ from the mean is also about 0.3%. As a result, the rule "four of five consecutive points beyond the one standard deviation limit" is applied to detect a shift larger than one standard deviation. As shown in Figure 2, similar rules summarized as the Western Electric Company Rules (NIST/SEMATECH 2003) are developed to detect shifts of less than three-sigma.

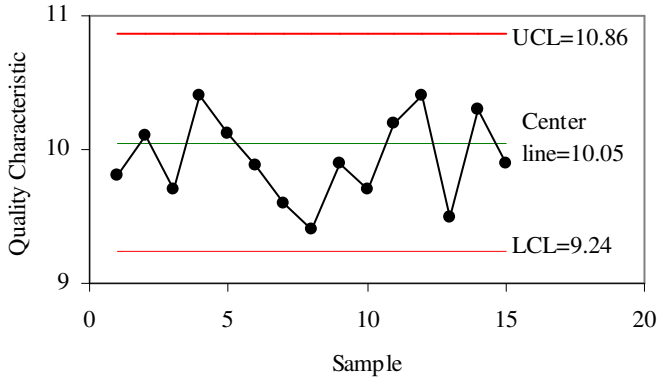

Figure 1 An example Shewhart chart

Haberl and Vajda (1988) and later Haberl (1992) applied Shewhart charts in analyzing both the cumulative energy and cost of the difference between measured and predicted heating or cooling consumption. Fasolo and Seborg (1994) used Shewhart charts to detect faults for AHUs. Since the normal distribution assumption may not be satisfied in the HVAC field, the paper didn't choose the three-sigma limits to be the fault detection criterion. Instead, two other criteria were used to determine if there was a fault:

chart limits were exceeded more than two times per every 1000 samples, or the data exhibited a marked lacked of variability.

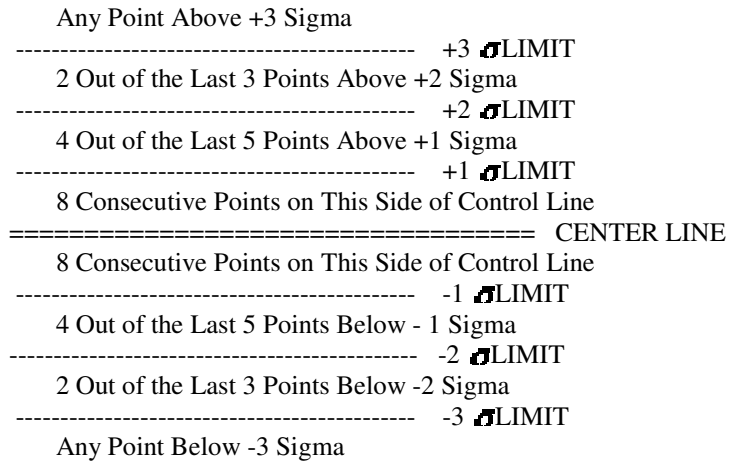

Figure 2 Western company rules for Shewhart chart

2.4 Similarity Measures and Their Application in FDD

Use of similarity measures is a popular technique in pattern recognition. It is widely applied in information retrieval to judge whether a document matches a query, or to measure the similarity of two documents. According to McGill et al. (1979), there are more than 60 different similarity measures. Among them, the most popular are cosine similarity and Euclidean distance similarity.

2.4.1 Cosine Similarity

Cosine similarity is a fundamental angle-based measure of similarity between two vectors of n dimensions using the cosine of the angle between them. It measures the similarity between two vectors based only on the direction, ignoring the impact of the distance between them. Given two vectors of attributes, $X = (x_1, x_2, \ldots, x_n)$ and $Y = (y_1,$

$y_2, \ldots, y_n$), the cosine similarity, cosθ, is represented using a dot product and magnitude as

$$\cos\theta = \frac{X \cdot Y}{\|X\|\|Y\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{(\sum_{i=1}^{n} x_i^2)}\sqrt{(\sum_{i=1}^{n} y_i^2)}} \quad (1)$$

The resulting similarity ranges from -1 meaning exactly opposite in direction, to 1 meaning exactly the same, with 0 indicating independence, and intermediate values indicating intermediate similarity or dissimilarity (Candan and Sapino 2010).

2.4.2 Euclidean Distance Similarity

Euclidean distance similarity is a common distance-based measure of similarity between two vectors of n dimensions using the distance between the vectors (Candan and Sapino 2010). The distance-based similarity measure considers only the impact of the distance between vectors regardless of the direction of the vectors. Given two vectors of attributes, $X = (x_1, x_2, \ldots, x_n)$ and $Y = (y_1, y_2, \ldots, y_n)$, the Euclidean Distance d from vector X to Vector Y is

$$d(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (2)$$

Shepard (1987) proposed as a universal law that distance and perceived similarity are related via an exponential function

$$s(X, Y) = e^{-d(X,Y)} \quad (3)$$

The resulting similarity ranges from 0 to 1 with 1 meaning the two vectors are identical.

2.4.3 The Application of Similarity Measure in FDD

Similarity measures show the effectiveness in FDD in many industries (Yoon and MacGregor 2001, Li and Dai 2005, Huang et al. 2007, Kabir 2009, Lee et al. 2009). Following are some references related to cosine similarity or Euclidean distance similarity.

Yoon and MacGregor (2001) introduced use of cosine similarity for fault isolation in the chemical industry. Joint plots of the cosine similarities in the residual and model spaces are applied to illustrate the diagnosis result. The cosines of the angle between the new fault vectors and the existing fault vectors in the library go to the (+1, +1) corner of the plot indicating essentially perfect co-linearity with the existing fault in both spaces. Li and Dai (2005) applied cosine similarity to diagnose oil-immersed transformer faults. A table records the cosine similarities between a fault symptom set and fault model sets. The test results showed the capability of the proposed approach. Lee et al. (2009) applied Euclidean distance similarity in the fault diagnosis of a turbine. Euclidean distance measures the similarity between interval-valued fuzzy sets. Kabir (2009) discussed the cosine similarity for fault diagnosis in an automotive infotainment application. The cosine similarities are presented on a bar graph in which each bar shows the similarity between investigated fault signature and one known fault signature in the fault database.

In the HVAC field, similarity measures have not yet been used in fault diagnosis in the published literature. Two fault detection studies combining similarity measures to determine days with similar energy consumption or similar AHU operation condition

have been found. Seem (2005) described a pattern recognition algorithm using Euclidean distance similarity for determining days of the week with similar energy consumption profiles. Li (2009) used cosine similarity to locate historical data that have similar operation conditions as the investigated data.

2.5 Summary

Fault detection and diagnosis is able to save energy, reduce maintenance costs, increase equipment life, and improve building control and occupant comfort. The above discussion reviews the general HVAC FDD methods and presents representative studies about whole building FDD. Physical redundancy, expert system rules, statistical bands and user selected thresholds are commonly used to detect faults. Fuzzy logic scoring systems, ANN classifiers, rule-based classifiers, and other classifiers are usually used to diagnose faults.

Computer models of whole buildings, systems and components form the basis of HVAC FDD for revealing the normal performance. Most developed whole building FDD systems are based on data-driven models, except ABCAT (Curtin 2007) which is based on a first-principles model. Building energy use is influenced by many factors such as weather, schedule, cooling and heating load, and occupancy. The first-principles simulation model in ABCAT, which includes the weather data, building electricity consumption and operational set-points as simulation inputs, can reasonably predict building cooling and heating consumption under different weather and operation conditions. Therefore, the FDD approaches to be undertaken in this research will use the first-principles simulation model in ABCAT to predict the normal building energy use.

The “Cumulative Cost Difference” plot developed in Curtin (2007) for ABCAT had a clear dependence on the user’s expertise. It lacked the ability to quantitatively define an alarm level. A general fault diagnosis scheme to provide possible fault reasons is seldom mentioned in the studies reviewed. This suggests a need to develop new whole building FDD techniques. Statistical process control is available for detecting changes in process variables and is a good candidate for identifying abnormal energy consumption. The most common SPC approach is the Shewhart chart used with Western Electric Company (WECO) Rules. Similarity measures are widely used in pattern matching. They quantitatively represent the degree of compliance within vectors. Cosine similarity and Euclidean distance similarity are two fundamental similarity measures. The use of similarity measures will be new in the HVAC fault diagnosis field.

3. OVERVIEW OF THE METHODOLOGY

This section provides an overview of the methodology developed in this dissertation for whole building fault detection and diagnosis. It begins by reviewing the methodology of ABCAT, follows by the detailed introduction of the proposed whole building fault detection and diagnosis approaches. In the next several sections, the proposed approaches are examined with synthetic simulation data and real field data.

3.1 Automated Building Commissioning Analysis Tool (ABCAT)

ABCAT is a tool that combines a calibrated simulation with detection and diagnostic techniques at the whole-building level. The simulation model in ABCAT is used to predict the building energy performance in the proposed FDD approaches. The methodology of ABCAT is briefly described below.

Figure 3 is a process flow diagram which visually describes the main steps in the ABCAT methodology. First, a building energy simulation model using the American Society of Heating, Refrigerating and Air- Conditioning Engineers (ASHRAE) Simplified Energy Analysis Procedure (SEAP) (Knebel 1983) is established and calibrated based on the building cooling and heating consumption in the baseline period. This period is chosen from a post-commissioning time period when the building's operation is considered to be optimal. The SEAP model is a first-principles model which effectively addresses the physical principles governing the building system behavior. The SEAP model has been found to be very useful for commissioning applications in existing buildings because it is simple enough to be easily implemented, and incorporates all of the major system and plant components needed to predict the results

of commissioning measures (Liu et al., 2003, 2004) with accuracy comparable to that of the more complex and newer simulation programs. The baseline period used is typically immediately after completion of commissioning, so it is assumed there are no faults during the baseline period. For the simulation model construction, the baseline period should be at least one month long and exhibit a wide range of ambient temperature and humidity fluctuations. Second, subsequent cooling and heating consumption is predicted by the model using future weather data and building electricity consumption. Third, both the simulated and measured consumption are passed to the data analysis routine that generates building performance plots, compares and performs calculations on the simulated and measured consumption data, applies fault detection methods, and reports diagnostic and energy consumption statistics. Finally, the user of the tool evaluates the data presented and determines whether or not there is a fault that requires action. If a fault is identified, the user or other experts can use the diagnostic information provided by ABCAT to help identify and correct the fault, and follow-up observations should see a return to expected performance. More details about ABCAT are presented in Lee et al. (2007) and the thesis of Curtin (2007).

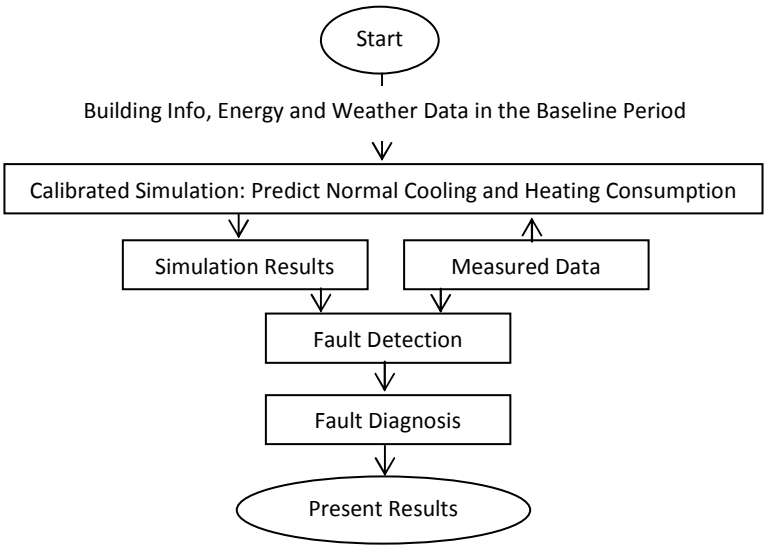

Figure 3 ABCAT process flow diagram

3.2 Fault Detection

3.2.1 Background

Small faults are difficult to detect in daily building energy consumption because they may increase consumption by an amount that is small compared to variation caused by weather and changing building activities. But they can have a significant impact when allowed to continue for a period of weeks, months or sometimes years. The proposed fault detection approach is intended to detect persisting small increases or decreases in the normal energy consumption. In this dissertation, normal consumption deviations are defined as one standard deviation (SD) of the residuals between measured and simulated consumption in a baseline operational period. This deviation level is called the "$SD_{\_baseline}$".

The Shewhart chart with the WECO rule "four of five consecutive points beyond the one standard deviation limit" (NIST/SEMATECH 2003) can be applied to detect shifts larger than one standard deviation. In this dissertation, the observed variables are the deviations between daily measured and simulated cooling or heating consumption. The mean of the deviations in the baseline period is adjusted to zero in the calibrated simulation, so the shift from the mean is equal to the deviation. The standard deviation in the limit is $SD_{\_baseline}$.

It is noted that the underlying assumption of the limits of the Shewhart chart is that the variable plotted is independent and normally distributed. However, in the HVAC field, this assumption is generally not valid. For example, the frequency distribution in the histogram of the HW consumption deviations in the baseline period for a simulated test building in Section 4.1 is left-skewed (Figure 4). Therefore, the application of the Shewhart chart with traditional control limits in abnormal energy consumption detection may lead to high false alarm rates. If we assume there are no faults during the baseline period, the time limit in the fault detection criterion should at least meet the condition that no faults were identified in the baseline period. In the investigation of 11 buildings' baseline periods, the "four out of five consecutive days" rule has detected 60 abnormal energy consumption faults. The investigation results are in conflict with the fault-free baseline assumption. Therefore, the "four of five consecutive days" rule is not a good choice for the fault detection criterion. The time limit needs to be redefined according to practical conditions.

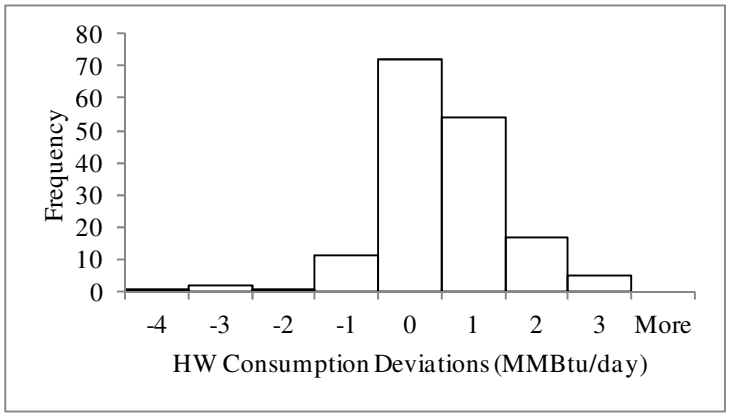

Figure 4 Histogram of the HW consumption deviations in the baseline period for a sample building

The proposed fault detection criterion named the Days Exceeding Threshold (DET) method may be stated as:

An abnormal energy consumption fault is identified if the deviation between the measured and simulated consumption is greater than one standard deviation of the residuals between measured and simulated consumption in the baseline period and persists for at least a user-defined number of days.

Two fault detection criteria, a Days Exceeding Threshold-Date (DET-Date) and a Days Exceeding Threshold-Outside Air Temperature (DET-Toa) are described in this dissertation based on the above idea. Residuals (differences between actual and modeled values) are used to investigate the normal/abnormal system behaviors. Both DET-Date and DET-Toa methods define a fault as identified when the absolute deviation between the measured and simulated consumption greater than $SD_{\_baseline}$ lasts for a user-defined number of "consecutive" days. The difference is that in the DET-Date method the fault-

flag days must be consecutive in time, while in the DET-Toa method the fault-flag days must be consecutive in outside air temperature.

The selection of the length of time limit is tricky. A longer time limit can decrease false alarm rates, but makes it easier to miss existing faults. A short time limit is creates excessive false alarms. In this dissertation, the standard that no fault can be identified in the 11 tested baseline periods is used to determine the number of user-defined days in the fault detection criteria. The next sections introduce the two fault detection methods and their time limit selection process.

3.2.2 Days Exceeding Threshold-Date Method

It is natural to conclude that there is a fault when the same fault symptom lasts for some number of consecutive days. The full description of the DET-Date method is:

An abnormal energy consumption fault is identified if the deviation between the measured and simulated consumption is greater than one standard deviation of the residuals between measured and simulated consumption in the baseline period and persists for at least a user-defined number of chronologically consecutive days.

Table 2 shows that in the investigation of the residuals for 11 buildings during their baseline periods, there are at most 11 days when the deviation between measured and simulated consumption is consecutively greater than $SD_{\_baseline}$. In this dissertation, 20 days is chosen as a conservative time limit for the DET-Date method to determine if the energy consumption is abnormal.

Table 2 The longest time in 11 baseline periods when the deviation between CHW/HW measured and simulated consumption is consecutively greater than SD_Baseline

Building	(Days)	
	CHW	HW
Eller	9	5
Harrington	7	4
Kleberg	6	4
Wehner	5	5
VMA	4	11
Coke	2	3
Bush Academic	4	6
Gilchrist	4	5
IBM	4	6
Sbisa	5	9
DASNY	3	2

Figure 5 illustrates the procedure of the DET-Date Method. First, we assign a "fault index" to each day's cooling and heating energy consumption in the investigated period. The fault index is defined as:

$$\text{Fault Index} = \begin{cases} 1 & \mathrm{E_{mea}} - \mathrm{E_{sim}} > SD_{baseline} \\ -1 & \mathrm{E_{mea}} - \mathrm{E_{sim}} < -SD_{baseline} \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

where E_{mea} is the daily measured cooling or heating energy consumption and E_{sim} is the daily fault-free cooling or heating energy consumption predicted by the calibrated simulation model in ABCAT. The expression means that when the daily consumption is at least one standard deviation above expected consumption, the fault index is 1; when the daily consumption is at least one standard deviation below expected consumption, the fault index is -1; for all other conditions, the fault index is defined to be 0. Next, draw the DET-Date plot. The date is plotted along the horizontal axis on the plot, and

every point on the plot represents the sum of the fault indexes for the next 20 days (including the day on which the point is plotted). Thus a point at 20 or -20 means the measured consumption for each of the next 20 days is more than one standard deviation above or below the simulated consumption. A fault period appears as one or more points at ±20 on the plot.

Figure 6 is an example of a DET-Date Plot.

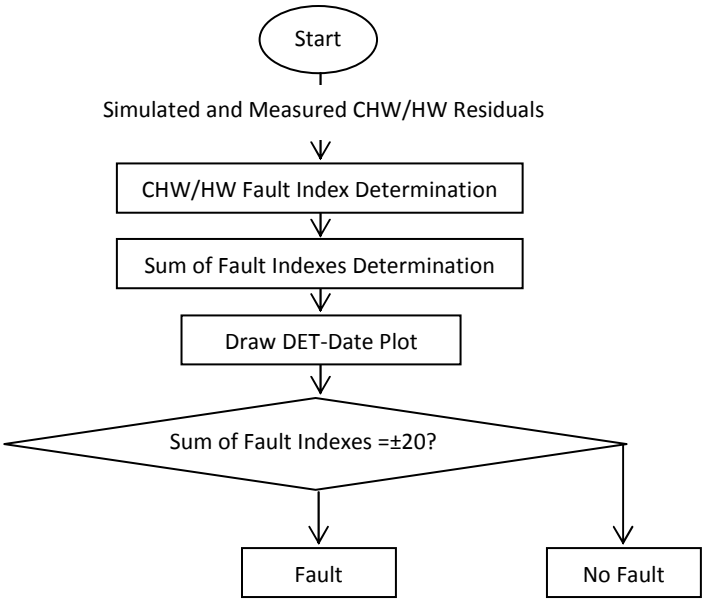

Figure 5 Days Exceeding Threshold-Date method procedure diagram

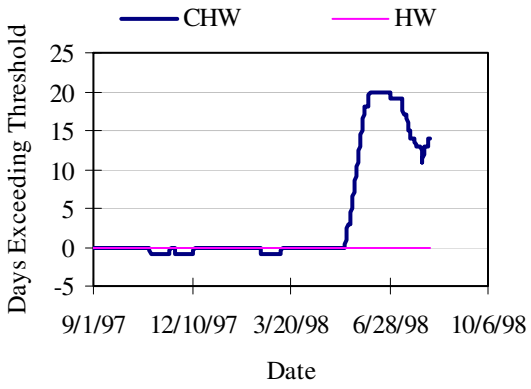

Figure 6 An example of Days Exceeding Threshold-Date plot

3.2.3 Days Exceeding Threshold-Outside Air Temperature Method

Building cooling and heating consumption variations due to HVAC operation changes are temperature-dependent. Claridge et al. (2001) addressed use of a "characteristic signatures" methodology for the rapid calibration of cooling and heating energy consumption simulations for commercial buildings. A characteristic signature is a normalized plot of cooling and heating consumption deviation caused by a specified change in HVAC system operation as a function of outdoor air temperature. The document contains the characteristic signatures of four different system types: single-duct variable-air-volume (SDVAV), single-duct constant-air-volume (SDCV), dual-duct variable-air-volume (DDVAV) and dual-duct constant-air-volume (DDCAV). For each operational change in a specified system, the characteristic signature has a characteristic shape as a function of outside air temperature. The cooling or heating consumption deviation due to a fault has similar behavior as a function of outside air temperature. As an example, Figure 7 shows the cooling coil leaving air temperature characteristic

signatures for CHW and HW of a SDVAV system. The curve on the left shows the change in cooling energy use and the curve on the right the change in heating energy use, when the temperature at which air leaves the cooling coil was decreased from 55 to 54°F. The change is expressed on the Y axis as percent of the maximum baseline cooling or heating consumption, respectively for a cooling or heating characteristic signature. The cooling or heating consumption deviations have similar values when the outside air temperature is close.

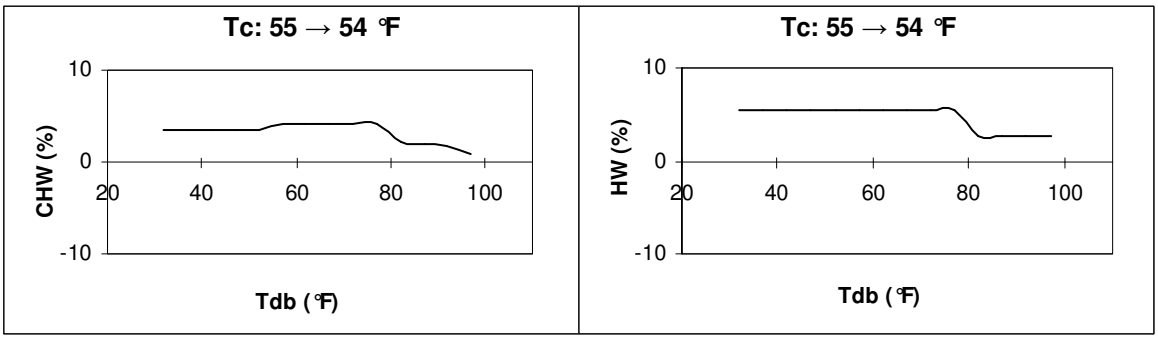

Figure 7 Cooling coil leaving air temperature characteristic signatures

The results of ABCAT simulation also show the temperature dependence. Figure 8 illustrates the outside airflow characteristic signatures for CHW and HW of a DDVAV system. The outside airflow ratio (outside airflow volume/maximum designed airflow volume of the system) increases 5%. The plot on the left shows that the CHW characteristic signature shows a consumption decrease of about 3% from 40-60°F and gradually consumption gradually increases to 8% over the range of 60-90°F. The plot on

the right shows that the HW characteristic signature gradually decreases from 5% to 0% from 40-60°F and stays at 0% from 60-90°F.

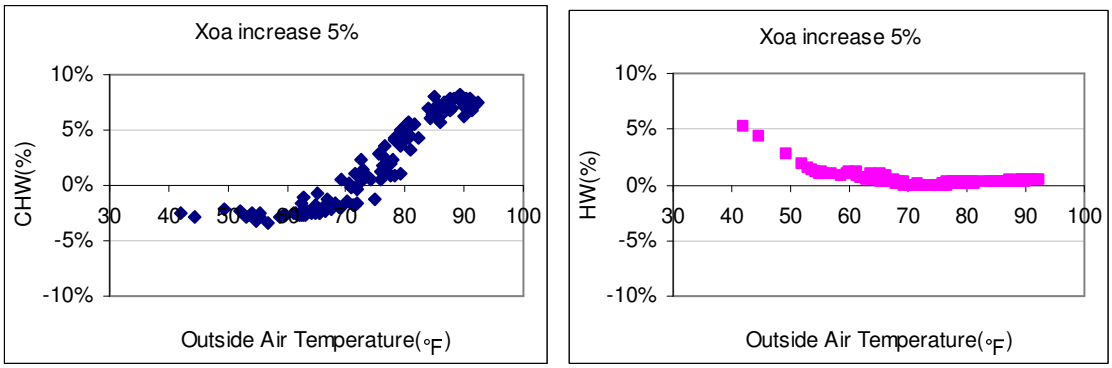

Figure 8 Outside airflow ratio (X_{oa}) characteristic signatures

Since building cooling and heating consumption variations due to HVAC operational changes are temperature-dependent, it would be beneficial if we take temperature dependence into account when we set up the fault detection metric. The DET-Toa method is developed based on this concept. The whole building fault detection criterion may be stated as:

An abnormal energy consumption fault is identified if the deviation between the measured and simulated consumption is greater than one standard deviation of the residuals between measured and simulated consumption in the baseline period and persists for at least a user-defined number of days which are consecutive when ordered according to increasing or decreasing outside air temperature.

Table 3 shows that 14-days is the shortest time period for which the DET-Toa method doesn't detect a fault in the 11 buildings during their baseline periods. As a result, similar to the DET-Date method, 20 days was chosen as a conservative time limit for the DET-Toa method to determine if the energy consumption is abnormal.

Figure 9 is a flow chart showing use of the DET–Toa method to detect abnormal cooling or heating consumption in buildings. The procedure includes three major steps.

The first step is to create the fault detection matrix. The main function of this step is to import the data for the investigated period into the program. The building energy consumption and weather data are assigned to a matrix named the fault detection matrix. The structure of the fault detection matrix is [Date, Toa, Fault Index]. The first column contains the date and the second column contains the corresponding daily average outside air temperature. The fault index placed in the third column is the CHW fault index for that day when investigating abnormal cooling consumption faults, and is the HW fault index for that day when investigating abnormal heating consumption faults. The rows of the fault detection matrix are arranged in ascending temporal order and each row represents a day. The CHW and HW fault indexes are derived according to expression (4). The program chooses the latest year of data if the investigated period is more than one year.

Table 3 The shortest time limit for which the DET-Toa method does not detect fault in the 11 buildings' baseline periods

Building	(Days)	
	CHW	HW
Eller	10	11
Harrington	10	11
Kleberg	8	8
Wehner	7	14
VMA	7	14
Coke	7	9
Bush Academic	7	9
Gilchrist	7	6
IBM	7	7
Sbisa	7	11
DASNY	6	5

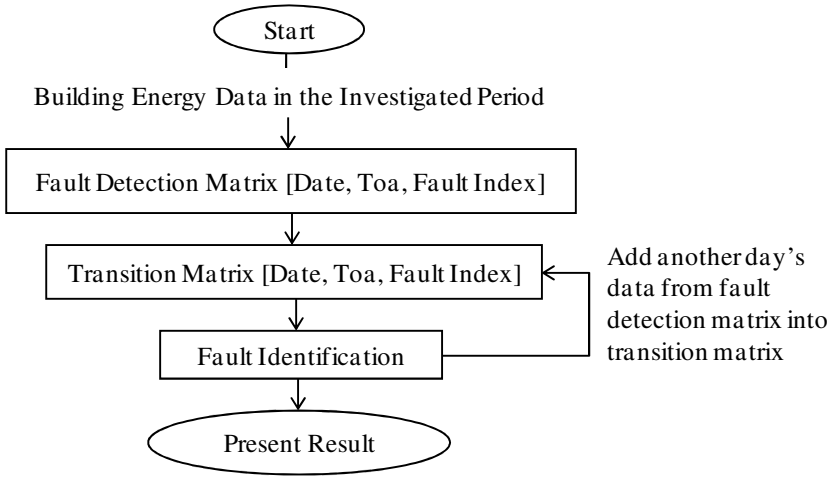

Figure 9 Flow chart of DET-Toa method

The second step is creation of the transition matrix. The intent of this step is to determine the date when a fault is identified. The structure of the transition matrix is the

same as the fault detection matrix: [Date, Toa, Fault index], and the rows are also arranged in ascending temporal order. The program first imports the data of the initial 20 days (e.g. December 1-20) in the fault detection matrix to the transition matrix. Then it runs the third step to see if a fault can be identified with the initial 20-days of data. If the answer is no, then add the next day's data in the fault detection matrix into the transition matrix and run the third step again to see if a fault can be identified with the 21-days (e.g. December 1- 21) of data. The transition matrix expands by one more day with each iteration until a fault is identified or the end of the fault detection matrix is reached. The last day entered in the transition matrix before a fault is identified is called the fault identification day.

The last step is fault identification. The flow chart of this step is plotted in Figure 10. In this step, we examine the fault index in the transition matrix day by day. If the fault index of day i is 0, the program checks the fault index of the next day. If the fault index of day i is 1 or -1, we will assign the rows in the transition matrix whose date is later than or equal to the date of day i to a sub transition matrix. Then we sort the data in the sub transition matrix by outside air temperature in ascending order. The top and bottom rows in the sorted sub transition matrix are the days with the lowest and highest outside air temperatures, respectively. The sorted sub transition matrix has four columns: [Date, Toa, Fault Index, the Sum of Fault Indexes]. The sum of fault indexes value placed in each row is calculated by adding the fault indexes in the next 20 rows (including the current row) in the sorted sub transition matrix. A fault is identified when the program finds 20 or -20 in the sum of fault indexes column.

The flow chart of the DET-Toa method is more complex than that of the DET–Date method, because the following problems must be solved by the program.

1. Determination of the fault identification day: The fault identification day is the earliest day on which the abnormal consumption fault can be identified. Answering this question in the DET-Date method is easy. We only need to locate the latest date in the 20 days which contribute the first +/-20 on the DET-Date plot. In the DET-Toa method, the process is a little more complicated. The investigated period may contain months of data, e.g. January to November. But it is possible that the program can already identify the fault when the data is only available from January to June. As a result, a transition matrix is defined in the program. It initially contains the first 20 days in the investigated period (the initial number of days should equal the time limit in the criterion) and expands by one more day with each iteration. Then, the program checks if a fault can be identified with the available data in the transition matrix. The fault identification day is the last day in the transition matrix when a fault is identified.

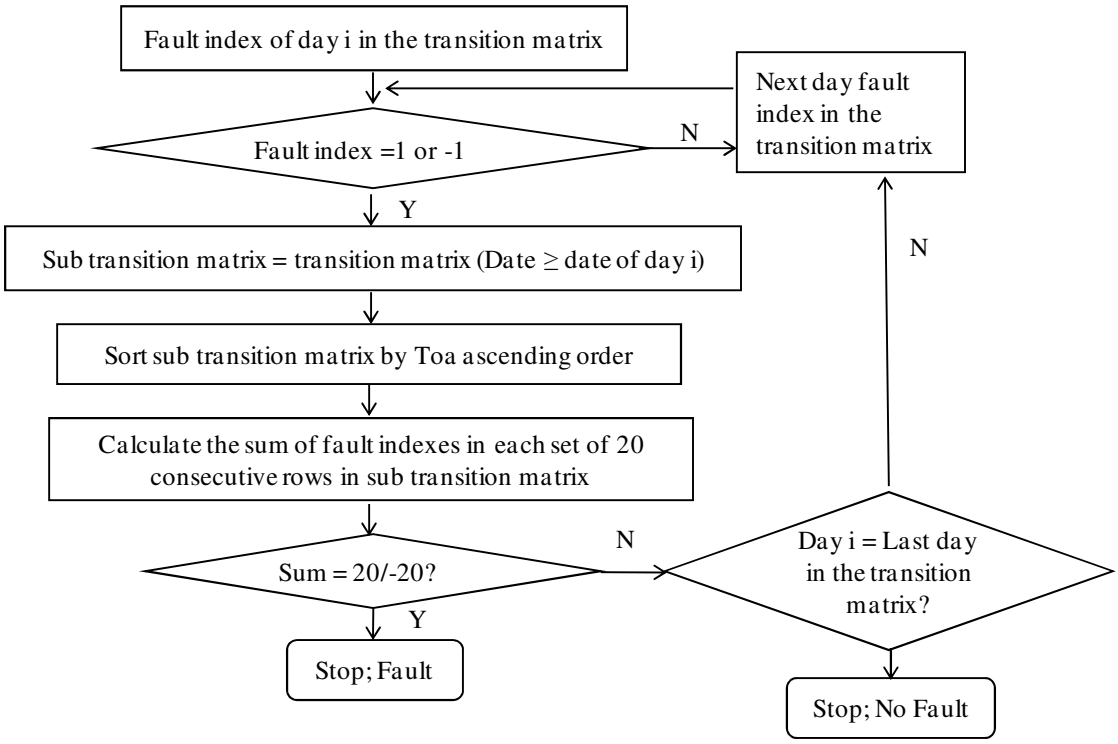

Figure 10 Flow chart of the "fault identification" step in the procedure

2. The impact of the fault index in the historical fault-free period: The fault index in the historical fault-free period needs special treatment; otherwise, it would influence the fault detection results. Assume the transition matrix contains the data in May, and a fault occurred on May 2 which increases CHW consumption at least one CHW $SD_{\_baseline}$ whenever outside air temperature is higher than 73°F. On May 1, the outside air temperature is 74°F, and the fault index is 0, since there was no fault at that time. In the remaining days when the outside air temperature is above 73°F, the fault indexes are all 1. If we directly sort the transition matrix by outside air temperature in ascending order and calculate the sum of indexes after that, we may find none of the sum of fault indexes reaches 20, because the zero fault index of May 1 is in the middle of the other 30 fault

indexes of 1. As a result, the sub transition matrix is defined to avoid such problems. It removes any row whose date is earlier than day i before the matrix is sorted and ordered according to the outside air temperature.

3.2.4 Fault Evaluation-Modified Z-Score

Fault evaluation assesses the impact of a fault on overall system performance. To let the user know the severity of a detected fault, a modified z-score (Hoaglin 1993) is used to quantify the direction and distance of the consumption from the normal value. The modified-z-score is a measure of how far the sum of the $SD_{\_baseline}$ residuals in the 20 fault-flag days is from their normal 20-day value. In the DET-Date method, the 20 fault-flag days are the 20 days contributing to the first ±20 on the DET-Date plot. In the DET-Toa method, the 20 fault-flag days are the 20 days which are consecutive in the order of outside air temperature and during which the deviation is consecutively larger than $SD_{\_baseline}$. In equation form, we determine the modified z-score with

$$z_m = \frac{\sum_{i=1}^{20}(Residual_i - \overline{Residual}_{baseline})}{20SD\_Baseline} \quad (5)$$

The residual is the difference between actual and modeled values. $\overline{Residual}_{baseline}$ is the mean of the residuals in the baseline period and is always adjusted to zero in the calibrated simulation. A modified z-score of two means the measured consumption averages two $SD_{\_baselines}$ higher than the simulated consumption in the 20 fault-flag days.

3.3 Fault Diagnosis

Fault diagnosis on the whole building level has been somewhat neglected in previous HVAC FDD studies. The level of diagnostics proposed emphasizes limiting the

possible causes to several options and ranking the options according to their probability. It will not attempt to "find a needle in a haystack", but instead will attempt to effectively reduce the size of the haystack in which the operator must look.

Figure 11 shows the major steps required to diagnose possible causes of abnormal cooling and heating consumption in buildings using similarity measures. The method is referred to as the Cosine Similarity method if cosine similarity is adopted and is referred to as the Euclidean Distance Similarity method if Euclidean similarity is implemented.

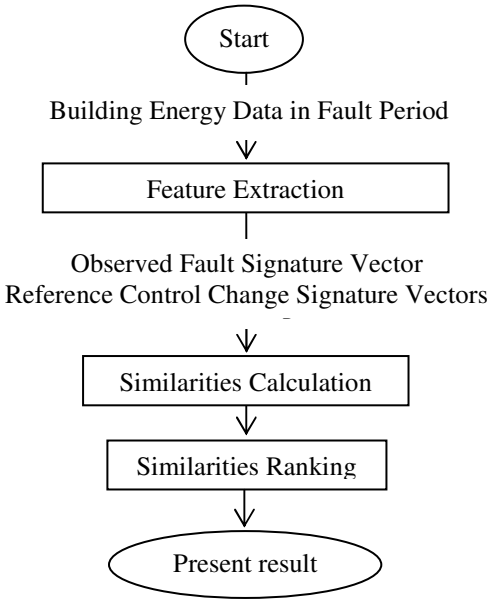

Figure 11 Block diagram for whole building fault diagnosis with similarity measures

Step 1: Reference Control Change Library Determination

The building HVAC system considered here is composed of AHUs and terminal boxes. Various AHU and terminal box faults are described in the literature. Technically,

all AHU and terminal box components including control devices could develop faults. IEA Annex 25 studied the typical faults in VAV systems (IEA 1996). According to the study, faults related to air dampers, coil valves and fan speed are the most important faults causing energy loss.

Whole building fault diagnosis is different from component level fault diagnosis. It can only give a general clue, for example, that there is excess outside air flow in the building, but can't tell which specific component, e.g. the fully closed outside air damper of AHU 3-1, is causing the problem. A technician still needs to investigate in the field to determine and correct the specific cause. The reference control change library collecting known whole building level faults is pre-determined initially. Liu et al. (2003) presents seven typical whole building level faults: terminal box reheat valve leakage, improper minimum terminal box airflow, improper minimum outside airflow, poor outside air damper quality, excessive maximum supply airflow (constant air volume system), improper supply air static pressure, and improper building positive pressure. Table 4 gives 10 whole building level fault examples. Each fault listed in Table 4 will be called a reference control change in the subsequent discussion.

The signatures of the reference control changes are used as the reference symptoms in fault diagnosis. The energy pattern may be different when the control change severity is different, so the number of levels of severity for a control change will be defined in advance in the library.

Table 4 Whole building level fault examples

Outside air flow volume increase/decrease
Preheat/precool temperature increase/decrease
Preheat/precool coil valve leakage
Cooling coil (SD)/cold deck (DD) leaving air temperature increase/decrease
Hot deck (DD) leaving air temperature increase/decrease
Heating coil valve leakage (DD)
Minimum airflow volume increase/decrease
Maximum airflow volume (CV) increase/decrease
Room set-point temperature increase/decrease
Terminal box reheat coil valve leakage (SD)
Terminal box damper leakage (DD)

Note: SD – Single Duct System; DD-Dual Duct System; CV – Constant Volume System

Step 2: Feature Extraction

The feature extraction block generates the observed fault signature vector and a number of reference control change signature vectors, each corresponding to a known control change in the reference control change library.

It is assumed some fault detection mechanism has already determined that an abnormal consumption fault is present and has persisted for a certain time. This period is referred to as the fault period. In this block, first, the calibrated simulation model in ABCAT is used to produce the fault-free cooling and heating consumption in the fault period. Second, the calibrated simulation model in ABCAT is used to predict the cooling and heating consumption when there is a known control change from the reference library persisting during the fault period. For a specified control change, a specific input parameter of the calibrated simulation model will be changed. Since there are several levels of severity for a control change, the corresponding input parameter will be

changed several times to simulate various fault sizes. Finally, the observed fault signature vector and reference control change signature vectors are generated using the following expression:

$$V = \begin{bmatrix} fs_{CHW} \\ fs_{HW} \end{bmatrix} \quad (6)$$

where $fs_{CHW,i} = \frac{CHW_{mea,i} - CHW_{sim,i}}{E_{AveBaseline}}$, $fs_{HW,i} = \frac{HW_{mea,i} - HW_{sim,i}}{E_{AveBaseline}}$ (Observed fault)

$fs_{CHW,i,j} = \frac{CHW_{ref\,C,i,j} - CHW_{sim,i}}{E_{AveBaseline}}$, $fs_{HW,i,j} = \frac{HW_{ref\,C,i,j} - HW_{sim,i}}{E_{AveBaseline}}$ (Reference control change)

A signature vector includes two parts: the CHW signature fs_{CHW} and the HW signature fs_{HW}. In this way, the similarity of both CHW and HW features can be considered. $CHW_{mea,i}$ and $HW_{mea,i}$ are the daily measured cooling and heating energy consumption values respectively on the i^{th} day of the fault period; $CHW_{sim,i}$ and $HW_{sim,i}$ are the daily fault-free cooling and heating energy consumption values respectively predicted by the calibrated simulation model on the i^{th} day of the fault period. $CHW_{ref\,C,i,j}$ and $HW_{ref\,C,i,j}$ are the daily cooling and heating energy consumption values respectively on the i^{th} day of the fault period when there is the j^{th} control change from the reference library persisting during the fault period. $E_{AveBaseline}$ is the average cooling plus heating energy consumption values in the baseline period. For a specified control change, a specific input parameter of the calibrated simulation model will be changed. As the energy pattern may be distinctive when the control change severity is different, it is necessary to take into account the change magnitude when altering the input parameter. As a result, when the reference control change signature vectors are generated, the number of severity levels for a control change must be defined in

advance. Then the corresponding input parameter will be changed several times to simulate various fault sizes. Assuming the fault severity has five levels, there would be five reference control change signature vectors for a single reference control change.

Step 3: Similarities Calculation

In this block, cosine similarity and Euclidean distance similarity between the observed fault signature vector and each of the reference control change signature vectors are calculated.

As mentioned in Section 2.4, given two vectors of attributes, $X = (x_1, x_2, \ldots, x_n)$ and $Y = (y_1, y_2, \ldots, y_n)$, the cosine similarity, $\cos\theta$, is represented using a dot product and magnitude as

$$\cos\theta = \frac{X \cdot Y}{\|X\|\|Y\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{(\sum_{i=1}^{n} x_i^2)}\ \sqrt{(\sum_{i=1}^{n} y_i^2)}} \quad (1)$$

The Euclidean Distance d from vector X to Vector Y and perceived similarity are

$$d(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (2)$$

$$S_d(X, Y) = e^{-d(X,Y)} \quad (3)$$

X in expressions (1-3) is the observed fault signature vector and Y is the reference control change signature vector. Substituting the expressions of observed and reference signatures, the expressions for cosine similarity $\cos\theta$ and Euclidean distance similarity S (X, Y) become

$$\cos\theta = \frac{\sum_{i=1}^{n}[(CHW_{mea}-CHW_{sim})_i(CHW_{ref\,C}-CHW_{sim})_i+(HW_{mea}-CHW_{sim})_i(HW_{ref\,C}-HW_{sim})_i]}{\sqrt{[\sum_{i=1}^{n}[(CHW_{mea}-CHW_{sim})_i^2+(HW_{mea}-HW_{sim})_i^2]}\sqrt{[\sum_{i=1}^{n}[(CHW_{ref\,C}-CHW_{sim})_i^2+(HW_{ref\,C}-HW_{sim})_i^2]}} \quad (7)$$

$$S_d(X,Y) = e^{-\sqrt{\sum_{i=1}^{n}(\frac{CHW_{mea,i}-CHW_{ref\,C,i}}{E_{AveBaseline}})^2+\sum_{i=1}^{n}(\frac{HW_{mea,i}-HW_{ref\,C,i}}{E_{AveBaseline}})^2}} \quad (8)$$

where i is the i^{th} day in the fault period and n is the number of days in the fault period. If the reference control change doesn't cause any energy shift over the fault period, $CHW_{ref\,C}$ and $HW_{ref\,C}$ would be the same as CHW_{sim} and HW_{sim} respectively. In this context, the cosine similarity is defined as zero.

Step 4: Similarities Ranking

The similarities ranking block sorts different types of reference control changes by the similarity in descending order. As mentioned above, a reference control change would have more than one fault signature vector. Thus, two steps will be taken when ranking the similarities. First, choose the largest cosine similarity/Euclidean distance similarity from the cases with the same reference control change to be representative of that control change. Next, compare the representative similarities of all the reference control changes and sort them by descending order to create a rank-ordered list of control changes.

Step 5: Present Results

The rank-ordered list of control changes basically ranks the probability that the reference control change is the cause of the observed fault. The similarity measures

compare the symptoms of the current fault against the symptoms of the reference control changes. A larger similarity corresponds to a higher probability that the known control change is the cause of the observed abnormal consumption.

4. ANALYSIS OF WHOLE BUILDING FAULT DETECTION APPROACHES WITH SIMULATION TEST

Generating data through simulation is much less costly than generating data in a field test. Various synthetic faults can be easily introduced into a simulation program. The results are also valuable for illustrating the success of the proposed whole building fault detection approaches – the Days Exceeding Threshold-Date (DET-Date) method and the Days Exceeding Threshold-Outside Air Temperature (DET-Toa) method. The validation process and results of the simulation tests with a SDVAV system and a DDVAV system are presented in this section. In addition, a comparison of the test results is performed to evaluate the effectiveness of the two methods. All the simulation data were produced by the simulation module of ABCAT (Curtin 2007). Simulation was performed for College Station weather data and the measured building electricity consumption.

4.1 Simulated for DDVAV System

4.1.1 Simulated Data Sets

The simulated test building is the Eller Oceanography and Meteorology (EOM) Building located on the Texas A&M University campus in College Station. It is primarily comprised of offices, classrooms and laboratories. The building has 14 floors with a basement and 180,000 ft^2 of conditioned space. Thermal energy is supplied to the building in the form of hot water for heating and chilled water for cooling from the central utility plant. The majority of the building is served by four dual-duct VAV AHUs. None of the AHUs have economizer capabilities. The commissioning work on

this building was finished in March 1997. The ABCAT simulations were calibrated to the baseline consumption period from March 19-August 31, 1997. The daily average temperature during the baseline period was between 47-88°F.The standard deviations of the residuals in the baseline period were 5.0 MMBtu/day and 3.3 MMBtu/day respectively for CHW and HW energy consumption. They are 7.7% and 5.1% of the average daily combined CHW and HW consumption during the baseline calibration period. The simulation period is September 1, 1997-September 30, 1998. The building usage was consistent during the period of study. The occupant density was assigned weekday and weekend values within the simulation. The measured daily building electricity use (none of which is directly used for heating or cooling) was assumed to be converted into heat gain within the building by the simulation.

Since there are many different faults related to the building HVAC systems, it is not time-effective to test the approach with all possible faults. Ten important faults causing significant energy changes were chosen in this test. It was assumed one of the following ten synthetic control changes that happened on October 1, 1997 and lasted to September 30, 1998:

1. Outside airflow ratio (X_{oa}) increase of 3.1%
2. Outside airflow ratio decrease of 3.1%
3. Cold deck leaving air temperature (T_{cl}) increase of 4°F
4. Cold deck leaving air temperature decrease of 4.5°F
5. Hot deck leaving air temperature (T_{hl}) increase of 10°F
6. Hot deck leaving air temperature decrease of 12°F

7. Minimum airflow ratio (X_{min}) increase of 16%
8. Minimum airflow ratio decrease of 20%
9. Room cooling set-point temperature (T_{rc}) increase of 1.8°F
10. Room cooling set-point temperature decrease of 1.6°F

Note: The denominator of the outside airflow ratio and minimum airflow ratio is the maximum design airflow volume in the system.

The energy consumption with no control change and under the ten synthetic control changes was produced by the simulation module of ABCAT (Curtin 2007). Simulation was performed for College Station weather data and the measured building electricity consumption during that period. The yearly average cooling plus heating consumption (October, 1997 – September, 1998) was 63.75MMBtu/day when there was no control change present. The ten synthetic control changes are referred to as the ten simulation test cases. The monthly average energy (cooling plus heating) use change index is defined as

$$\text{Monthly average energy use change index} = \frac{\text{Monthly average energy use change}}{\text{Yearly average energy use when there is no fault}} \quad (9)$$

The influence of the ten control changes on the monthly average energy use change index in the period from October, 1997 – September, 1998 is shown in Figure 4.1. The parameter changes were chosen so the maximum monthly average cooling and heating consumption deviation (Blank bars in Figure 12) caused by each synthetic control change was 10% of the yearly average cooling and heating consumption (63.75MMBtu/day) when there was no fault.

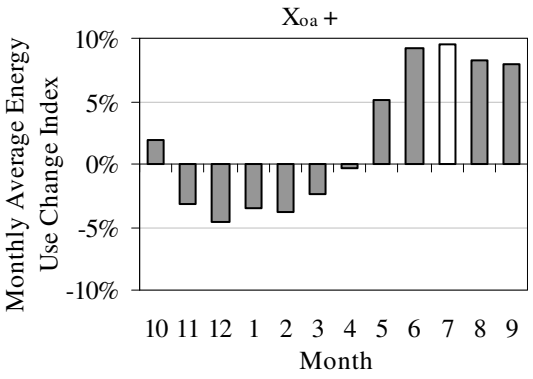

Figure 12.1

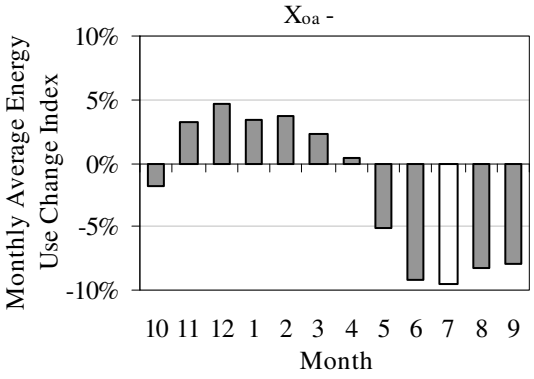

Figure 12.2

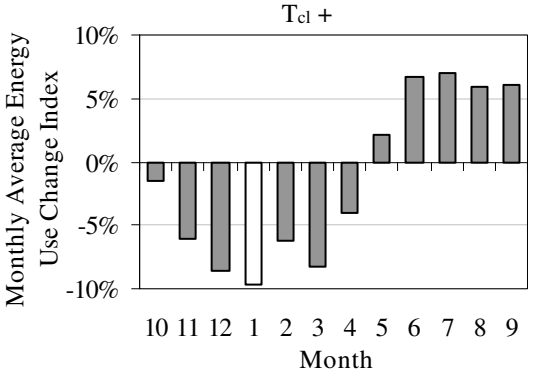

Figure 12.3

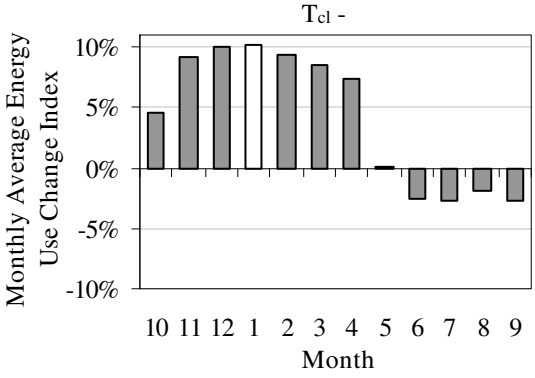

Figure 12.4

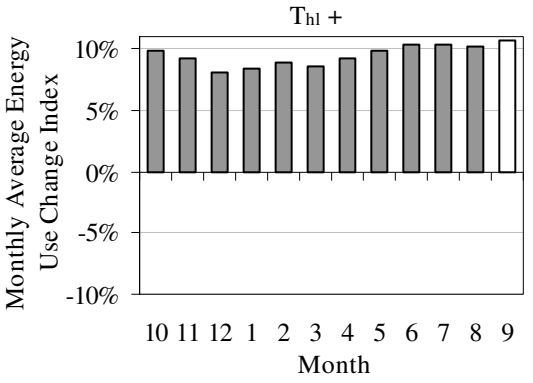

Figure 12.5

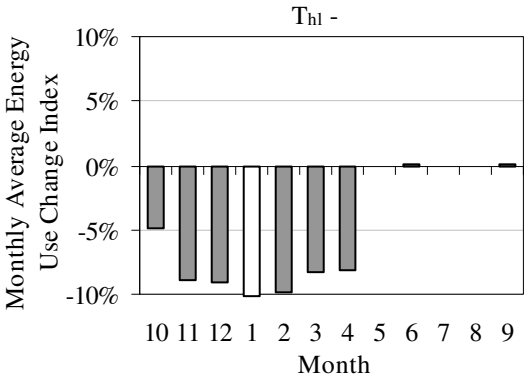

Figure 12.6

Figure 12 Monthly average energy (cooling plus heating) use change indexes under different synthetic control changes from September 1, 1997 to September 30, 1998 for the EOM Building. (12.1)) X_{oa} increase of 3.1%, (12.2) X_{oa} decrease of 3.1%, (12.3) T_{cl} increase of 4°F, (12.4) T_{cl} decrease of 4.5°F, (12.5) T_{hl} increase of 10°F, (12.6) T_{hl} decrease of 12.1°F, (12.7) X_{min} increase of 16%, (12.8) X_{min} decrease of 20%, (12.9) T_{rc} increase of 1.8°F, (12.10) T_{rc} decrease of 1.6°F

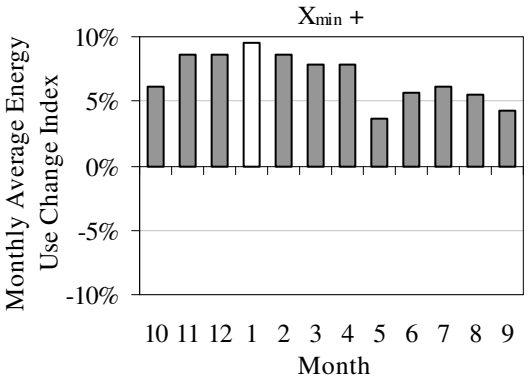

Figure 12.7

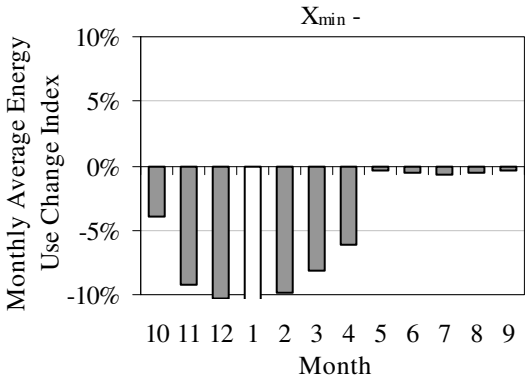

Figure 12.8

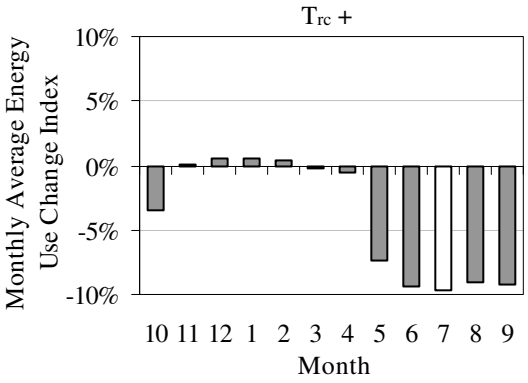

Figure 12.9

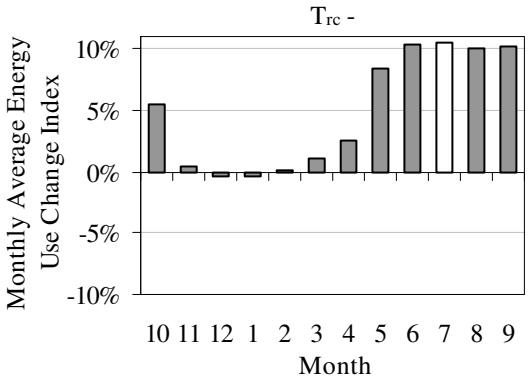

Figure 12.10

Figure 12 Continued

4.1.2 Fault Detection Results with Simulated Test Data - DET-Date Method

The simulated consumption under synthetic control changes were treated as the real measured data and the simulated consumption with no control change were treated as simulated data. The measured data were compared with the simulated data. Daily CHW and HW fault indexes were calculated according to expression (4) in Section 3 for each day of the simulation. DET-Date plots based on the sum of the 20 consecutive fault indexes in time series are shown in Figure 13. This is a new type of figure, so it is

explained in detail. The CHW line in Figure 13.1 begins with a value of zero on September 1, 1997. This value indicates that the CHW consumption is within one standard deviation from the expected value on all 20 of the days from September 1-20, 1997. The CHW line reaches +20 on June 8, 1998 indicating that CHW consumption is at least one standard deviation above expected consumption for June 8-June 27, 1998 and hence corresponds to the beginning of an identified CHW Fault. The CHW line remains at +20 until June 28, 1998, meaning that the CHW consumption remains at least one standard deviation above expected consumption every day through July 17, 1998, so the actual duration of the detected CHW fault extends 40 days past the first day when the CHW line is plotted as +20.

Figure 13.1 indicates that the DET-Date plot successfully identified the abnormal CHW energy consumption due to excess outside airflow. Likewise, the DET-Date plot has detected the abnormal energy consumption caused by five of the other synthetic control changes (Figure 13.2, Figure 13.4, and Figures 13.8-10) but has failed to identify four of the other synthetic control changes.

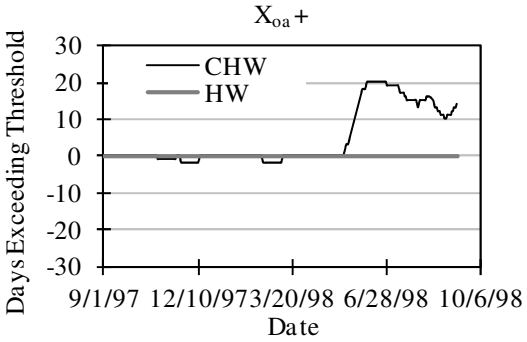

Figure 13.1

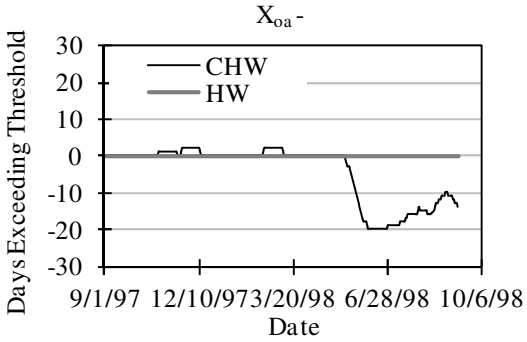

Figure 13.2

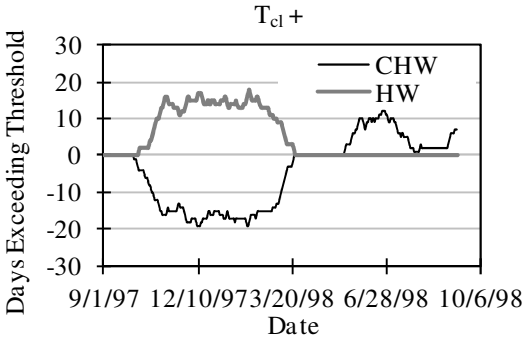

Figure 13.3

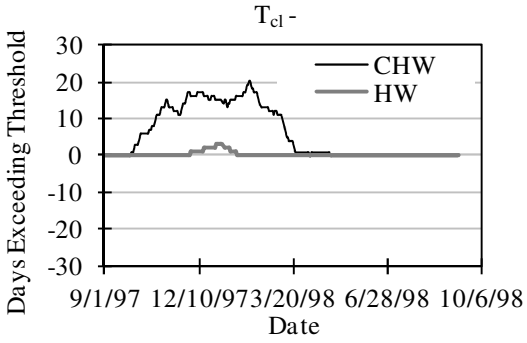

Figure 13.4

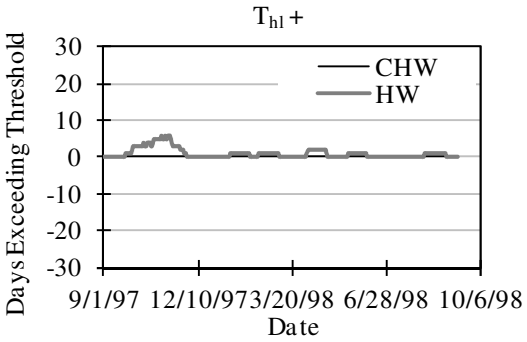

Figure 13.5

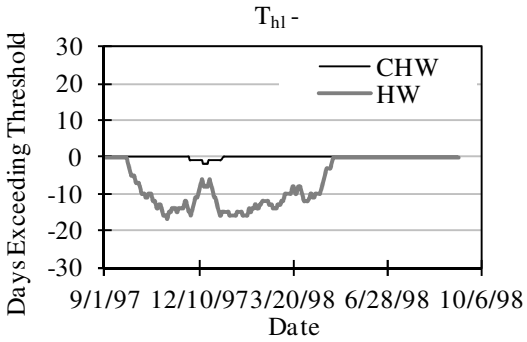

Figure 13.6

Figure 13 Days Exceeding Threshold-Date plots under different synthetic control changes from September 1, 1997 to September 30, 1998 for the EOM Building. (13.1) X_{oa} increase of 3.1%, (13.2) X_{oa} decrease of 3.1%, (13.3) T_{cl} increase of 4°F, (13.4) T_{cl} decrease of 4.5°F, (13.5) T_{hl} increase of 10°F, (13.6) T_{hl} decrease of 12°F, (13.7) X_{min} increase of 16%, (13.8) X_{min} decrease of 20%, (13.9) T_{rc} increase of 1.8°F, (13.10) T_{rc} decrease of 1.6°F

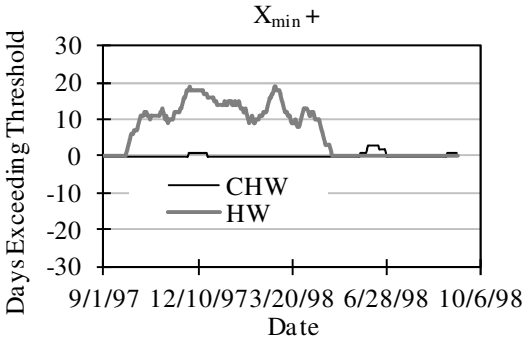

Figure 13.7

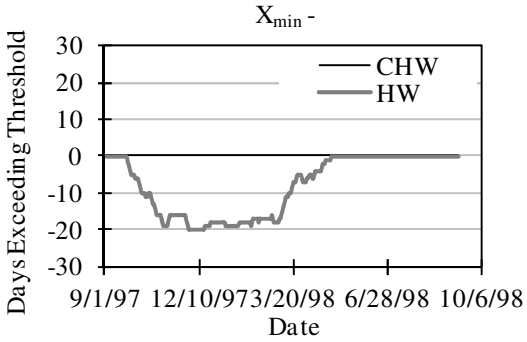

Figure 13.8

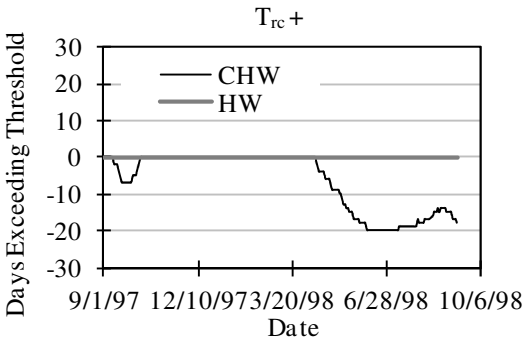

Figure 13.9

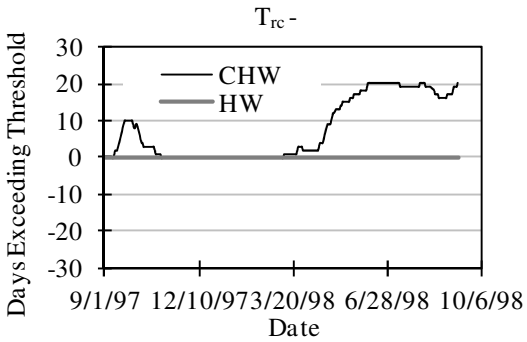

Figure 13.10

Figure 13 Continued

In the lower room cooling set-point temperature case (ID 10), the DET-Date plot detected three abnormal CHW consumption periods (Figure 13.10). The end of the first period is July 31, 1998 and the beginning of the second period is August 2, 1998. This is because the CHW consumption is within one standard deviation of the expected value on August 1, 1998. The end of the second period is August 29, 1998 and the beginning of the third period is September 11, 1998. There are four days in the period from August 29-September 11, 1998 in which the CHW consumption is within one standard deviation of the expected value. The fault detection criterion needs the deviation above one

$SD_{\_baseline}$ to be continuous for at least 20 days; thus the abnormal CHW consumption appears as the three disconnected periods in Figure 13.10.

The fault detection results with the DET-Date plot in the ten simulation cases are summarized in Table 4.1. The DET-Date plot detected abnormally high and low CHW consumption due to the excess outside airflow and the deficient outside airflow respectively in June 1998; detected abnormally high CHW consumption due to the lower cold deck leaving air temperature in February 1998; detected abnormally low HW consumption due to the deficient minimum airflow in November 1997; and detected abnormally low and high CHW consumption due to the lower and higher room cooling set-point temperature respectively in June 1998. In summary, the DET-Date plot successfully identified the synthetic fault in six of the ten simulation cases.

The fault identification day in Table 5 is the earliest day on which the abnormal consumption fault can be identified. For example, the CHW line reaches +20 on June 8, 1998 in Figure 13.1, which indicates that CHW consumption is at least one standard deviation above expected consumption for June 8-27, 1998. Consequently, the abnormal CHW consumption cannot be identified until we update the data on June 27, 1998. The fault identification day in the simulation case 1 "X_{oa} increase of 3.1%" is June 27, 1998. The modified z-score (Z_m) listed in Table 5 shows the fault magnitude in the 20 fault-flag days. It was calculated according to expression (5) in Section 3.

Table 5 Summary of detected abnormal energy consumption faults by the DET-Date method for the EOM Building

ID	Synthetic Fault	Detection	Consumption	Fault Identification Day	Z_m
1	X_{oa} increase of 3.1%	Y	CHW Increase	6/27/1998	1.2
2	X_{oa} decrease of 3.1%	Y	CHW Decrease	6/27/1998	-1.2
3	T_{cl} increase of 4°F	N			
4	T_{cl} decrease of 4.5°F	Y	CHW Increase	2/21/1998	1.3
5	T_{hl} increase of 10°F	N			
6	T_{hl} decrease of 12°F	N			
7	X_{min} increase of 16%	N			
8	X_{min} decrease of 20%	Y	HW Decrease	12/18/1997	-1.3
9	T_{rc} increase of 1.8°F	Y	CHW Decrease	6/27/1998	-1.3
10	T_{rc} decrease of 1.6°F	Y	CHW Increase	6/27/1998	1.4

Note: Consumption "Increase/Decrease" means the measured consumption is higher/lower than the fault-free simulated consumption. "Z_m" is the modified z-score, which is the ratio of the average daily CHW/HW increase/decrease during the 20 fault-flag days to the standard deviation of CHW/HW residuals during the calibration baseline period.

4.1.3 Fault Detection Results with Simulated Test Data – DET-Toa Method

The simulated consumption under synthetic control changes were treated as the real measured data and the simulated consumption with no control change were treated as simulated data. A program based on the fault detection procedure described in Section 3.2.3 is written. The subroutines of the DET-Toa program are presented in Appendix A. The standard deviations of the residuals in the baseline period, dates, daily average outside air temperatures as well as measured and simulated CHW and HW consumption data are input into the program. The program was run for the ten simulation cases. The program would first determine whether there was any abnormal CHW consumption in the period September 1997-September 1998, and then determine if there was any

abnormal HW consumption in the same period. The fault detection results are summarized in Table 6.

Table 6 Summary of detected abnormal energy consumption faults by the DET-Toa method for the EOM Building

ID	Synthetic Fault	Detection	Consumption	Fault Identification Day	Z_m
1	X_{oa} increase of 3.1%	Y	CHW Increase	6/24/1998	1.2
2	X_{oa} decrease of 3.1%	Y	CHW Decrease	6/24/1998	-1.2
3	T_{cl} increase of 4°F	Y	CHW Decrease	11/23/1997	-2.8
			HW Increase	12/1/1997	3.4
4	T_{cl} decrease of 4.5°F	Y	CHW Increase	12/6/1997	1.3
5	T_{hl} increase of 10°F	N			
6	T_{hl} decrease of 12°F	Y	HW Decrease	12/20/1997	-1.1
7	X_{min} increase of 16%	Y	HW Increase	12/15/1997	1
8	X_{min} decrease of 20%	Y	HW Decrease	11/12/1997	-1.3
9	T_{rc} increase of 1.8°F	Y	CHW Decrease	6/21/1998	-1.3
10	T_{rc} decrease of 1.6°F	Y	CHW Increase	6/21/1998	1.3

Note: Consumption "Increase/Decrease" means the measured consumption is higher/lower than the fault-free simulated consumption. "Z_m" is the modified z-score, which is the ratio of the average daily CHW/HW increase/decrease during the 20 fault-flag days to the standard deviation of CHW/HW residuals during the calibration baseline period

Table 6 shows that the DET-Toa program successfully identified the synthetic fault in nine of the ten simulation cases. It detected abnormally high CHW consumption and low CHW consumption due to excess outside airflow and deficient outside airflow respectively in June 1998; detected abnormally low and high CHW consumption due to the higher and lower cold deck leaving air temperature in November and December 1997 respectively; detected abnormally low HW consumption due to the lower hot deck leaving air temperature in December 1997; detected abnormally high and low HW

consumption due to the excess and deficient minimum airflow in December and November 1997 respectively; and detected abnormally low CHW consumption and high CHW consumption due to the lower and higher room cooling set-point temperature respectively in June 1998.

The fault identification day in Table 6 is the earliest day in which the abnormal consumption fault can be identified with the DET-Toa method. For instance, in case 1 "X_{oa} increase of 3.1%", the fault identification day is June 24, 1998. This means that the DET-Toa method did not identify the excess outside airflow fault until June 24, 1998, when the outside airflow initially increased on October 1, 1997.

4.1.4 Comparison of the Fault Detection Results

The fault detection results of the DET-Date and DET-Toa methods are compared in Table 7. Table 7 shows that the earliest date when a synthetic control change fault was identified (fault identification day) was November 12, 1997, even though the fault first occurred on October 1, 1997. This is because both the DET-Date and the DET-Toa methods can identify faults only when they significantly increase/decrease the energy consumption. The influence of a control change on the CHW/HW consumption depends on weather conditions, building activities and other factors; these factors vary over time. A control change may have little impact on energy use in winter and swing seasons, but may significantly increase the energy consumption in summer. Therefore, the control change that occurred in October may not be noticed until winter or even the next summer. To give an example, in the excess outside airflow case (ID 1), the DET-Date and the DET-Toa methods detected abnormally high CHW consumption in the summer

of 1998 (Table 7). This suggests the control change significantly raised the CHW consumption in summer and has no significant impact on energy consumption during the rest of the year.

Table 7 shows that for the 10 faults tested, the DET-Toa method gives better performance than the DET-Date method. The DET-Toa method identified three faults that were not identified by the DET-Date method. In addition, for the six cases in which both methods successfully identified faults, the DET-Toa method detected the abnormal consumption three days to 2.5 months earlier than the DET-Date method in every case.

Table 7 Comparison of fault detection results for the EOM Building

ID	Synthetic Fault	Detection		Fault Identification Day	
		DET-Date	DET-Toa	DET-Date	DET-Toa
1	X_{oa} increase of 3.1%	Y	Y	6/27/1998	6/24/1998
2	X_{oa} decrease of 3.1%	Y	Y	6/27/1998	6/24/1998
3	T_{cl} increase of 4°F	N	Y		11/23/1997 (CHW fault) 12/1/1997(HW fault)
4	T_{cl} decrease of 4.5°F	Y	Y	2/21/1998	12/6/1997
5	T_{hl} increase of 10°F	N	N		
6	T_{hl} decrease of 12°F	N	Y		12/20/1997
7	X_{min} increase of 16%	N	Y		12/15/1997
8	X_{min} decrease of 20%	Y	Y	12/18/1997	11/12/1997
9	T_{rc} increase of 1.8°F	Y	Y	6/27/1998	6/21/1998
10	T_{rc} decrease of 1.6°F	Y	Y	6/27/1998	6/21/1998

The reason for the better performance of the DET-Toa method can be explained by the dependence of the building energy performance on the outside air temperature. The CHW and HW energy consumption changes during the period from October, 1997

to September, 1998 for the 10 cases are summarized in Appendix B. The plots show that for each control change, the cooling/heating consumption change has a characteristic shape as a function of outside air temperature. The results of case 1 "Outside Airflow Ratio Increase of 3.1%" and case 3 "Cold Deck Leaving Air Temperature Increase of 4°F"are discussed as examples.

4.1.4.1 Case 1 "Outside Airflow Ratio Increase of 3.1%"

Both of the DET-Date and the DET-Toa methods detected an abnormal CHW energy consumption fault in 1998. The fault identification day of the DET-Toa method is June 24, 1998, and it is three days earlier than the fault identification day of the DET-Date method.

Figure 14 illustrates the CHW and HW energy consumption changes caused by the control change "X_{oa} Increase of 3.1%" in the period from October, 1997 to September, 1998, described in the units of standard deviations of the CHW and HW residuals in the baseline. The expression used for the energy consumption change is

$$\text{Energy Consumption Change} = (E_{mea}-E_{sim})/SD_{\_baseline} \quad (10)$$

The CHW or the HW consumption changes due to "X_{oa} Increase of 3.1%" have fixed patterns as a function of outside air temperature (Figure 14). When the outside air temperature increases from 40°F to 90°F, the CHW energy consumption change gradually increases from -1 to 1.5 standard deviation, and the HW energy consumption change slightly decreases from 0.3 to 0 standard deviation. The CHW energy consumption is all below one $SD_{\_baseline}$ when the outside air temperature is below 76°F.

When the DET-Toa program detected the CHW fault, the sub transition matrix included the data from June 3 to June 24. The following analysis focuses on the data in the period from June 3 to June 24 for explaining the reason for the early fault identification day with the DET-Toa method.

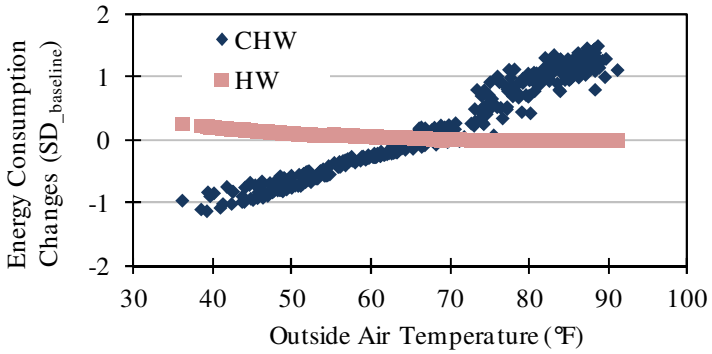

Figure 14 Cooling and heating energy changes plotted as functions of outside air temperature for the period from October 1, 1997 to September 30, 1998 for the EOM Building (Case 1)

As mentioned in Section 4.2.2, the fault identification day is the earliest day on which the abnormal consumption fault can be identified. Therefore, if the fault identification day of the DET-Date method is the same day as the DET-Toa method, the CHW energy consumption change would be all greater than one $SD_{\_baseline}$ during the period from June 5-24. However, the sum of the CHW fault indexes on June 5 is 18, because the CHW energy consumption changes of June 5 and June 6 are less than one $SD_{\_baseline}$ as shown in Figure 15. The small CHW energy consumption change is related to the low outside air temperature of those two days. As shown in Figure 14, the CHW

energy consumption change increases nearly linearly with the increase of outside air temperature. The daily average outside air temperatures of June 5 and June 6 are only 71°F and 74°F respectively (Figure 16). Both of them are lower than the changing-point temperature (76°F) below which the CHW energy consumption changes are all below one $SD_{\_baseline}$.

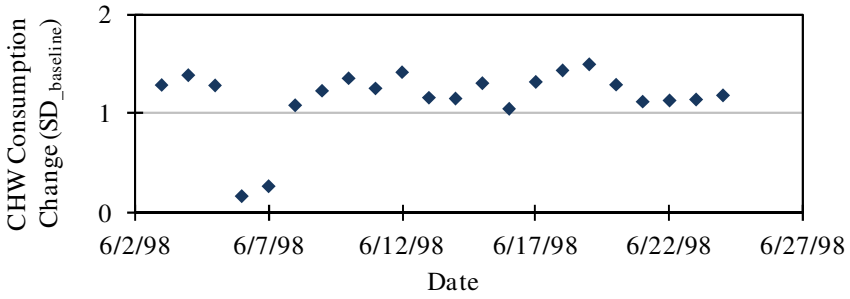

Figure 15 The CHW energy consumption change in the period from June 3-24, 1998 for the EOM Building (Case 1)

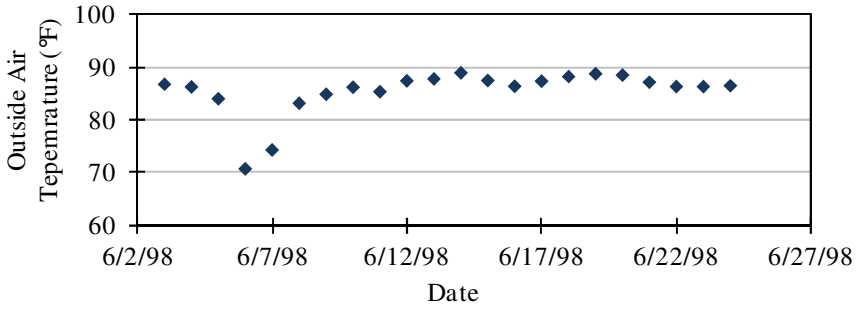

Figure 16 The outside air temperature for the period from June 3-24, 1998

When the data were sorted by the outside air temperature in the ascending order, the data of June 5 and June 6 will move to the top two lines of the sub transition matrix,

and the data of June 3 and June 4 will move to the later part of the sub transition matrix. The time limit 20 was reached in the third line of the matrix, which indicated the appearance of a fault. The sub transition matrix after sorting is provided in Figure 17.

Date	Toa	CHW Fault Index	Sum of Fault Indexes
6/6/1998	70.6	0	18
6/7/1998	74.2	0	19
6/8/1998	83.1	1	20
…	…	…	…
6/4/1998	86.2	1	15
…	…	…	…
6/3/1998	86.7	1	10
…	…	…	…
6/14/1998	88.9	1	1

Figure 17 A portion of the sub transition matrix produced by the DET-Toa program when the CHW fault is identified for the EOM Building (Case 1)

4.1.4.2 Case 3 "Cold Deck Leaving Air Temperature Increase of 4°F"

In Case 3, the DET - Toa method detected an abnormal CHW energy consumption fault on November 23, 1997 and detected an abnormal HW energy consumption fault on December 3, 1997. No fault was detected when the DET-Date method was applied.

The CHW and HW energy consumption changes caused by the control change "Cold Deck Leaving Air Temperature Increase of 4°F" in the period from October, 1997 to September, 1998 are graphed in Figure 18, described in the units of standard deviations of the CHW and HW residuals in the baseline. In Figure 18, the HW energy consumption change first decreases rapidly and then remains relatively constant with

further increase of outside air temperature. The change-point temperature is about 55°F. The CHW energy consumption generally increases when the outside air temperature is lower than 57°F and remains relatively constant when the outside air temperature is higher than 57°F. Figure 18 shows that the CHW energy consumption change is all below minus one $SD_{\_baseline}$ when the outside air temperature is smaller than 57°F, and the HW energy consumption change is all above one $SD_{\_baseline}$ when the outside air temperature is smaller than 55°F.

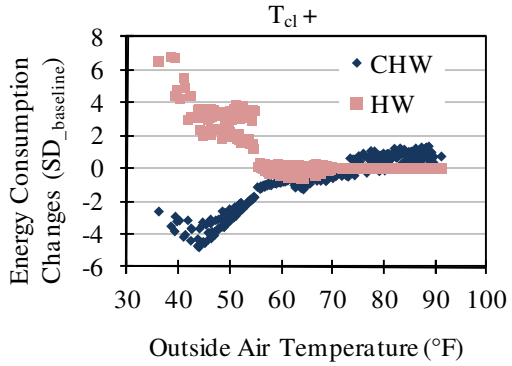

Figure 18 Cooling and heating energy changes plotted as functions of outside air temperature for the period from October 1, 1997 to September 30, 1998 for the EOM Building (Case 3)

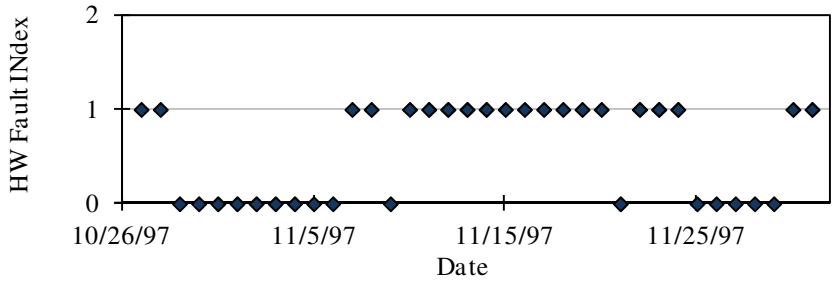

Figure 19 The HW fault indexes during the period from October 27 - December 1, 1997 for the EOM Building (Case 3)

When the DET-Toa program identified the HW fault, the sub transition matrix contained the data from October 27 to December 1. Figure 19 shows that there are no more than 11 consecutive days where the HW fault index is 1 during this period. The DET-Date method could not detect a fault using its criterion. Notice in Figure 20, the outside air temperatures of the days where the HW fault index is 1 are all below 55°F, and the outside air temperatures of the days where the HW fault index is 0 is are above 55°F. Therefore, the days where the HW fault index is 1 move to the top of the sub transition matrix when the matrix is sorted by the outside air temperature in ascending order. The time limit 20 was reached in counting the sum of the fault indexes. The sub transition matrix after sorting is provided in Figure 21.

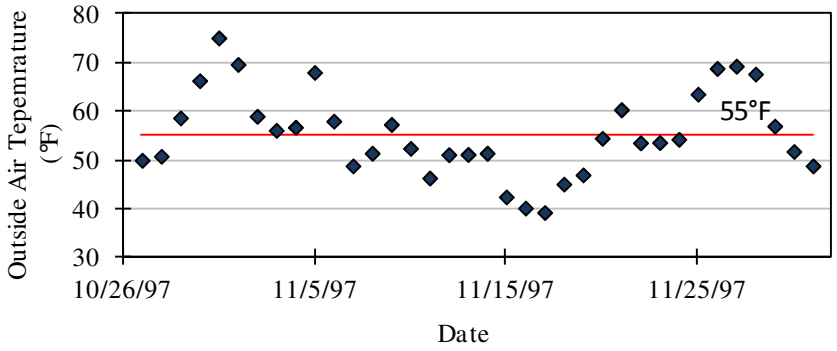

Figure 20 The outside air temperature in the period from October 27 - December 1, 1997

Similar analysis was conducted for the detected CHW fault. When the DET-Toa program detected a CHW fault, the sub transition matrix contained the data from October 23 to November 27. The DET-Date method did not detect the fault as there was no 20 consecutive day period during which the fault index of each day was -1. For the

DET-Toa method, the matrix moved the days where the outside air temperature is below 57°F to the top when the matrix is sorted by the outside air temperature in ascending order. As shown in Figure 18, the CHW energy consumption change values are all below minus one $SD_{\_baseline}$ when the outside air temperature is lower than 57°F. As a result, the 20 limit was reached when the sum of the fault indexes was counted.

Date	Toa	HW Fault Index	Sum of Fault Indexes
11/17/1997	39.3	1	20
11/16/1997	40.2	1	19
11/15/1997	42.5	1	18
…	…	…	…
11/20/1997	54.5	1	1
…	…	…	…
10/31/1997	75.1	0	0

Figure 21 A portion of the sub transition matrix produced by the DET-Toa program when the HW Fault is identified for the Eller Building (Case 3)

4.1.4.3 Summary

The comparison of the results of the DET-Toa and the DET-Date methods in 10 cases shows that the DET-Toa fault detection method is superior. The DET-Toa method detected more faults and detected them earlier then the DET-Date method. The analysis of Case 1 and Case 3 results implies that the better performance of the DET-Toa method is due to the energy-use changes' dependence on the outside air temperature.

4.2 Simulated for SDVAV System

4.2.1 Simulated Data Sets

The simulated test building is the Veterinary Research Building on the Texas A&M University campus in College Station. It is a five story building with 115,000 ft^2 of conditioned space. The building is comprised primarily of laboratories but also contains classrooms and offices. Thermal energy is supplied to the building as hot water for heating and chilled water for cooling from the central utility plant. The majority of the building is served by five SDVAV AHUs. The commissioning was completed in November of 1996. The ABCAT simulations were calibrated to the baseline consumption period from January 1-July 20, 1998. The standard deviations of the residuals in the baseline period were 5.2 MMBtu/day and 3.7 MMBtu/day respectively for CHW and HW energy consumption. They are 7.9% and 5.7% of the average daily CHW and HW consumption during the baseline calibration period. The building usage was consistent in the study. The occupant density was assigned weekday and weekend values within the simulation. The measured daily building electricity use (none of which is directly used for heating or cooling) was assumed to be converted into heat gain within the building by the simulation.

It is assumed one of the following ten synthetic control changes happened in the simulation period (January 1-December 31, 2000):

1. Outside airflow ratio (X_{oa}) increase of 6.5%
2. Outside airflow ratio decrease of 6.5%
3. Cooling coil leaving air temperature (T_{cl}) increase of 1.7°F

4. Cooling coil leaving air temperature decrease of 1.5°F
5. Preheat temperature (T_{preh}) increase of 3.3°F
6. Preheat temperature decrease of 4°F
7. Minimum airflow ratio (X_{min}) increase of 4.9%
8. Minimum airflow ratio decrease of 6.5%
9. Room cooling set-point temperature (T_{rc}) increase of 2.8°F
10. Room cooling set-point temperature decrease of 1.7°F

Note: The denominator of outside airflow ratio and minimum airflow ratio is the maximum design airflow volume in the system.

The energy consumption with no control change and under the ten synthetic control changes was generated by the simulation module of ABCAT (Curtin 2007). Simulation was conducted for College Station weather data and the measured building electricity consumption during that period. The yearly average cooling plus heating consumption in 2000 was 67.88MMBtu/day in the simulation when there was no fault.

The influence of the ten control changes on the monthly average energy use change index (expression (9) in Section 4.1.1) in 2000 is demonstrated in Figure 22. The parameter changes were chosen so the maximum monthly average cooling and heating consumption deviation (Blank bars in Figure 22 caused by each synthetic control change was 10% of the yearly average cooling and heating consumption (67.88MMBtu/day) when there was no fault.

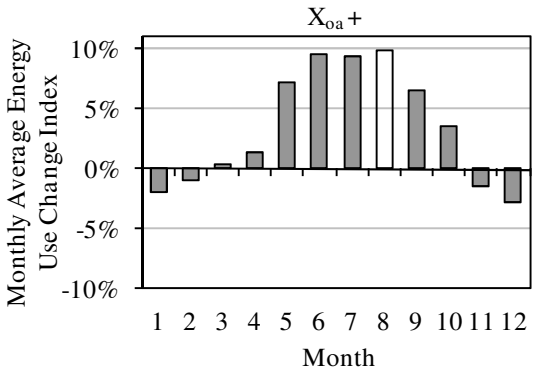

Figure 22.1

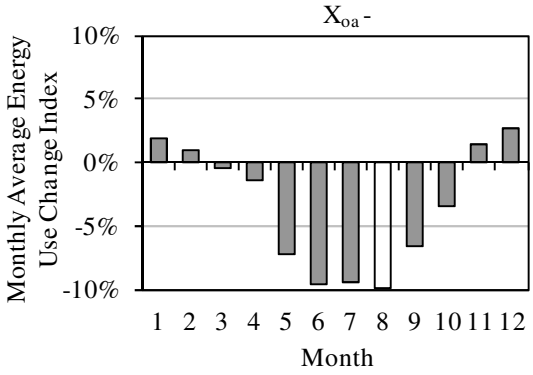

Figure 22.2

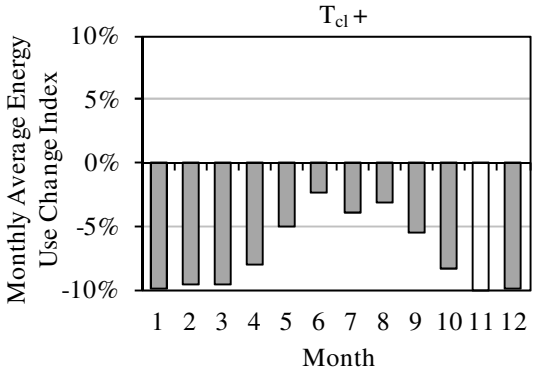

Figure 22.3

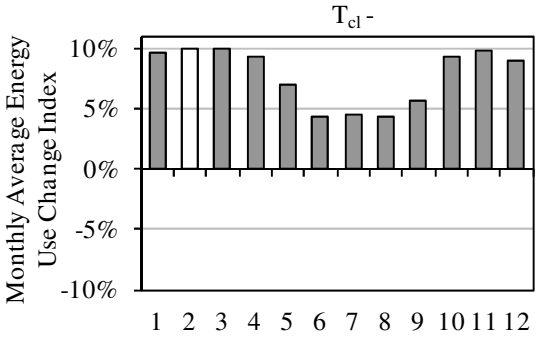

Figure 22.4

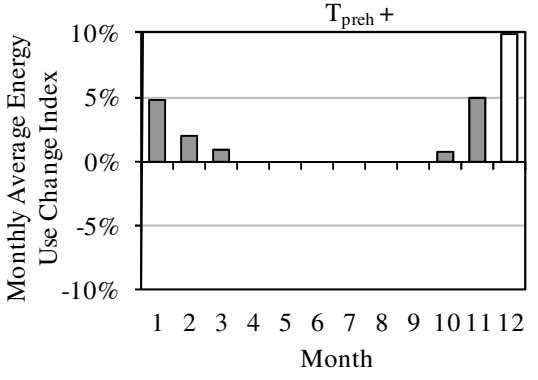

Figure 22.5

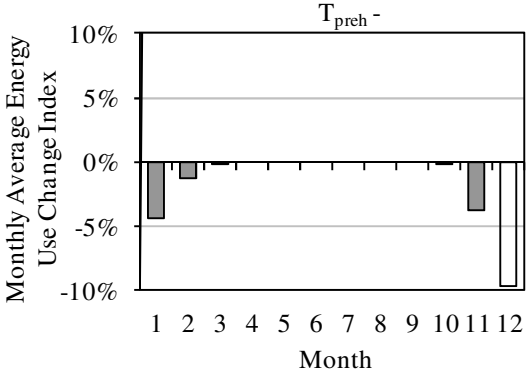

Figure 22.6

Figure 22 Monthly average energy consumption use changes under different synthetic control changes in 2000 for the Veterinary Research Building. (22.1) X_{oa} increase of 6.5%, (22.2) X_{oa} decrease of 6.5%, (22.3) T_{cl} increase of 1.7°F, (22.4) T_{cl} decrease of 1.5°F, (22.5) T_{preh} increase of 3.3°F, (22.6) T_{preh} decrease of 4°F, (22.7) X_{min} increase of 4.9%, (22.8) X_{min} decrease of 6.5%, (22.9) T_{rc} increase of 2.8°F, (22.10) T_{rc} decrease of 1.7°F

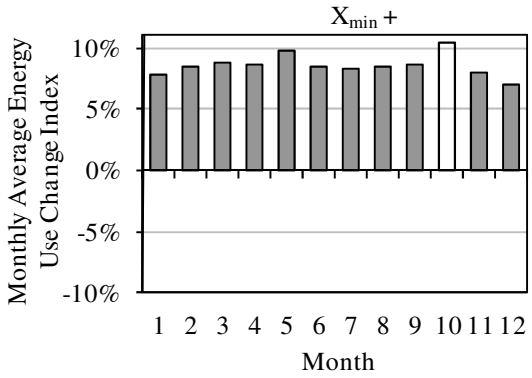

Figure 22.7

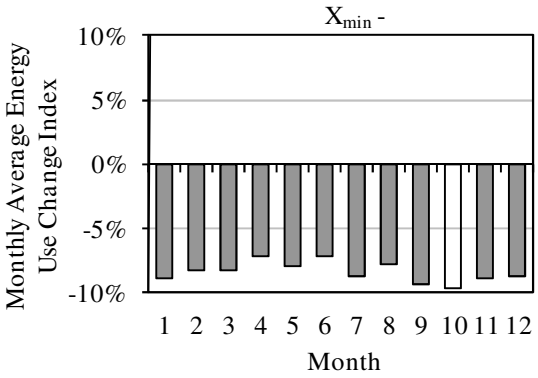

Figure 22.8

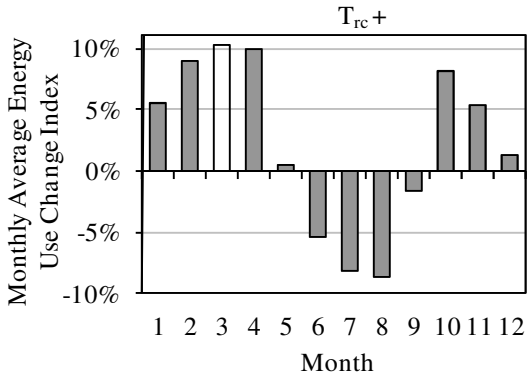

Figure 22.9

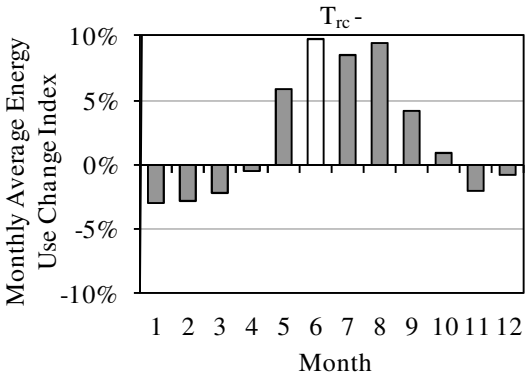

Figure 22.10

Figure 22 Continued

4.2.2 Fault Detection Results with Simulated Test Data - DET-Date Method

The simulated consumption under synthetic control changes and the simulated consumption with no control change were treated as the real measured data and the simulated data respectively. DET-Date plots based on the sum of the 20 consecutive fault indexes in time series are summarized in Figure 23. See expression (4) in Section 3 for the definition of the fault index. The meaning of the DET-Date plot is illustrated in

detail in Section 4.1.2. The CHW line in Figure 23.1 reaches +20 on June 6. This means that CHW consumption is at least one standard deviation above expected consumption for June 9-28, 1998 and hence corresponds to an identified CHW Fault. Similarly, the abnormal consumption caused by four of the other synthetic control changes is identified in the rest of the DET-Date plots (Figure 23.2 and Figures 23.8-10). The DET-Date plots failed to detect five of the synthetic control changes.

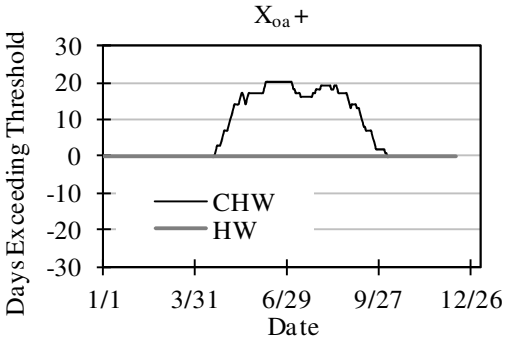

Figure 23.1

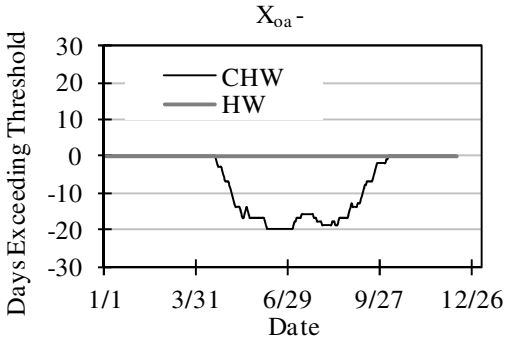

Figure 23.2

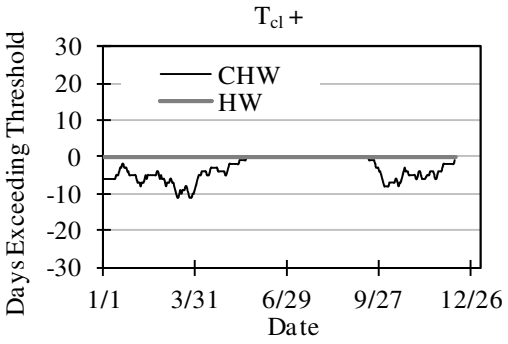

Figure 23.3

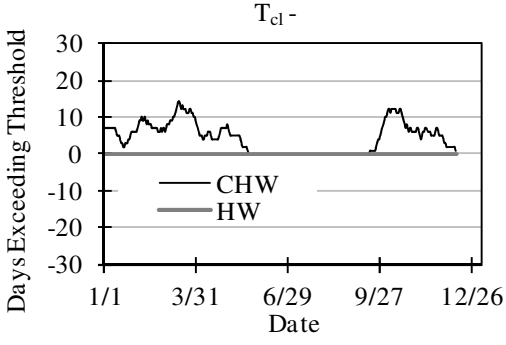

Figure 23.4

Figure 23 Days Exceeding Threshold-Date Plots under different synthetic control changes in 2000 for the Veterinary Research Building. (23.1) X_{oa} increase of 6.5%, (23.2)X_{oa} decrease of 6.5%, (23.3) T_{cl} increase of 1.7°F, (23.4) T_{cl} decrease of 1.5°F, (23.5) T_{preh} increase of 3.3°F, (23.6) T_{preh} decrease of 4°F, (23.7) X_{min} increase of 4.9%, (23.8) X_{min} decrease of 6.5%, (23.9) T_{rc} increase of 2.8°F, (23.10) T_{rc} decrease of 1.7°F

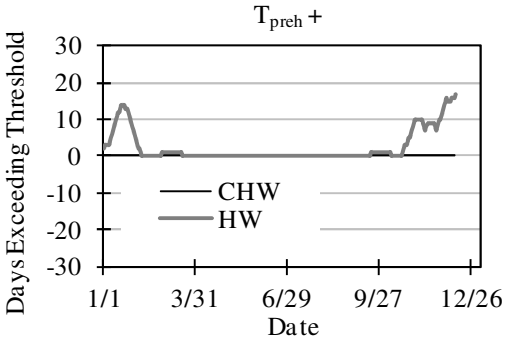

Figure 23.5

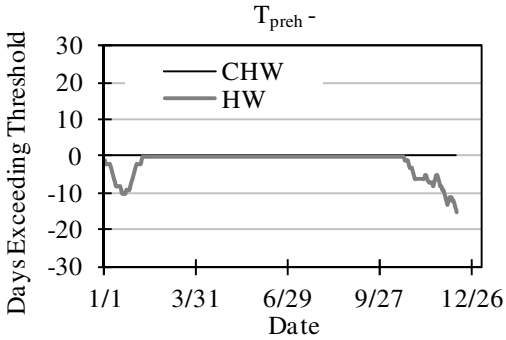

Figure 23.6

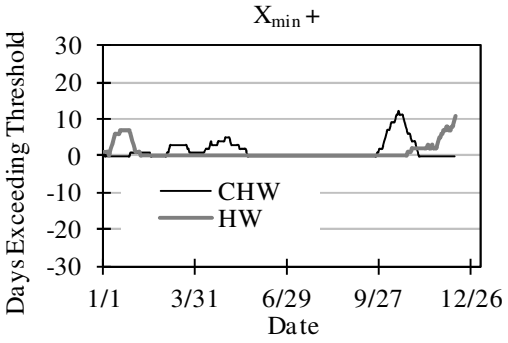

Figure 23.7

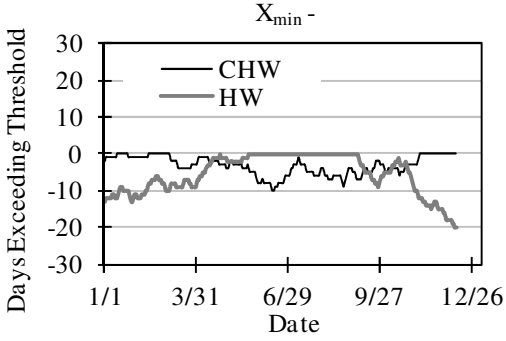

Figure 23.8

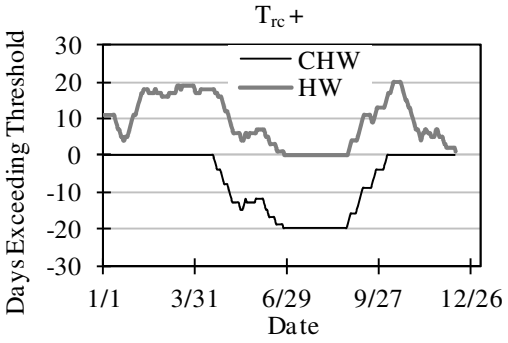

Figure 23.9

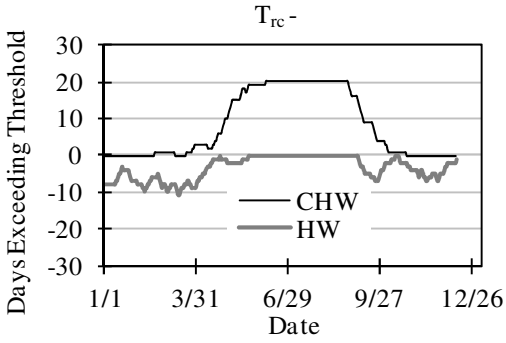

Figure 23.10

Figure 23 Continued

The fault detection results with the DET-Date method in the ten simulation cases are summarized in Table 8. The DET-Date method noticed abnormally high and low CHW consumption due to the excess and deficient outside airflow respectively in June;

abnormally low HW consumption due to the deficient minimum airflow in December; abnormally low CHW consumption and high HW consumption due to the lower room cooling set-point temperature in July and October respectively; and abnormally high CHW consumption due to the higher room cooling set-point temperature in June. Overall, the DET-Date method effectively detected the abnormal consumption fault in five of the ten simulation cases (Table 8).

Table 8 Summary of detected abnormal energy consumption faults by the DET-Date method for the Veterinary Research Building

ID	Synthetic Fault	Detection	Consumption	Fault Identification Day	Z_m
1	X_{oa} increase of 6.5%	Y	CHW Increase	6/29/2000	1.4
2	X_{oa} decrease of 6.5%	Y	CHW Decrease	6/29/2000	-1.4
3	T_{cl} increase of 1.7°F	N			
4	T_{cl} decrease of 1.5°F	N			
5	T_{preh} increase of 3.3°F	N			
6	T_{preh} decrease of 4°F	N			
7	X_{min} increase of 4.9%	N			
8	X_{min} decrease of 6.5%	Y	HW Decrease	12/29/2000	-1.3
9	T_{rc} increase of 2.8°F	Y	CHW Decrease	7/15/2000	-1.6
			HW Increase	10/30/2000	1.8
10	T_{rc} decrease of 1.7°F	Y	CHW Increase	6/28/2000	1.5

Note: Consumption "Increase/Decrease" means the measured consumption is higher/lower than the fault-free simulated consumption. "Z_m" is the modified z-score, which is the ratio of the average daily CHW/HW increase/decrease during the 20 fault-flag days to the standard deviation of CHW/HW residuals during the calibration baseline period.

4.2.3 Fault Detection Results with Simulated Test Data – DET-Toa Method

The DET-Toa program was run for the ten simulation cases. The fault detection results are tabulated in Table 9. The DET-Toa method successfully identified the

synthetic fault in each of the ten simulation cases. The meaning of the fault identification day in Table 9 has been illustrated in Section 4.1.3.

Table 9 Summary of detected abnormal energy consumption faults by the DET-Toa method for the Veterinary Research Building

ID	Synthetic Fault	Detection	Consumption	Fault Identification Day	Z_m
1	X_{oa} increase of 6.5%	Y	CHW Increase	6/4/2000	1.3
2	X_{oa} decrease of 6.5%	Y	CHW Decrease	6/4/2000	-1.3
3	T_{cl} increase of 1.7°F	Y	CHW Decrease	4/2/2000	-1.1
4	T_{cl} decrease of 1.5°F	Y	CHW Increase	3/15/2000	1
5	T_{preh} increase of 3.3°F	Y	HW Increase	11/9/2000	1.1
6	T_{preh} decrease of 4°F	Y	HW Decrease	12/4/2000	-1.3
7	X_{min} increase of 4.9%	Y	HW Increase	12/28/2000	1.1
8	X_{min} decrease of 6.5%	Y	HW Decrease	12/27/2000	-1.3
9	T_{rc} increase of 2.8°F	Y	CHW Decrease	7/14/2000	-1.6
			HW Increase	2/14/2000	1.9
10	T_{rc} decrease of 1.7°F	Y	CHW Increase	5/30/2000	1.4
			HW Decrease	3/18/2000	-1.2

Note: Consumption "Increase/Decrease" means the measured consumption is higher/lower than the fault-free simulated consumption. "Z_m" is the modified z-score, which is the ratio of the average daily CHW/HW increase/decrease during the 20 fault-flag days to the standard deviation of CHW/HW residuals during the calibration baseline period.

4.2.4 Comparison of the Fault Detection Results

Similar to the situation of the EOM Building, the DET-Toa method gives better performance than the DET - Date method for detecting faults (Table 10). The DET-Toa method identified five faults that were not identified by the DET-Date method. Furthermore, for the five cases in which both methods successfully identified faults, the DET-Toa method noticed the abnormal consumption one day to 8.5 months earlier than the DET-Date method in every case.

The CHW and HW energy consumption changes versus daily average outside air temperature in 2000 are graphed in Appendix B. The figure shows that the outside air temperature has a significant impact on the behavior of the energy consumption variation. The influence of the temperature dependence on the difference of the fault detection results of the DET-Date and the DET-Toa methods has been illustrated in Section 4.1.4 through two case studies. One more example is provided below.

Table 10 Comparison of fault detection results for the Veterinary Research Building

ID	Synthetic Fault	Detection		Fault Identification Day	
		DET-Date	DET-Toa	DET-Date	DET-Toa
1	X_{oa} increase of 6.5%	Y	Y	6/29/2000	6/4/2000
2	X_{oa} decrease of 6.5%	Y	Y	6/29/2000	6/4/2000
3	T_{cl} increase of 1.7°F	N	Y		4/2/2000
4	T_{cl} decrease of 1.5°F	N	Y		3/15/2000
5	T_{preh} increase of 3.3°F	N	Y		11/9/2000
6	T_{preh} decrease of 4°F	N	Y		12/4/2000
7	X_{min} increase of 4.9%	N	Y		12/28/2000
8	X_{min} decrease of 6.5%	Y	Y	12/29/2000	12/27/2000
9	T_{rc} increase of 2.8°F	Y	Y	7/15/2000 (CHW fault) 10/30/2000(HW fault)	7/14/2000 (CHW fault) 2/14/2000 (HW fault)
10	T_{rc} decrease of 1.7°F	Y	Y	6/28/2000 (CHW fault)	5/30/2000 (CHW fault) 3/18/2000 (HW fault)

4.2.4.1 Case 4 Preheat Temperature Decrease of 4°F

The DET-Toa method identified an abnormally low HW energy consumption fault on March 15, while no fault was identified by the DET-Date method. Figure 24 shows that the HW energy consumption change remains constant at -1.3 $SD_{\_baseline}$ over the low outside air temperature range, then sharply increases to 0 and remains constant at

0 over the high outside air temperature range. The HW energy consumption change is all below minus one $SD_{\_baseline}$ when the outside air temperature is below 47°F.

The sub transition matrix provided by the DET-Toa program included the data from January 5 to December 4, when the program detected the abnormally low HW consumption. There are nine days in January, two days in February, six days in November, and three days in December when the outside air temperature is below 47°F. The fault indexes in these days are all -1. The DET-Date method cannot detect faults using its criterion because these days are not consecutive in time. The DET-Toa method sorted the matrix by the outside air temperature in ascending order; thus, all of the twenty days mentioned above would move to the top of the sub transition matrix. The time limit -20 was reached in counting the sum of fault indexes. The sub transition matrix after sorting is demonstrated in Figure 25.

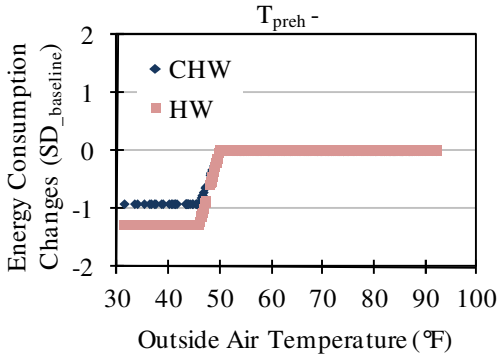

Figure 24 Cooling and heating energy changes plotted as functions of outside air temperature for the period from October, 1997 to September, 1998 for the Veterinary Research Building (Case 4)

Date	Toa	CHW Fault Index	Sum of Fault Indexes
11/17/1997	39.3	-1	-20
11/16/1997	40.2	-1	-19
11/15/1997	42.5	-1	-18
11/18/1997	45.1	-1	-17
11/11/1997	46.4	-1	-16
11/19/1997	47.0	-1	-15
11/7/1997	48.9	-1	-15
10/27/1997	50.0	-1	-14
10/28/1997	50.8	-1	-13
11/12/1997	51.2	-1	-12
11/13/1997	51.2	-1	-11
11/14/1997	51.5	-1	-10
11/8/1997	51.5	-1	-9
11/10/1997	52.5	-1	-8
11/22/1997	53.6	-1	-7
11/23/1997	53.7	-1	-6
11/20/1997	54.5	-1	-5
11/3/1997	56.2	-1	-4
10/26/1997	56.5	-1	-3
11/4/1997	56.8	-1	-2
11/9/1997	57.4	0	-1
11/6/1997	58.0	0	-1
10/29/1997	58.7	0	-1
11/2/1997	59.0	0	-1
11/21/1997	60.4	0	-1
10/23/1997	65.1	-1	-1
10/30/1997	66.4	0	0
11/5/1997	68.0	0	0
11/1/1997	69.7	0	0
10/25/1997	75.0	0	0
10/31/1997	75.1	0	0
10/24/1997	76.0	0	0

Figure 25 The sub transition matrix produced by the DET-Toa program when the CHW fault is identified for the Veterinary Research Building (Case 4)

4.3 Summary

The simulation tests were conducted to investigate the fault-detection capability of the DET-Date method and the DET-Toa method. Two buildings, one with DDVAV systems and the other one with SDVAV systems, were selected for the tests. Ten synthetic faults were assumed to happen and last for one year in each of the two buildings. The DET-Date method successfully identified the synthetic faults in six of the ten simulation cases in the building with DDVAV system, and in five of the ten simulation cases in the building with SDVAV system. The DET-Toa method effectively identified the synthetic faults in nine of the ten simulation cases in the building with DDVAV system, and identified the synthetic faults in all of the ten simulation cases in the building with SDVAV system.

Detection accuracy is determined as the ratio of the number of cases where the fault detection method successfully identified the abnormal consumption to the total number of the simulation cases. The detection accuracy for the DET-Date and DET-Toa methods are 55% and 95% respectively in the total 20 simulation test cases. It is concluded that the DET-Toa method gives better performance not only because it detected more synthetic faults, but also because it detected the faults earlier than the DET-Date method. Analysis indicates that the superior performance of the DET-Toa method results from using outside air temperature dependence in its fault detection metric.

5. ANALYSIS OF WHOLE BUILDING FAULT DIAGNOSIS APPROACHES WITH SIMULATION TEST

Two methods, namely, the Cosine Similarity method and the Euclidean Distance Similarity method that use similarity measures for determining the underlying cause of abnormal CHW and HW energy consumption are presented in Section 3. Both methods can order list of control changes according to the probability to that each caused the observed abnormal energy consumption fault. In this section, both methods are validated using simulated data for the whole building fault diagnosis. The simulation test was performed on SDVAV and DDVAV systems. The fault diagnosis performance of the two methods is compared and discussed in this section.

All the simulation data were produced by the simulation module of ABCAT (Curtin 2007). Simulation was performed for College Station weather data and the measured building electricity consumption.

5.1 Simulated for DDVAV System

5.1.1 Simulated Data Sets

The simulated test building is the Bush Academic Building located on the Texas A&M University west campus in College Station. It consists primarily of offices and classrooms. The building has three floors for a total area of 133,326 ft^2. It is generally occupied on weekdays during the day. Thermal energy is supplied to the building in the form of hot water and chilled water from the central utility plant. The HVAC system in the building is a DDVAV system. The commissioning work on this building was

completed in May of 2007. The ABCAT simulation was calibrated to the baseline consumption period of weekdays from June 1, 2007 to April 20, 2008.

It is assumed one of the following ten synthetic control changes lasted in the weekday period from July, 2008 - June, 2009:

1. Outside airflow ratio (X_{oa}) increase of 3.9%
2. Outside air preheat temperature (T_{preh}) increase of 14°F
3. The amount of heat leakage from preheat coil (PreHL) is 40kBtu/hr
4. Cold deck leaving air temperature (T_{cl}) decrease of 2°F
5. Hot deck leaving air temperature (T_{hl}) increase of 30°F
6. The amount of heat leakage from heating coil (HL) is 60kBtu/hr
7. Minimum airflow ratio (X_{min}) increase of 2.5%
8. Room heating set-point temperature (T_{rh}) increase of 3°F
9. Room cooling set-point temperature (T_{rc}) decrease of 2.2°F
10. Terminal box damper leakage (TDL) increases 7.5%

Note: The denominator of X_{oa}, X_{min}, TDL is the maximum design airflow volume in the system.

The ten synthetic control changes are referred to as ten simulation test cases. The influence of the ten control changes on the monthly average energy use change index from July 2008 to June 2009 are demonstrated in Figure 26. The parameter changes were chosen so the maximum monthly average cooling and heating consumption change caused by each synthetic control change was 10% of the yearly average cooling and heating consumption (20.11MMBtu/day) when there was no fault.

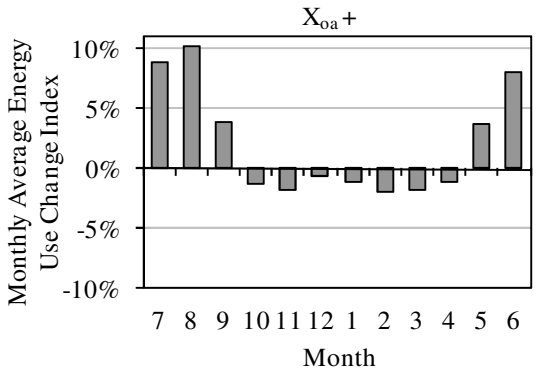

Figure 26.1

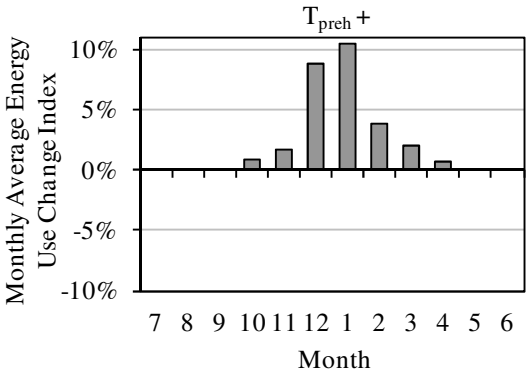

Figure 26.2

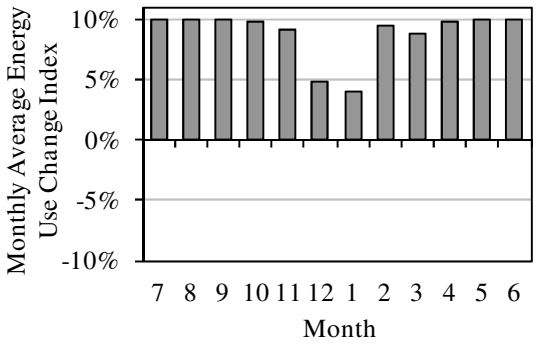

Figure 26.3

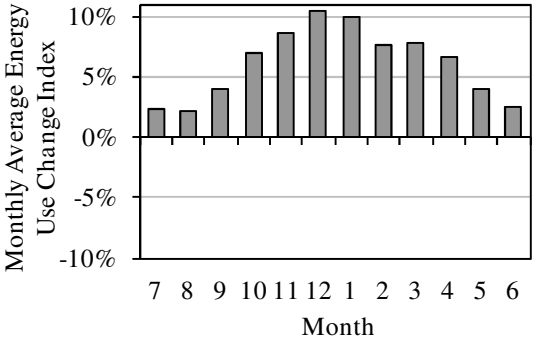

Figure 26.4

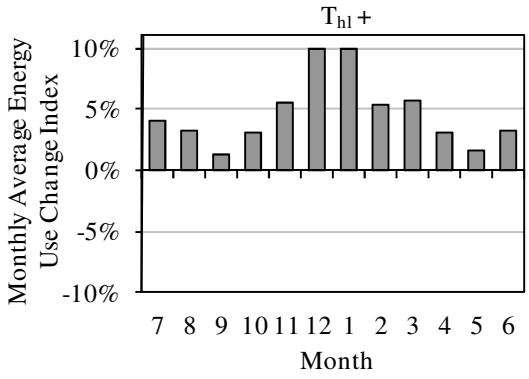

Figure 26.5

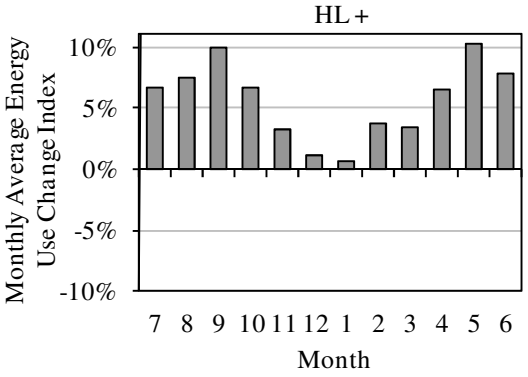

Figure 26.6

Figure 26 Monthly average energy (cooling plus heating) consumption changes under different synthetic control changes in the period from July, 2008 - June, 2009 for the Bush Academic Building. (26.1) X_{oa} increase of 3.9%, (26.2) T_{preh} increase of 14°F, (26.3) PreHL is 40kBtu/hr, (26.4) T_{cl} decrease of 2°F, (26.5) T_{hl} increase of 30°F, (26.6) HL is 60kBtu/hr, (26.7) X_{min} increase of 2.5%, (26.8) T_{rh} increase of 3°F, (26.9) T_{rc} decrease of 2.2°F, (26.10) TDL increase of 7.5%

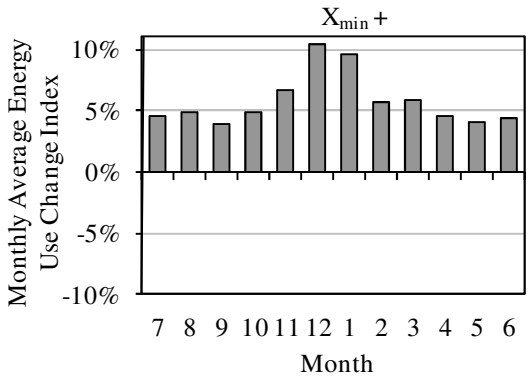

Figure 26.7

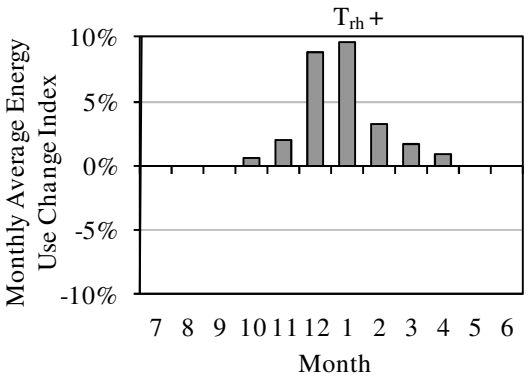

Figure 26.8

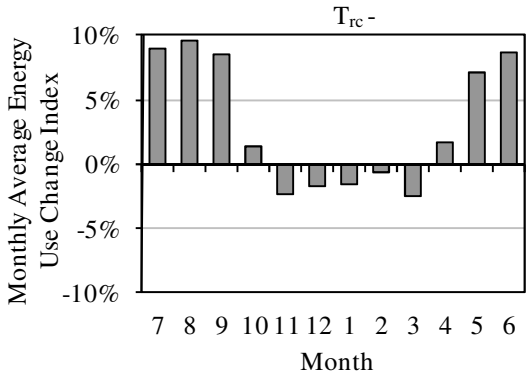

Figure 26.9

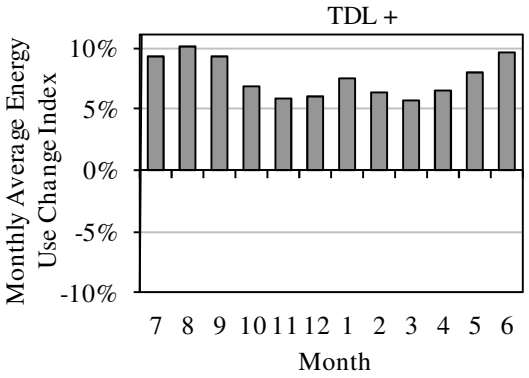

Figure 26.10

Figure 26 Continued

The reference control change library including 17 types of control changes is shown on Table 11. Each row includes a type of reference control change and there are five levels of fault severity for each control change. The first column contains the ID of the reference control change. The second column provides the key words describing the control change. The remaining columns present the different magnitudes of the control change. For example, "-2%" in the first row means outside airflow ratio decreased 2% and "2%" in the second row means outside airflow ratio increased 2%.

Table 11 Reference control change library for the Bush Academic Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%	-4%	-6%	-8%	-10%	
2	X_{oa} increase	2%	4%	6%	8%	10%	
3	T_{preh} decrease	-3	-6	-9	-12	-15	°F
4	T_{preh} increase	3	6	9	12	15	°F
5	PreHL increase	10	20	30	40	50	kBtu/hr
6	T_{cl} decrease	-2	-4	-6	-8	-10	°F
7	T_{cl} increase	2	4	6	8	10	°F
8	T_{hl} decrease	-2	-4	-6	-8	-10	°F
9	T_{hl} increase	2	4	6	8	10	°F
10	HL increase	20	40	60	80	100	kBtu/hr
11	X_{min} decrease	-2%	-4%	-6%	-8%	-10%	
12	X_{min} increase	2%	4%	6%	8%	10%	
13	T_{rc} decrease	-1	-2	-3	-4	-5	°F
14	T_{rc} increase	1	2	3	4	5	°F
15	T_{rh} decrease	-1	-2	-3	-4	-5	°F
16	T_{rh} increase	1	2	3	4	5	°F
17	TDL increase	2%	4%	6%	8%	10%	

Note: The denominator of X_{oa}, X_{min} and TDL is the maximum design airflow volume in the system.

The specific input parameter was changed in the calibrated simulation model in ABCAT to generate the daily cooling and heating energy consumption when there was a known control change from the reference library persisting during the fault period (July 2008-June 2009). The simulated consumption under the ten synthetic control changes was treated as the real measured data in the fault period.

5.1.2 Diagnostic Results with Simulated Test Data

5.1.2.1 Results of Case 1 "Outside Airflow Ratio Increase of 3.9%"

Case 1 "Outside airflow ratio increase of 3.9%" is selected as an example to illustrate the process of fault diagnosis specifically. The observed fault signature vector and the reference control change signature vectors were generated according to expression (6) in Section 3.3.3. The observed fault signature vector components versus outside air temperature are graphed in Figure 27.

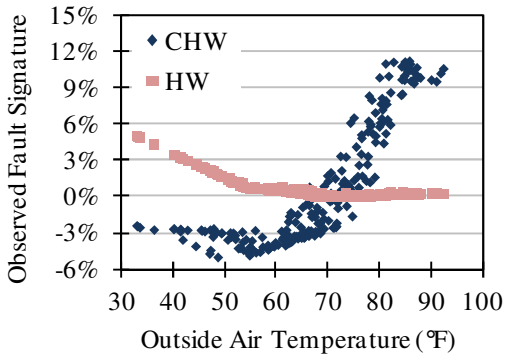

Figure 27 The observed fault signature vector components plotted as a function of outside air temperature for the weekday period from July, 2008 to June, 2009 for the Bush Academic Building (Case 1)

The cosine similarity and the Euclidean distance similarity values were computed between the observed fault signature vector and each of the reference control change signature vectors following expressions (7) and (8) in Section 3. There are 17 different types of reference control changes. For a single reference control change there are five reference control change signature vectors. The similarity results are demonstrated in Table 12 and Table 13. In Table 12 and Table 13, column one includes the ID of the

reference control change and column two describes the type of reference control change. As mentioned above, there are five severity levels (I-V) for a control change (Table 11). The next five columns present the values of the cosine similarity (Table 12)/the Euclidean distance similarity (Table 13) between the observed fault signature vector and the different reference control change signature vectors with different magnitude levels. For example, the cosine similarity in the cell (Row 1, Column "Magnitude I") is -0.99. This suggests the cosine similarity between the observed fault signature vector and the reference control change "X_{oa} - 2%" signature vector is -0.99. The corresponding value of the magnitude levels of each control change (I, II, III, IV and V) can be found in Table 11.

Table 12 Cosine similarity results in Case 1 "Outside Airflow Ratio Increase of 3.9%" for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	-0.99	-0.99	-0.99	-0.99	-0.99	-0.99
2	X_{oa} increase	1.00	1.00	1.00	1.00	1.00	1.00
3	T_{preh} decrease	0.06	0.04	0.04	0.03	0.02	0.06
4	T_{preh} increase	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
5	PreHL increase	0.33	0.32	0.32	0.32	0.31	0.33
6	T_{cl} decrease	-0.06	-0.06	-0.05	-0.01	0.03	0.03
7	T_{cl} increase	0.06	0.11	0.39	0.62	0.75	0.75
8	T_{hl} decrease	-0.06	-0.07	-0.07	-0.03	0.02	0.02
9	T_{hl} increase	0.06	0.05	0.05	0.04	0.04	0.06
10	HL increase	-0.01	0.25	0.41	0.44	0.45	0.45
11	X_{min} decrease	-0.38	-0.37	-0.30	-0.25	-0.21	-0.21
12	X_{min} increase	0.34	0.30	0.30	0.31	0.35	0.35
13	T_{rc} decrease	0.70	0.70	0.71	0.72	0.73	0.73
14	T_{rc} increase	-0.70	-0.70	-0.70	-0.68	-0.66	-0.66
15	T_{rh} decrease	-0.11	-0.11	-0.10	-0.09	-0.09	-0.09
16	T_{rh} increase	0.12	0.11	0.11	0.12	0.12	0.12
17	TDL increase	0.60	0.60	0.59	0.59	0.59	0.60

The largest cosine similarity/ Euclidean distance similarity among the cases with the identical reference control change are listed in the rightmost column (Column "Max") in Table 12 and Table 13. They are chosen to be representative similarities of the corresponding control changes.

Table 13 Euclidean distance similarity results in Case 1 "Outside Airflow Ratio Increase of 3.9%" for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.20	0.11	0.06	0.04	0.02	0.20
2	X_{oa} increase	0.60	0.97	0.57	0.34	0.20	0.97
3	T_{preh} decrease	0.33	0.31	0.29	0.28	0.27	0.33
4	T_{preh} increase	0.36	0.35	0.31	0.27	0.23	0.36
5	PreHL increase	0.35	0.31	0.26	0.20	0.15	0.35
6	T_{cl} decrease	0.22	0.08	0.03	0.01	0.00	0.22
7	T_{cl} increase	0.23	0.14	0.11	0.06	0.02	0.23
8	T_{hl} decrease	0.33	0.31	0.29	0.27	0.25	0.33
9	T_{hl} increase	0.36	0.36	0.36	0.35	0.34	0.36
10	HL increase	0.35	0.33	0.27	0.17	0.10	0.35
11	X_{min} decrease	0.21	0.13	0.09	0.07	0.05	0.21
12	X_{min} increase	0.32	0.16	0.07	0.02	0.01	0.32
13	T_{rc} decrease	0.41	0.32	0.18	0.08	0.03	0.41
14	T_{rc} increase	0.21	0.11	0.06	0.04	0.02	0.21
15	T_{rh} decrease	0.29	0.23	0.19	0.15	0.12	0.29
16	T_{rh} increase	0.36	0.29	0.22	0.18	0.14	0.36
17	TDL increase	0.40	0.30	0.17	0.08	0.04	0.40

Figures 28 and 29 compare the representative similarities of the 17 types of reference control change and show control changes ordered in descending order of their representative similarity values with the results in percentage format. Each bar indicates the representative similarity between the observed fault and a type of reference control change. The reference control change IDs shown on the X axis correspond to the IDs

listed in the first column of Table 11. In Figures 28 and 29, the bars neatly decrease in height from left to right. The first bar appearing on the left has the largest similarity value while the last bar appearing on the right has the smallest similarity value.

Based on the description of the cosine similarity and the Euclidean distance similarity in Section 2, the higher the value of similarity measure, the higher the degree of compliance between two vectors. The cosine similarity between the observed fault and the reference control change (ID 2) is 100% and ranks No.1 among the 17 types of reference control change (Figure 28). This suggests that the energy pattern of the signature vector of control change "outside air flow ratio increase" is closest to the energy pattern of the observed fault signature vector. As a result, the diagnosis result with the Cosine Similarity method shows that the control change "outside air flow ratio increase" has the highest probability as the cause of the observed fault in the simulation case 1.

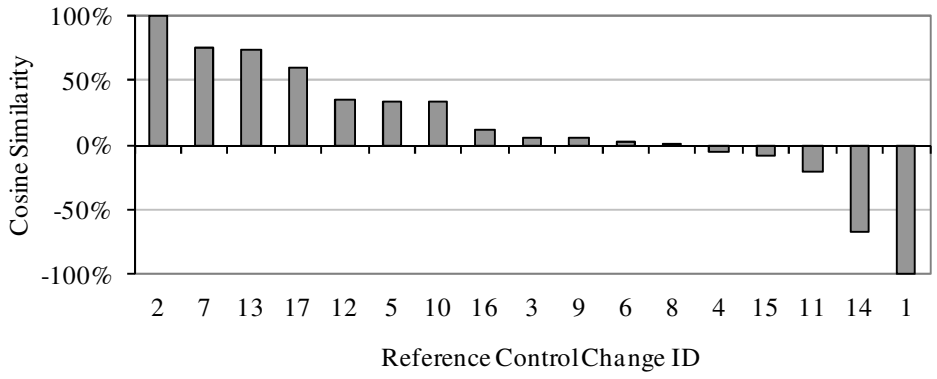

Figure 28 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 1)

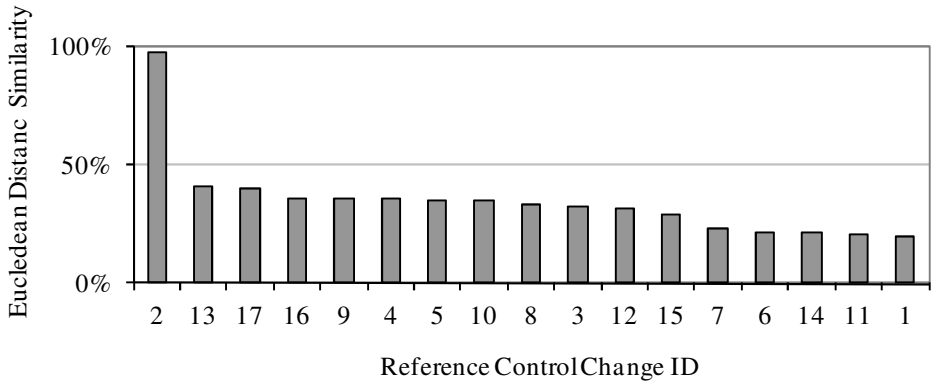

Figure 29 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 1)

Similarly, the Euclidean distance similarity between the observed fault and the reference control change (ID 2) is 97% and ranks No.1 among the 17 types of reference control change (Figure 29). It is also concluded that the control change "outside air flow ratio increase" is the most probable cause for the observed fault.

In summary, in Case 1 the fault diagnoses with both the Cosine Similarity method and the Euclidean Distance Similarity method indicate that "outside airflow ratio increase" is the most probable cause for the observed abnormal energy consumption fault. The diagnosis result matches the synthetic setting perfectly.

5.1.2.2 Results of Other Cases

A similar procedure was used in the other nine simulation cases. In Table 14, the first column indicates the ID of simulation test cases, the second column provides the description of synthetic control change leading to the abnormal energy consumption in each simulation case, and the remaining two columns present the fault diagnosis results

with the Cosine Similarity method and the Euclidean Distance Similarity method respectively. It is found that the fault diagnosis results using cosine similarity are coincident with the synthetic control changes setting in each of the ten cases, and the fault diagnosis results using Euclidean distance similarity match the synthetic control changes setting in nine of the ten cases. In Case 5, the Euclidean Distance Similarity method indicates the incorrect fault reason. The correct answer "T_{hl} increase" ranks second in the 17 types of reference control change.

Please see Appendix C for the values of cosine similarity and Euclidean distance in other cases. The reference control changes ordered by the representative similarity in descending order are also shown in Appendix C.

Table 14 Summary of diagnosis results for the Bush Academic Building

Case ID	Synthetic Control Change	Fault Diagnosis Results	
		Cosine Similarity	Euclidean Distance Similarity
1	X_{oa} increase of 3.9%	X_{oa} increase	X_{oa} increase
2	T_{preh} increase of 14°F	T_{preh} increase	T_{preh} increase
3	The heat leakage of preheat coil increase of 40kBtu/hr	The heat leakage of preheat coil increase	The heat leakage of preheat coil increase
4	T_{cl} decrease of 2°F	T_{cl} decrease	T_{cl} decrease
5	**T_{hl} increase of 30°F**	**T_{hl} increase**	**X_{min} increase**
6	The heat leakage of heating coil increase of 60kBtu/hr	The heat leakage of heating coil increase	The heat leakage of heating coil increase
7	X_{min} increase of 2.5%	X_{min} increase	X_{min} increase
8	T_{rh} increase of 3°F	T_{rh} increase	T_{rh} increase
9	T_{rc} decrease of 2.2°F	T_{rc} decrease	T_{rc} decrease
10	Terminal box damper leakage increase of 7.5%	Terminal box damper leakage increase	Terminal box damper leakage increase

5.1.3 Discussion

The fault diagnosis result "X_{oa} increase" using the Euclidean distance similarity is different from the defined synthetic control change in Case 5 "T_{hl} + 30°F". This result is due to the high dependence of Euclidean distance similarity on the reference control change magnitude setting. Euclidean distance similarity is a distance-based similarity measure. It considers only the impact of the distance between the vectors. The control change magnitude has a significant impact on the values of energy consumption. Therefore, the distance between the observed fault signature vector under a synthetic control change and the signature vector under a certain reference control change might be distinctive when the synthetic control change magnitude is far away from the magnitude range stated in the reference library. The largest fault magnitude is +10°F as set for "T_{hl} increase" in the library (Table 11). The signature of "T_{hl}+10°F" is much lower than the signature of "T_{hl}+30°F" (Figure 30).

As shown on Table 15, the representative Euclidean distance similarity in the reference control change category "T_{hl} increase" is 0.6 when the magnitude is "+10°F". It is smaller than the representative Euclidean distance similarity of the reference control change "X_{min} increase" 0.65. However, the representative Euclidean distance similarity reaches 1.0 when a larger T_{hl} magnitude setting is implemented in the reference library (Table 16). "T_{hl} increase" would rank No.1 among all the possible fault reasons after enlarging the magnitude setting of "T_{hl} increase". The influence of the reference control change magnitude setting on the effectiveness of the Cosine Similarity and the Euclidean Distance Similarity methods will be further studied in Section 8.

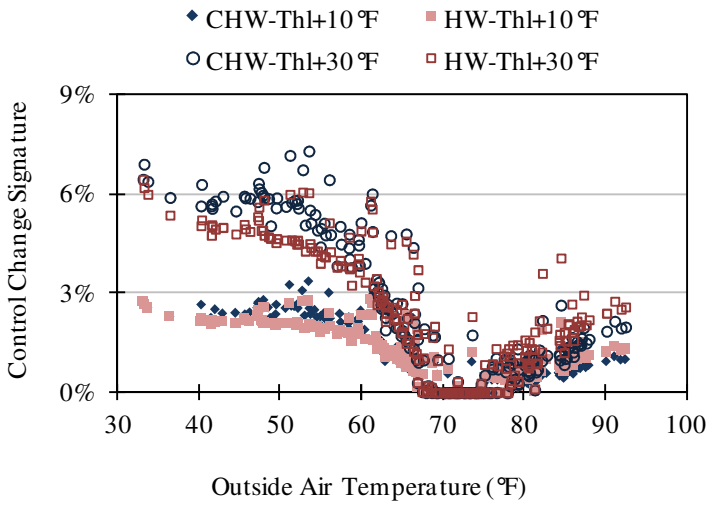

Figure 30 The signature of control changes "T_{hl}+10°F" and "T_{hl}+30°F" as functions of outside air temperature for the Bush Academic Building

Table 15 Euclidean distance similarity between the observed fault signature vector and the reference control change "T_{hl} increase" signature vectors in the simulation Case 5 for the Bush Academic Building

Reference Control Change	Magnitude				
	+2°F	+4°F	+6°F	+8°F	+10°F
T_{hl} increase	0.43	0.48	0.52	0.56	0.60

Table 16 New Euclidean distance similarity values between the observed fault signature vector and the reference control change "T_{hl} increase" signature vectors in the simulation Case 5 for the Bush Academic Building

Reference Control Change	Magnitude				
	+6°F	+12°F	+18°F	+24°F	+30°F
T_{hl} increase	0.52	0.65	0.77	0.89	1.00

5.2 Simulated for SDVAV System

5.2.1 Simulated Data Sets

The simulated test building is the Veterinary Research Building on the Texas A&M University campus in College Station. The majority of the building is served by five SDVAV AHUs. The commissioning was completed in November of 1996. The ABCAT simulations were calibrated to the baseline consumption period from January 1-July 20, 1998. More information about the building is provided in Section 4.

It was assumed one of the following six synthetic control changes happened in the period from January 1 to December 31, 2000:

1. Outside airflow ratio (X_{oa}) increase of 6.5%
2. Preheat temperature (T_{preh}) increase of 3.3°F
3. Cooling coil leaving air temperature (T_{cl}) decrease of 1.5°F
4. Minimum airflow ratio (X_{min}) increase of 4.9%
5. Room heating set-point temperature (T_{rh}) increase of 2.3°F
6. Room cooling set-point temperature (T_{rc}) decrease of 1.7°F

Note: The denominator of X_{oa} and X_{min} is the maximum design airflow volume in the system.

The six synthetic control changes are referred to as six simulation test cases. The parameter changes were chosen so the maximum monthly average cooling and heating consumption deviation caused by each synthetic control change in 2000 was 10% of the yearly average cooling and heating consumption (67.88MMBtu/day) when there was no fault.

The reference control change library containing the known whole building level faults is presented in Table 17. The structure of Table 17is the same as Table 11. In general, in the adopted reference control change library, there are a total of 12 different types of reference control change and each type includes five different severity levels.

Table 17 Reference control change library for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%	-4%	-6%	-8%	-10%	
2	X_{oa} increase	2%	4%	6%	8%	10%	
3	T_{preh} decrease	-2	-4	-6	-8	-10	°F
4	T_{preh} increase	2	4	6	8	10	°F
5	T_{cl} decrease	-2	-4	-6	-8	-10	°F
6	T_{cl} increase	2	4	6	8	10	°F
7	X_{min} decrease	-2%	-4%	-6%	-8%	-10%	
8	X_{min} increase	2%	4%	6%	8%	10%	
9	T_{rc} decrease	-1	-2	-3	-4	-5	°F
10	T_{rc} increase	1	2	3	4	5	°F
11	T_{rh} decrease	-1	-2	-3	-4	-5	°F
12	T_{rh} increase	1	2	3	4	5	°F

Note: The denominator of X_{oa}, and X_{min} is the maximum design airflow volume in the system.

The specific input parameter was varied several times in the calibrated simulation model in ABCAT to simulate various fault sizes. The simulated consumption under the six stated synthetic control changes was treated as the real measured data in the test.

5.2.2 Diagnostic Results with Simulated Test Data

5.2.2.1 Results of Case 3 "Cooling Coil Leaving Air Temperature Decrease of 1.5°F"

Case 3 "Cooling Coil Leaving Air Temperature Decrease of 1.5°F" is listed as an example to illustrate the process of fault diagnosis. The observed fault signature vector and the reference control change signature vectors are derived according to expression (6) in Section 3. The observed fault signature vector components versus outside air temperature are plotted in Figure 31.

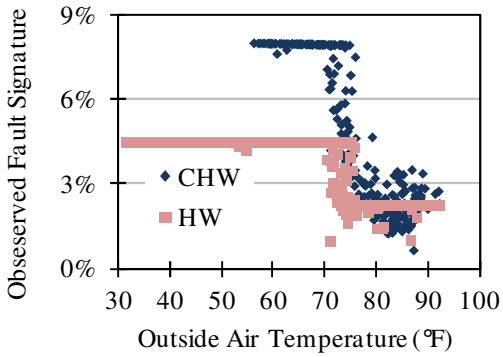

Figure 31 The observed fault signature vector components plotted as a function of outside air temperature for the period from January – December, 2000 for the Veterinary Research Building (Case 3)

The cosine similarity and the Euclidean distance similarity between the observed fault signature vector and each of the reference control change signature vectors are demonstrated in Tables 18 and 19 respectively. They were calculated following expressions (7) and (8) in Section 3. As an example, the cosine similarity value in the cell (Row 1, Column "Magnitude I") of -0.21 suggested that the cosine similarity

between the observed fault signature vector and the "X_{oa} - 2%" signature vector is -0.21. The corresponding values of the different magnitude levels (I, II, III, IV and V) of each control change can be found in Table 17.

The maximum cosine similarity or Euclidean distance similarity from the cases with the identical reference control change are presented in the rightmost column (Column "Max") of Tables 18 and 19. These maximum similarities are called representative similarities of the corresponding reference control changes.

Table 18 Cosine similarity results in Case 3 "Cooling Coil Leaving Air Temperature Decrease of 1.5°F" for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	-0.21	-0.21	-0.21	-0.21	-0.21	-0.21
2	X_{oa} increase	0.21	0.21	0.21	0.21	0.21	0.21
3	T_{preh} decrease	-0.37	-0.36	-0.35	-0.34	-0.34	-0.34
4	T_{preh} increase	0.42	0.43	0.46	0.48	0.51	0.51
5	T_{cl} decrease	1.00	0.99	0.99	0.98	0.97	1.00
6	T_{cl} increase	-0.96	-0.89	-0.79	-0.65	-0.49	-0.49
7	X_{min} decrease	-0.85	-0.84	-0.84	-0.83	-0.82	-0.82
8	X_{min} increase	0.87	0.88	0.88	0.87	0.87	0.88
9	T_{rc} decrease	-0.03	0.11	0.21	0.28	0.33	0.33
10	T_{rc} increase	0.19	0.24	0.28	0.32	0.38	0.38
11	T_{rh} decrease	-0.41	-0.40	-0.40	-0.40	-0.40	-0.40
12	T_{rh} increase	0.41	0.40	0.40	0.40	0.41	0.41

Figures 32 and 33 indicate the reference control changes ordered by the representative similarity in descending order. Each bar in the figure represents the representative similarity between the observed fault and a reference control change. The

reference control change IDs shown on the X axis correspond to the IDs listed in Table 17. In the figures, the leftmost bar has the largest similarity value while the rightmost bar has the smallest similarity value.

Table 19 Euclidean distance similarity results in Case 3 "Cooling Coil Leaving Air Temperature Decrease of 1.5°F" for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.16	0.12	0.09	0.06	0.04	0.16
2	X_{oa} increase	0.20	0.17	0.13	0.10	0.07	0.20
3	T_{preh} decrease	0.15	0.12	0.09	0.08	0.08	0.15
4	T_{preh} increase	0.23	0.22	0.17	0.12	0.07	0.23
5	T_{cl} decrease	0.57	0.05	0.00	0.00	0.00	0.57
6	T_{cl} increase	0.03	0.01	0.00	0.00	0.00	0.03
7	X_{min} decrease	0.10	0.06	0.04	0.03	0.02	0.10
8	X_{min} increase	0.35	0.53	0.35	0.17	0.08	0.53
9	T_{rc} decrease	0.13	0.05	0.01	0.00	0.00	0.13
10	T_{rc} increase	0.20	0.12	0.06	0.03	0.01	0.20
11	T_{rh} decrease	0.15	0.11	0.08	0.07	0.05	0.15
12	T_{rh} increase	0.23	0.21	0.17	0.13	0.10	0.23

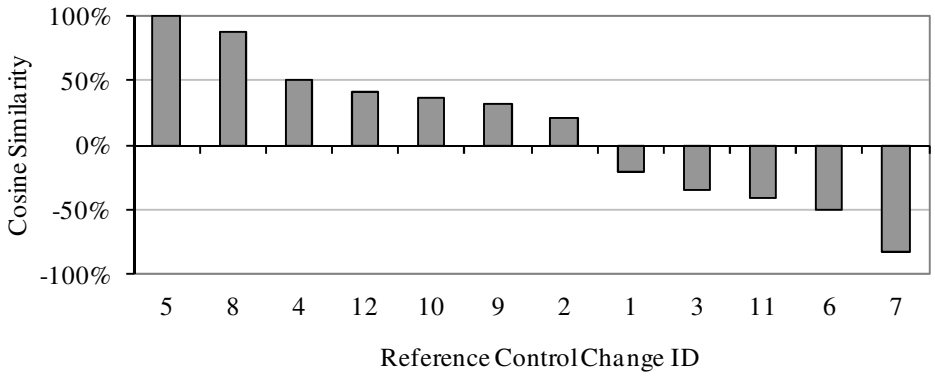

Figure 32 Representative cosine similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 3)

A larger similarity value corresponds to a higher coincidence between two vectors. The cosine similarity between the observed fault and the reference control change (ID 5) is 100% and ranks No.1 among the 12 types of reference control changes (Figure 32). This suggests that the symptom of the signature vector of reference control change "cooling coil leaving air temperature decrease" (ID 5) is most similar to the symptom of the observed fault signature vector. Therefore, the observed abnormal consumption is most likely due to a decrease of the cooling coil leaving air temperature.

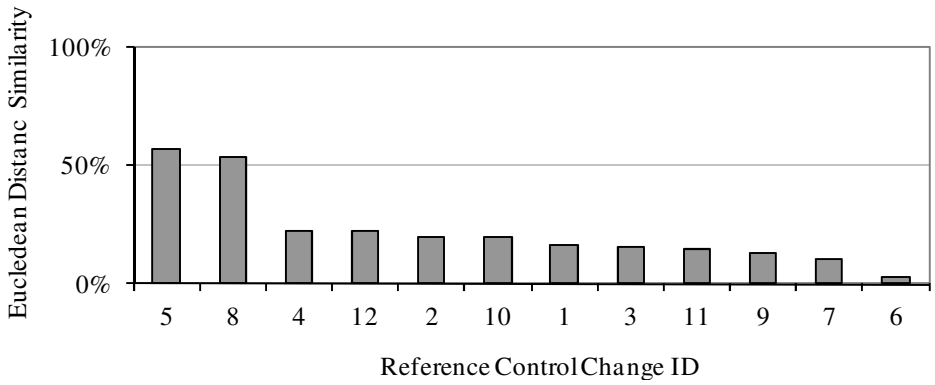

Figure 33 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 3)

Likewise, the Euclidean distance similarity between the observed fault and the reference control change (ID 5) is 57% and is No.1 among the 12 types of reference control changes (Figure 33). It also infers that "cooling coil leaving air temperature decrease" has the highest probability of causing the observed fault.

In short, in Case 3, the fault diagnosis with both the Cosine Similarity method and the Euclidean Distance Similarity method indicates "cooling coil leaving air

temperature decrease" is the most probable cause of the observed abnormal energy consumption fault. The diagnosis results both correctly identified the synthetic faults.

5.2.2.2 Results of Other Cases

A similar procedure was applied in the other five simulation cases for the Veterinary Research Building. In Table 20, the first column indicates the ID of simulation test cases, the second column describes the synthetic control change in each simulation case, and the third and fourth columns show the fault diagnosis results with the Cosine Similarity and the Euclidean Distance Similarity methods respectively. It is found that the fault diagnosis results with both cosine similarity and Euclidean distance similarity are consistent with the synthetic control changes setting in all six cases.

Table 20 Summary of the diagnosis results for the Veterinary Research Building

Case ID	Synthetic Control Change	Fault Diagnosis Results	
		Cosine Similarity	Euclidean Distance Similarity
1	X_{oa} increase of 6.5%	X_{oa} increase	X_{oa} increase
2	T_{preh} increase of 3.3°F	T_{preh} increase	T_{preh} increase
3	T_{cl} decrease of 1.5°F	T_{cl} decrease	T_{cl} decrease
4	X_{min} increase of 4.9%	X_{min} increase	X_{min} increase
5	T_{rh} increase of 2.3°F	T_{rh} increase	T_{rh} increase
6	T_{rc} decrease of 1.7°F	T_{rc} decrease	T_{rc} decrease

Please see Appendix C for the values of the cosine similarity and Euclidean distance similarity in each case. The reference control changes sorted by the representative similarity in descending order are also recorded in Appendix C.

5.3 Summary

This section carries out simulation tests on the two proposed whole building fault diagnosis methods – the Cosine Similarity method and the Euclidean Distance Similarity method. The fault diagnosis capability of the two methods was examined in the test. The test buildings included a building with DDVAV systems and a building with SDVAV systems. Synthetic control changes were assumed to happen and persist during a fault period lasting for one year.

For the DDVAV system building, the fault diagnosis results with the Cosine Similarity method were consistent with the synthetic control changes in each of the ten simulation cases; the fault diagnosis results with the Euclidean Distance Similarity method matched the synthetic control changes in nine of the ten simulation cases. The indicated fault reason using the Euclidean Distance Similarity method was different from the synthetic setting in the simulation case "T_{hl}+30°F". Analysis indicated that this false result is caused by the improper magnitude level setting in the reference control change library. It seems that the Euclidean Distance Similarity method is more sensitive to the reference control change magnitude setting than the Cosine Similarity method. Further investigation about this statement will be performed in Section 8 of this dissertation. For the SDVAV system building, the fault diagnosis results of both methods matched the stated synthetic control changes perfectly. The simulation test results suggest that the Cosine Similarity method and the Euclidean Distance Similarity method are both promising techniques for whole building fault diagnosis.

6. ANALYSIS OF WHOLE BUILDING FAULT DETECTION AND DIAGNOSIS APPROACHES WITH FIELD TEST

In the previous two sections, simulation tests were conducted for the proposed whole building fault detection and diagnosis (FDD) approaches to illustrate their FDD capabilities. A field test would provide opportunities to further test their robustness in practice. The performance of the whole building FDD approaches developed in a field test is presented in this section. Descriptions of the test buildings, field data sets, as well as detection and diagnosis results are given.

6.1 Field Test of Proposed Fault Detection Approaches

In order to further test the fault detection capabilities of the DET-Date and DET-Toa methods, a multiple building retrospective test was performed. This group of five buildings on the TAMU campus had previously been studied in a commissioning persistence study (for the years from 1996 to 2000). The application of the DET-Date and DET-Toa methods on two additional test cases are also presented. The fault detection results of these two buildings form the foundation of the following fault diagnosis test cases. General building information, fault detection procedure, and results are given below.

6.1.1 Retrospective Test Cases

6.1.1.1 Building Information

6.1.1.1.1 Kleberg Center

The Kleberg Center contains 165,000 ft^2 of conditioned space consisting of offices, classrooms and laboratories. The building has two large SDVAV AHUs with

large fresh air requirements to maintain proper makeup air for significant laboratory exhaust flows. Additionally, two smaller Single-Duct Constant Volume(SDCV) AHUs condition some lecture/teaching rooms. The building has temperature economizer control. The commissioning work on this building was completed in August 1996. The ABCAT simulation was calibrated to the baseline consumption period from November 1996 to July 1997.

6.1.1.1.2 Veterinary Research Building

The information for this building has been given in Section 4.

6.1.1.1.3 Wehner Building

The Wehner Building is a four-story building with 192,000 ft^2 of conditioned space consisting of offices, classrooms, and computer labs. The building has six DDVAV AHUs that serve the second to fourth floors, each with a separate constant volume outside air pretreat unit, and three SDVAV AHUs serving the first floor. The commissioning work on this building was completed in December 1996. The ABCAT simulation was calibrated to the baseline consumption period from January to July 1997.

6.1.1.1.4 EOM Building

The information for this building has been provided in Section 4.

6.1.1.1.5 Harrington Tower

Harrington Tower is an eight story building with 131,000 ft^2 of conditioned space. The building is comprised of multiple offices, classrooms, and computer laboratory spaces. The majority of the building (floors 2–8) is served by a single DDVAV AHU with an economizer. The 1st floor is served by three separate SDVAV

AHUs. The building has temperature economizer control. The building was commissioned in August of 1996. The ABCAT simulation was calibrated to the baseline consumption period from August 1996 to August 1997.

6.1.1.2 Detection Results with Field Data – DET-Date Method

The following procedure was applied to each building for fault detection with the DET-Date method:

1. The simulation model of ABCAT was calibrated based on the energy data in the baseline period.
2. The calibrated simulation model predicted the cooling and heating consumption in the period from the first day post-baseline to the end of 2000.
3. Calculated daily CHW and HW fault indexes according to expression (4) in Section 3.
4. Drew the DET-Date plot

The DET-Date plots of the five buildings are shown on Figures 34-38 respectively. According to the fault detection standard, a fault is identified if the deviation between the measured and the simulated consumption is greater than one standard deviation in the baseline period and persists for at least 20 days. The detection results are summarized in Table 21.

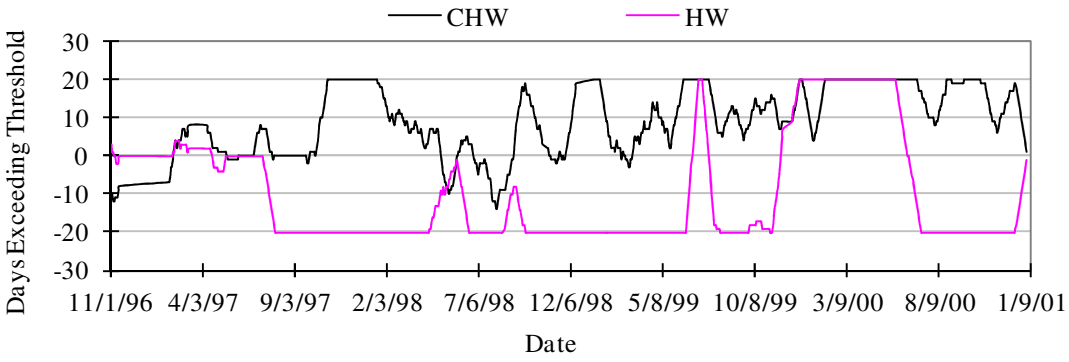

Figure 34 Days Exceeding Threshold-Date plots from November 1, 1996 to December 31, 2000 for the Kleberg Center

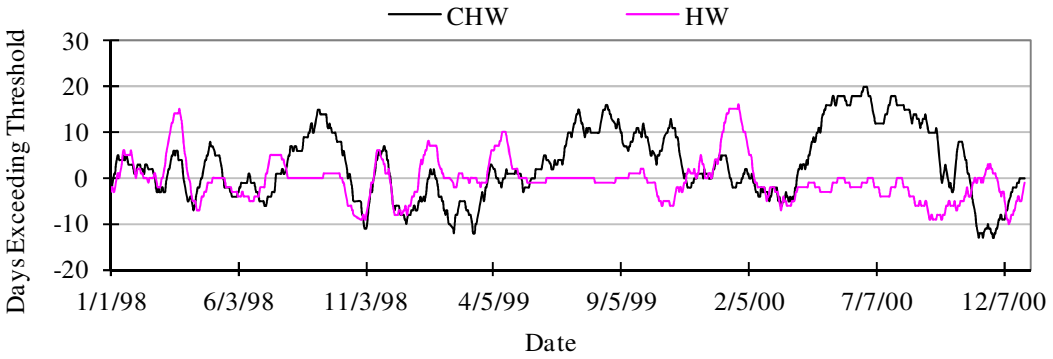

Figure 35 Days Exceeding Threshold-Date plots from March 19, 1997 to December 31, 2000 for the Veterinary Research Building

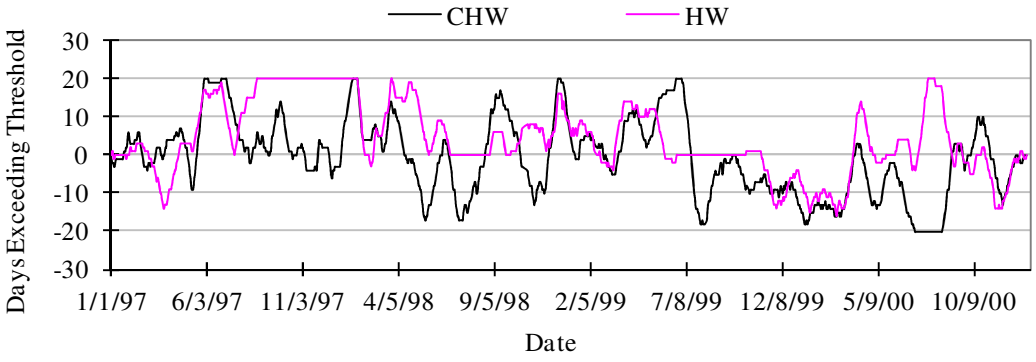

Figure 36 Days Exceeding Threshold-Date plots from January 1, 1997 to December 31, 2000 for the Wehner Building

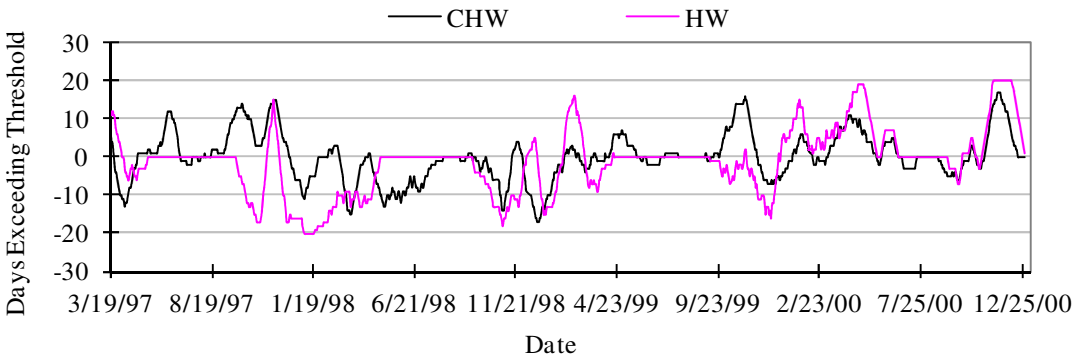

Figure 37 Days Exceeding Threshold-Date plots from March 19, 1997 to December 31, 2000 for the EOM Building

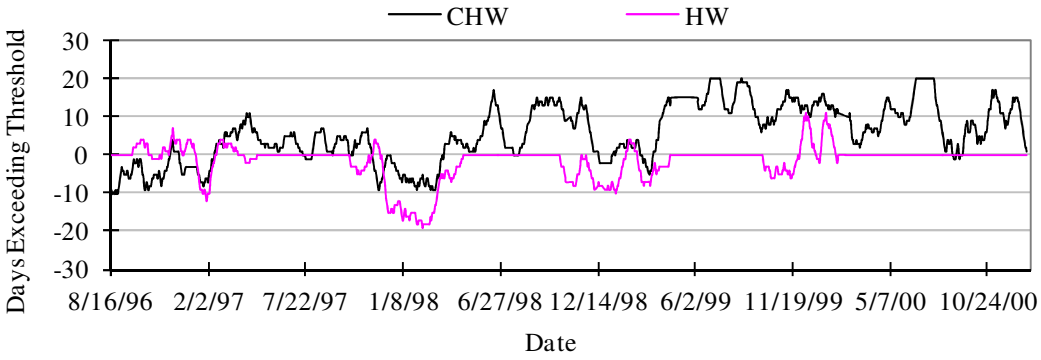

Figure 38 Days Exceeding Threshold-Date plots from August 16, 1996 to December 31, 2000 for the Harrington Tower

Table 21 Summary of detected abnormal energy consumption faults by the DET-Date method in the field test

ID	Consumption	Fault Identification Day	Duration	Zm
Kleberg Center				
HW 1	Decrease	8/20/1997	8/1/1997-5/2/1998	-2.8
HW 2	Decrease	7/8/1998	6/19/1998-9/4/1998	-1.9
HW 3	Decrease	10/10/1998	9/21/1998-7/4/1999	-3.2
HW 4	Decrease	8/29/1999	8/10/1999-10/16/1999	-2.5
HW 5	Decrease	11/20/1999	11/1/1999-12/7/1999	-1.4
HW 6	Increase	7/24/1999	7/5/1999-7/30/1999	4.6
HW 7	Increase	1/8/2000	12/20/1999-6/15/2000	5.7

Table 21 Continued.

ID	Consumption	Fault Identification Day	Duration	Zm
Kleberg Center				
HW 8	Decrease	7/28/2000	7/9/2000-12/31/2000	-3.2
CHW 1	Increase	11/15/1997	10/27/1997-2/9/1998	8.5
CHW 2	Increase	1/31/1999	1/12/1999-2/19/1999	4.7
CHW 3	Increase	6/29/1999	6/10/1999-8/12/1999	4.0
CHW 4	Increase	1/8/2000	12/20/1999-1/12/2000	7.4
CHW 5	Increase	2/19/2000	1/31/2000-7/21/2000	4.6
CHW 6	Increase	9/8/2000	8/20/2000-9/16/2000	2.7
CHW 7	Increase	10/7/2000	9/18/2000-11/3/2000	5.6
Veterinary Research Building				
CHW 1	Increase	7/10/2000	6/21/2000-7/14/2000	1.9
Wehner Building				
HW 1	Increase	9/10/1997	08/22/1997 - 02/18/1998	7.2
HW 2	Increase	4/12/1998	3/24/1998-4/14/1998	4.0
HW 3	Increase	8/15/2000	07/27/2000 - 08/25/2000	3.2
CHW 1	Increase	6/18/1997	5/30/1997-6/24/1997	2.8
CHW 2	Increase	7/15/1997	6/26/1997-7/24/1997	3.4
CHW 3	Increase	2/10/1998	1/22/1998-2/18/1998	4.3
CHW 4	Increase	1/3/1999	12/15/1998-1/7/1999	5.0
CHW 5	Increase	7/10/1999	6/21/1999-7/19/1999	2.5
CHW 6	Decrease	7/25/2000	7/6/2000 - 9/6/2000	-3.3
EOM Building				
HW 1	Decrease	1/25/1998	1/6/1998-2/9/1998	-6.0
HW 2	Increase	12/3/2000	11/14/2000-12/31/2000	7.6
Harrington Tower				
CHW 1	Increase	7/17/1999	6/28/1999-8/3/1999	1.6
CHW 2	Increase	9/9/1999	8/21/1999-9/14/1999	1.6
CHW 3	Increase	7/10/2000	6/21/2000-8/10/2000	5.4

Note: Consumption “Increase/Decrease” means the measured consumption is higher/lower than the fault-free simulated consumption. “Z_m” is the modified z-score, which is the ratio of the average daily CHW/HW increase/decrease during 20 fault-flag days to the standard deviation of CHW/HW residuals during the calibration baseline period

Eight abnormal HW consumption faults and seven abnormal CHW consumption faults were detected in the Kleberg Center (Table 21). In HW faults 1-5 and 8, the measured HW consumption data were less than the simulated data. It was found that most of the measured HW data during these periods were zero. Facility personnel verified that the HW consumption meter had problems during the periods from mid 1997 to late 1999 and from June 2000 to December 2000. The reasons for faults CHW 1 and CHW 2 are unknown as no specific details of the control changes are available. Both CHW and HW consumption were higher than expected in July 1999 (HW 6 and CHW 3) and from the end of 1999 through the summer of 2000 (HW 7, CHW 4, and CHW 5), which were consistent with problems the Kleberg Center experienced as documented in Chen et al (2002). Chen et al (2002) reported that the Kleberg Center experienced several problems after April 1999. Leaking chilled water valves resulted in a lower air discharge temperature and more terminal reheat. They also caused the preheat coil to remain on, regardless of the outside air temperature. Failed CO_2 sensors and building static pressure sensors resulted in excessive outside airflow. Leaking damper actuators in some of the VAV boxes resulted in a higher minimum airflow ratio. There were also other problems; for example, two chilled water pump variable frequency drives were by-passed to full speed, increasing chilled water and hot water consumption due to high pressures in the water loops. These problems could also account for faults CHW 6 and 7.

Four HW-increase faults, one HW-decrease fault, nine CHW-increase faults and one CHW- decrease fault for the other four buildings are detected by the DET-Date method (Table 21). The HW fault 1 of the Veterinary Research Building appeared to be

linked to a new constant 56ºF set-point of the cooling coil discharge temperature in 2000 (Cho et al 2002). The causes of the other detected faults in Table 21 can't be verified because no specific details of control changes are available.

In summary, 30 abnormal energy consumption faults were detected in 15 building-years of consumption data with the DET-Date method. The absolute magnitudes of these faults, calculated as minimum, maximum, and median ratios to the standard deviation during the calibration baseline period, were 1.6/8.5/4.1 for the 17 CHW faults and 1.4/7.6/3.2 for the 13 HW faults.

6.1.1.3 Detection Results with Field Data – DET-Toa Method

In order to compare the fault detection capabilities of the DET-Toa method to the capabilities of the DET-Date method, the earliest abnormal CHW/HW energy consumption faults in the five buildings detected by the DET-Date method as shown in Table 21 were chosen as the foundation of the comparison. The DET-Toa method was used to examine the fault existence during the period from the first day post-baseline to the fault identification day in which the faults were detected by the DET-Date method. This period is called the investigated period and is recorded in Table 22.

Table 22 Comparison of the fault identification day of the DET-Toa and the DET-Date methods

ID	Consumption	Fault Identification Day		Investigated Period
		DET-Date	DET-Toa	
Kleberg Center				
HW 1	Decrease	8/20/1997	8/20/1997	8/1/1997-8/20/1997
CHW 1	Increase	11/15/1997	11/15/1997	8/1/1997-11/15/1997
Veterinary Research Building				
CHW 1	**Increase**	**7/10/2000**	**6/15/2000**	**7/12/1999-7/10/2000**
Wehner Building				
HW 1	**Increase**	**9/10/1997**	**9/4/1997**	**5/30/1997-9/10/1997**
CHW 1	Increase	6/18/1997	6/18/1997	5/30/1997-6/18/1997
Eller O&M Building				
HW 1	**Decrease**	**1/25/1998**	**11/10/1997**	**9/1/1997-1/25/1998**
Harrington Tower				
CHW 1	Increase	7/17/1999	7/17/1999	7/16/1998-7/17/1999

The superior performance of the DET-Toa method could be validated if it detected the faults earlier than the DET-Date method in the investigated period. The result of the comparison is presented in Table 22. It shows that compared with the DET-Date method, the DET-Toa method identified the abnormal energy consumption faults earlier in three of the seven cases and identified the faults on the same day in the remaining four cases.

6.1.2 Retrospective Test Case – Sbisa Dining Hall

6.1.2.1 Building Information

The Sbisa Dining Hall is an 82,000 ft^2 single story building with a partial basement on the campus of Texas A&M University in College Station. Its primary function is as a dining facility. The main AHUs are single duct constant volume (SDCV)

AHUs with terminal reheat boxes. Three constant volume dedicated outside air handling units (OAHUs) provide pretreated makeup air for the majority of the AHUs. Thermal energy is supplied to the building in the form of hot and chilled water from the central utility plant. The ABCAT simulation was calibrated to the baseline consumption period from February 2, 2004 to December 31, 2004. The DET-Date and the DET-Toa methods were used to identify the faults in the period from January 1, 2006 to June 4, 2006.

6.1.2.2 Detection Results with Field Data

The DET-Date plot is shown in Figure 39, and the detection results of the DET-Date and DET-Toa methods are summarized in Table 23. Curtin (2007) indicates that in 2006 the discharge air temperature in two of the three OAHUs in the building was exceptionally low, which would lead to low HW energy consumption in winter and high CHW consumption in summer. The unusual energy performance was first caught by the DET-Toa method. It detected abnormally high CHW energy consumption on May 20, 2006. The DET-Date method detected the abnormally high CHW energy consumption on June 1, 2006 (Table 23). In summary, the DET-Toa method identified the abnormal energy consumption fault 10 days earlier than the DET-Date method in 2006 for the Sbisa Dining Hall.

Table 23 Comparison of fault detection results for the Sbisa Dining Hall

Faults	Fault Identification Day	
	DET-Toa	DET-Date
CHW Increase	5/20/2006	6/1/2006

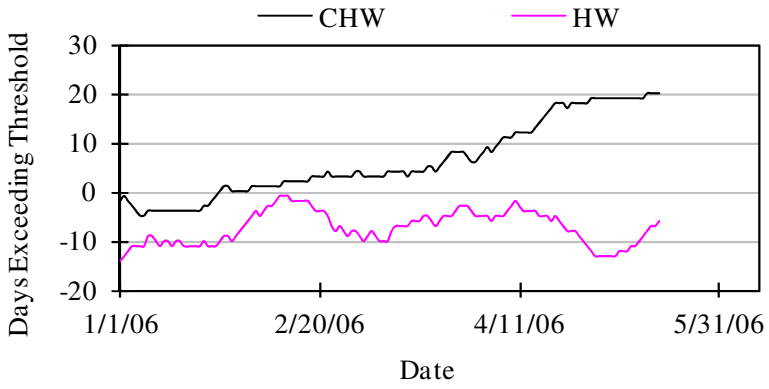

Figure 39 Days Exceeding Threshold-Date plot of January 1–June 4, 2006 for the Sbisa Dining Hall

6.1.3 Live Test Case – Bush Academic Building

6.1.3.1 Building Information

The building information is listed in Section 5. The DET-Date and DET-Toa methods were applied to investigate the fault appearance in the weekday period from November 1, 2008 to June 29, 2009.

6.1.3.2 Detection Results with Field Data

The DET-Date plot is shown in Figure 40. The detection results of the DET-Date and DET-Toa methods are summarized in Table 6.4. Claridge et al. (2009) shows that there was a preheat valve leaking by on a pre-treat unit during the investigated period. This fault would cause the increase of CHW and HW energy consumption in the summer. The DET-Toa method detected abnormally high HW energy consumption on April 23, 2009 and abnormally high CHW energy consumption on May 6, 2009. The DET-Date method detected only abnormally high CHW energy consumption on May 15,

2009 (Table 24). We can infer that the DET-Toa method identified the abnormal energy consumption about three weeks earlier than the DET-Date method.

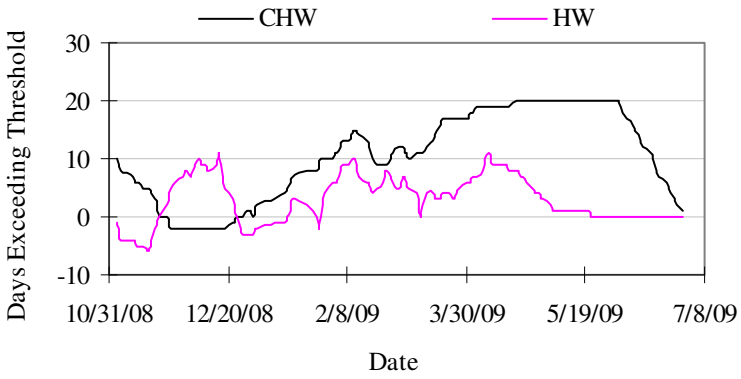

Figure 40 Days Exceeding Threshold-Date plot from November 1, 2008 to June 29, 2009 for the Bush Academic Building

Table 24 Comparison of fault detection results for the Bush Academic Building

Faults	Fault Identification Day	
	DET-Toa	DET-Date
CHW Increase	5/6/2009	5/15/2009
HW Increase	4/23/2009	

6.1.4 Summary

The fault detection capabilities of the DET-Date and DET-Toa methods were validated with a field test. In the test cases, 34 abnormal energy consumption faults were detected with the DET-Date method. The DET-Toa method had better performance in the field test. It identified five faults earlier than the DET-Date method in the

comparison of nine existing faults in the test cases. In the rest four cases, the fault identification day of the two methods are on the same day.

6.2 Field Test for Fault Diagnosis Approaches

Data from the Sbisa Dining Hall and the Bush Academic Building were used to test and validate the Cosine Similarity and Euclidean Distance Similarity methods as described in Section 3. Building information, diagnosis process, and diagnosis results are illustrated below.

6.2.1 Sbisa Dining Hall

6.2.1.1 Field Data Sets

The building information is described in previous section. As mentioned in Section 3, the application of the proposed fault diagnosis approaches requires the condition that some fault detection mechanism has already noticed that an abnormal consumption fault is present and has persisted for a certain time. The DET-Toa method had detected especially high CHW consumption on May 20, 2006 for the Sbisa Dining Hall in previous section. The period from April 29-June 4, 2006 is referred to as the fault period. April 29, 2006 is the earliest day in the 20 fault-flag days of the DET-Toa method, and June 4, 2006 is the last day in the investigated period of fault detection. The maximum monthly average cooling and heating consumption increase in the fault period was 15% of the average energy consumption in the baseline period (76.8MMBtu/day) (Figure 41).

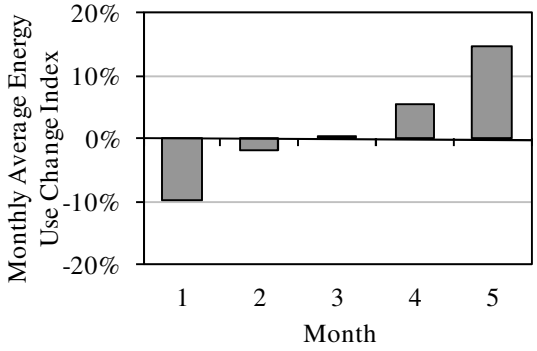

Figure 41 Monthly average energy consumption changes of January 1 – June 4, 2006 for the Sbisa Dining Hall

6.2.1.2 Diagnostic Results with Field Data

The following procedure was implemented in the fault period for the fault diagnosis using the Cosine Similarity and Euclidean Distance Similarity methods:

1. Determined a reference control change library including the known whole building level faults.
2. Used ABCAT to produce the expected cooling and heating consumption when there was a known control change from the reference library persisting during the fault period.
3. Generated the observed fault signature vector and reference control change signature vectors according to expression (6) in Section 3.
4. Calculated the similarity values between the observed fault signature vector and each of the reference control change signature vectors. Cosine similarity was calculated when the Cosine Similarity method was used, and Euclidean

distance similarity was calculated when the Euclidean Distance Similarity method was applied.

5. Selected the representative similarity for each type of the reference control change.
6. Sorted different types of reference control change by representative similarity in descending order. The ranking indicated the probability that the reference control change is the cause of the observed fault.

Table 25 defines 12 different types of reference control change with five levels (I-V) of magnitude. The magnitudes III, IV, and V of "X_{oa} decrease" are blank; they would have negative values since the original input parameter was 28% in the calibrated simulation model.

Table 25 Reference control change library for the Sbisa Dining Hall

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	Outside airflow ratio (X_{oa}) decrease	-10%	-20%				
2	Outside airflow ratio increase	10%	20%	30%	40%	50%	
3	Outside air precool temperature (T_{prec}) decrease	-2	-4	-6	-8	-10	°F
4	Outside air precool temperature increase	2	4	6	8	10	°F
5	Cooling coil leaving temperature (T_{cl}) decrease	-2	-4	-6	-8	-10	°F
6	Cooling coil leaving temperature increase	2	4	6	8	10	°F
7	Maximum airflow ratio (X_{max}) decrease	-10%	-20%	-30%	-40%	-50%	
8	Maximum airflow ratio increase	10%	20%	30%	40%	50%	
9	Room cooling set-point temperature (T_{rc}) decrease	-2	-4	-6	-8	-10	°F
10	Room cooling set-point temperature increase	2	4	6	8	10	°F
11	Room heating set-point temperature (T_{rh}) decrease	-2	-4	-6	-8	-10	°F
12	Room heating set-point temperature increase	2	4	6	8	10	°F

Note: The denominator of X_{oa} and X_{max} is the maximum design airflow volume in the system.

The observed fault signature vector components are plotted versus outside air temperature in Figure 42. The cosine similarity and Euclidean distance similarity between the observed fault signature vector and each of the reference control change signature vectors are summarized in Tables 26 and 27 respectively.

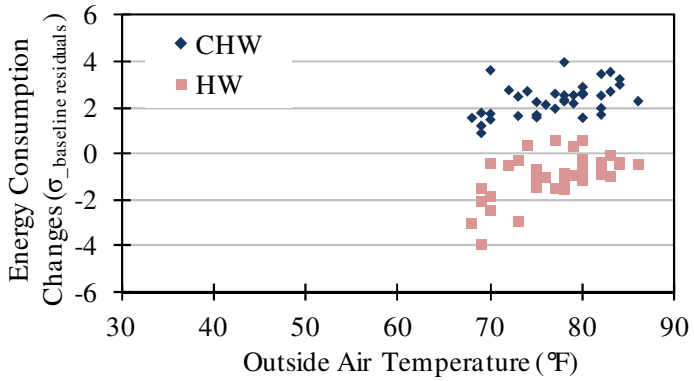

Figure 42 The observed fault signature vector components plotted as a function of outside air temperature in the period from 4/29-6/4/2006 for the Sbisa Dining Hall

Table 26 Cosine similarity results for the Sbisa Dining Hall

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	-0.86	-0.86				-0.86
2	X_{oa} increase	0.86	0.86	0.86	0.86	0.86	0.86
3	T_{prec} decrease	0.90	0.92	0.92	0.92	0.92	0.92
4	T_{prec} increase	0.13	0.21	0.23	0.24	0.29	0.29
5	T_{cl} decrease	0.60	0.60	0.59	0.59	0.59	0.60
6	T_{cl} increase	-0.50	-0.44	-0.52	-0.55	-0.61	-0.44
7	X_{max} decrease	-0.51	-0.51	-0.53	-0.55	-0.57	-0.51
8	X_{max} increase	0.51	0.51	0.51	0.51	0.51	0.51
9	T_{rc} decrease	-0.28	-0.30	-0.34	-0.39	-0.45	-0.28
10	T_{rc} increase	0.29	0.29	0.29	0.30	0.30	0.30
11	T_{rh} decrease	0.00	0.00	0.00	0.00	0.00	0.00
12	T_{rh} increase	0.00	0.00	0.00	0.00	0.00	0.00

In Figure 43, the first bar on the left has the highest cosine similarity, and the last bar on the right has the lowest cosine similarity. The reference control change IDs on the X axis correspond to the IDs in Table 25. Figure 43 shows that the control changes "T_{prec} decrease" and "X_{oa} decrease" have the largest and the smallest cosine similarity values respectively among the 12 types of reference control change. Therefore, the observed abnormal consumption is most probably due to a decrease of the outside air precool temperature and is least likely to be caused by a decrease of the outside airflow ratio.

Table 27 Euclidean distance similarity results for the Sbisa Dining Hall

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.12	0.07				0.12
2	X_{oa} increase	0.33	0.40	0.28	0.16	0.09	0.40
3	T_{prec} decrease	0.23	0.26	0.30	0.33	0.37	0.37
4	T_{prec} increase	0.21	0.21	0.21	0.21	0.22	0.22
5	T_{cl} decrease	0.24	0.23	0.12	0.05	0.02	0.24
6	T_{cl} increase	0.15	0.14	0.04	0.02	0.01	0.15
7	X_{max} decrease	0.16	0.10	0.06	0.03	0.02	0.16
8	X_{max} increase	0.22	0.22	0.20	0.13	0.08	0.22
9	T_{rc} decrease	0.17	0.10	0.06	0.03	0.02	0.17
10	T_{rc} increase	0.18	0.15	0.12	0.08	0.05	0.18
11	T_{rh} decrease	0.00	0.00	0.00	0.00	0.00	0.00
12	T_{rh} increase	0.00	0.00	0.00	0.00	0.00	0.00

The ranking of the reference control changes based on the results of Euclidean distance similarity concludes that the increase of the outside airflow ratio is the most possible reason for the observed fault (Figure 44). This result disagrees with the real situation. The correct answer ranks second in the 12 types of reference control changes.

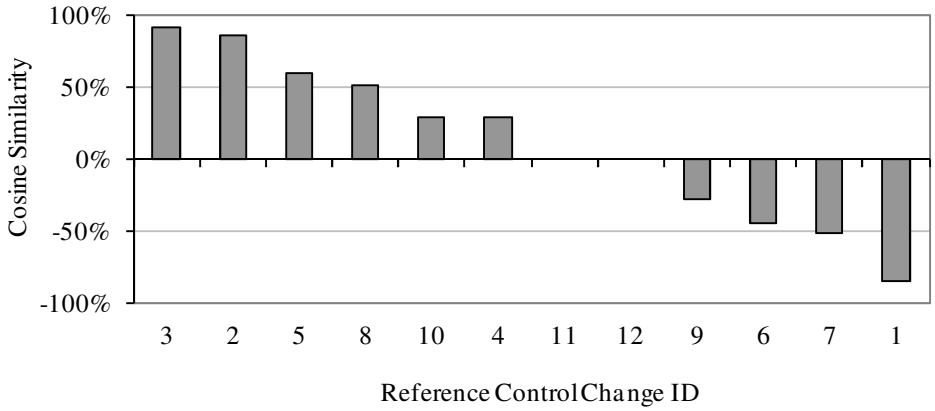

Figure 43 Representative cosine similarity values for different reference control changes sorted in descending order for the Sbisa Dining Hall

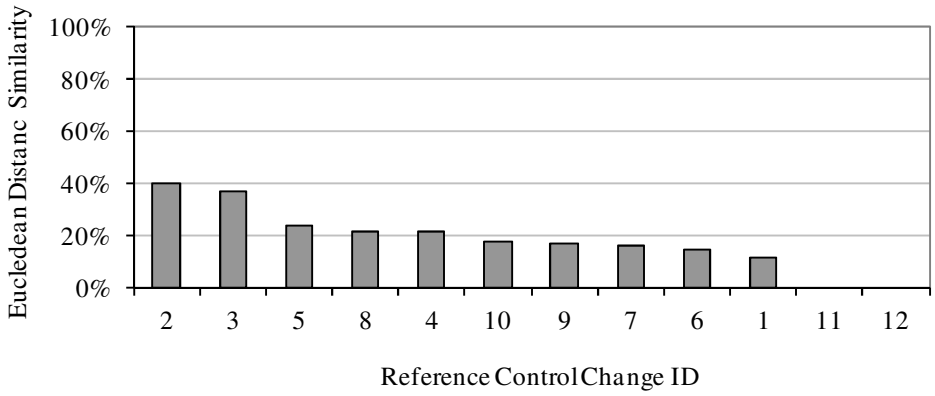

Figure 44 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Sbisa Dining Hall

Curtin (2007) reported that investigation into trended control data points had led to the discovery of exceptionally low discharge air temperature in two of the three OAHUs in the buildings. It is obvious that the diagnosis result with either cosine similarity or Euclidean distance similarity method is consistent with the field investigation conclusion.

6.2.2 Bush Academic Building

6.2.2.1 Building Information

The building information is described in Section 5. The DET-Toa method had detected remarkably high HW consumption on April 23, 2009 for the Bush Academic Building. The weekday period from November 04, 2008 to June 29, 2009 is referred to as the fault period below, where November 04, 2008 is the earliest day in the 20 fault-flag days of the DET-Toa method, and June 29, 2009 is the last day in the investigated period of fault detection. The maximum monthly average cooling and heating consumption increase in the fault period was 30% of the average energy consumption in the baseline period (21.7MMBtu/day) (Figure 45).

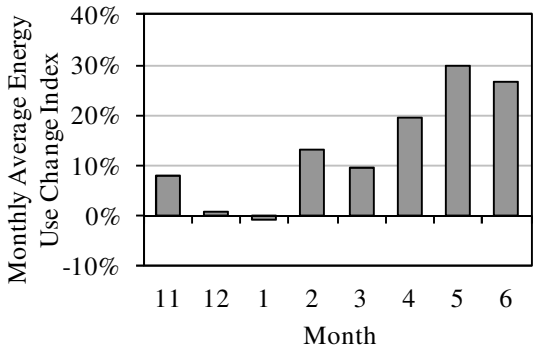

Figure 45 Monthly average energy consumption changes from November 4, 2008 to June 29, 2009 for the Bush Academic Building

6.2.2.2 Diagnostic Results with Field Data

The same procedure as described in Section 6.1.2.2 was implemented for the fault diagnosis in the fault period with the Cosine Similarity and Euclidean Distance

Similarity methods. Seventeen different types of reference control change with five levels of magnitude are presented in Table 28. The observed fault signature vector components are plotted versus outside air temperature in Figure 46. The cosine similarity and Euclidean distance similarity between the observed fault signature vector and each of the reference control change signature vectors are provided in Tables 29 and 30 respectively.

Table 28 Reference control change library for the Bush Academic Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%	-4%	-6%	-8%	-10%	
2	X_{oa} increase	2%	4%	6%	8%	10%	
3	T_{preh} decrease	-3	-6	-9	-12	-15	°F
4	T_{preh} increase	3	6	9	12	15	°F
5	PreHL increase	10	20	30	40	50	kBtu/hr
6	T_{cl} decrease	-2	-4	-6	-8	-10	°F
7	T_{cl} increase	2	4	6	8	10	°F
8	T_{hl} decrease	-2	-4	-6	-8	-10	°F
9	T_{hl} increase	2	4	6	8	10	°F
10	HL increase	10	20	30	40	50	kBtu/hr
11	X_{min} decrease	-2%	-4%	-6%	-8%	-10%	
12	X_{min} increase	2%	4%	6%	8%	10%	
13	T_{rc} decrease	-1	-2	-3	-4	-5	°F
14	T_{rc} increase	1	2	3	4	5	°F
15	T_{rh} decrease	-1	-2	-3	-4	-5	°F
16	T_{rh} increase	1	2	3	4	5	°F
17	TDL increase	2%	4%	6%	8%	10%	

Note: The denominator of outside airflow ratio and maximum airflow ratio is the maximum design airflow volume in the system.

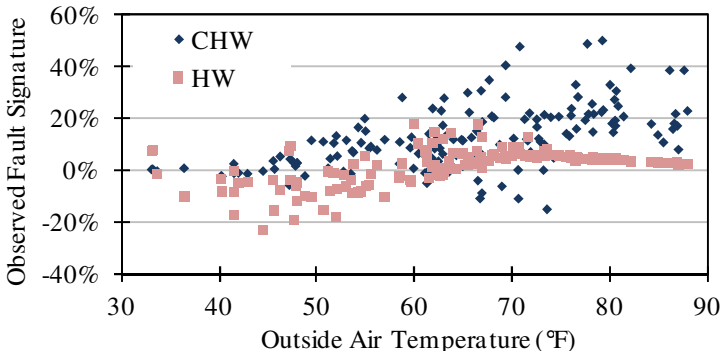

Figure 46 The observed fault signature vector components plotted as a function of outside air temperature in the weekday period from 11/1/2008 - 6/30/2009 for the Bush Academic Building

Table 29 Cosine similarity results for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	-0.19	-0.19	-0.19	-0.19	-0.18	-0.18
2	X_{oa} increase	0.22	0.23	0.24	0.27	0.29	0.29
3	T_{preh} decrease	0.06	0.07	0.07	0.07	0.07	0.07
4	T_{preh} increase	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06
5	PreHL increase	0.65	0.65	0.64	0.64	0.63	0.65
6	T_{cl} decrease	0.39	0.40	0.42	0.43	0.45	0.45
7	T_{cl} increase	-0.42	-0.30	-0.13	0.01	0.12	0.12
8	T_{hl} decrease	-0.24	-0.20	-0.16	-0.11	-0.07	-0.07
9	T_{hl} increase	0.21	0.21	0.20	0.19	0.19	0.21
10	HL increase	0.00	0.10	0.53	0.60	0.64	0.64
11	X_{min} decrease	-0.28	-0.23	-0.18	-0.14	-0.10	-0.10
12	X_{min} increase	0.29	0.31	0.34	0.37	0.40	0.40
13	T_{rc} decrease	0.47	0.51	0.55	0.59	0.62	0.62
14	T_{rc} increase	-0.47	-0.44	-0.41	-0.40	-0.41	-0.40
15	T_{rh} decrease	0.13	0.13	0.13	0.13	0.13	0.13
16	T_{rh} increase	-0.13	-0.14	-0.14	-0.14	-0.14	-0.13
17	TDL increase	0.54	0.54	0.55	0.57	0.57	0.57

Table 30 Euclidean distance similarity results for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.04	0.03	0.03	0.02	0.02	0.04
2	X_{oa} increase	0.04	0.04	0.04	0.04	0.04	0.04
3	T_{preh} decrease	0.04	0.04	0.04	0.04	0.04	0.04
4	T_{preh} increase	0.04	0.04	0.03	0.03	0.03	0.04
5	PreHL increase	0.05	0.05	0.06	0.06	0.07	0.07
6	T_{cl} decrease	0.05	0.04	0.03	0.02	0.01	0.05
7	T_{cl} increase	0.03	0.02	0.02	0.01	0.01	0.03
8	T_{hl} decrease	0.04	0.03	0.02	0.02	0.01	0.04
9	T_{hl} increase	0.04	0.04	0.04	0.04	0.04	0.04
10	HL increase	0.00	0.04	0.04	0.05	0.05	0.05
11	X_{min} decrease	0.03	0.03	0.02	0.02	0.01	0.03
12	X_{min} increase	0.04	0.04	0.03	0.02	0.01	0.04
13	T_{rc} decrease	0.04	0.05	0.05	0.06	0.05	0.06
14	T_{rc} increase	0.03	0.03	0.02	0.02	0.01	0.03
15	T_{rh} decrease	0.04	0.04	0.04	0.04	0.03	0.04
16	T_{rh} increase	0.03	0.03	0.03	0.02	0.02	0.03
17	TDL increase	0.05	0.05	0.05	0.05	0.05	0.05

Figure 47 indicates that the cosine similarity values for the observed fault and reference control changes “Heat leakage of preheat coil increase” (ID 5) and “Heat leakage of heating coil increase” (ID 10) are almost identical and rank in the top two places among the 17 reference control changes. This suggests that control changes “Heat leakage of preheat coil increase” and “Heat leakage of heating coil increase” are the two most similar energy change patterns to the energy change pattern of the observed fault. They are the two most probable causes of the observed abnormal energy consumption.

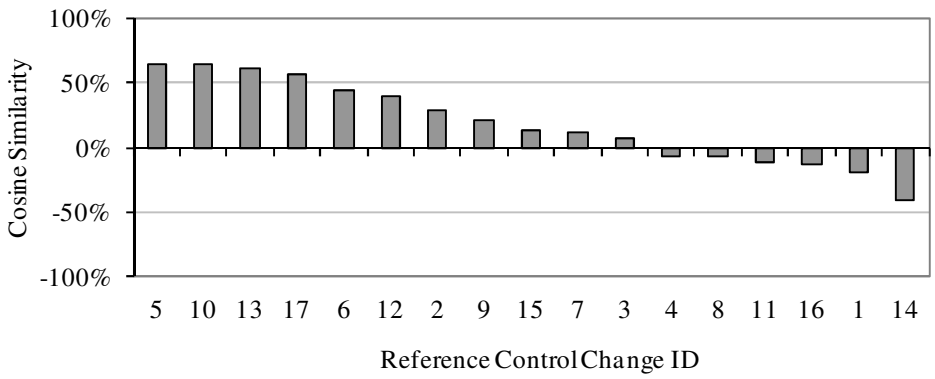

Figure 47 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building

The difference between the Euclidean distance similarity values of the different reference control changes range from 3% to 7% (Figure 48). The control change "Heat leakage of preheat coil increase" (ID 5) has the largest Euclidean distance similarity. The small value of Euclidean distance similarity is rooted in its definition. Recall from Section 2 that the expression to compute Euclidean distance similarity s(X,Y) is $s(X,Y) = e^{-d(X,Y)}$ (3), where d(X,Y) is the Euclidean distance within vectors X and Y. Figure 49 demonstrates that the Euclidean distance similarity exponentially falls with the increase of Euclidean distance. When Euclidean distance is 0.1, Euclidean distance similarity is 90%, and when the Euclidean distance is three, the similarity drops to only 5%. The Euclidean distance among the observed fault vector and all pre-determined reference control change vectors are above two in the test; thus, the corresponding Euclidean distance similarities based on expression (3) are all below 10%.

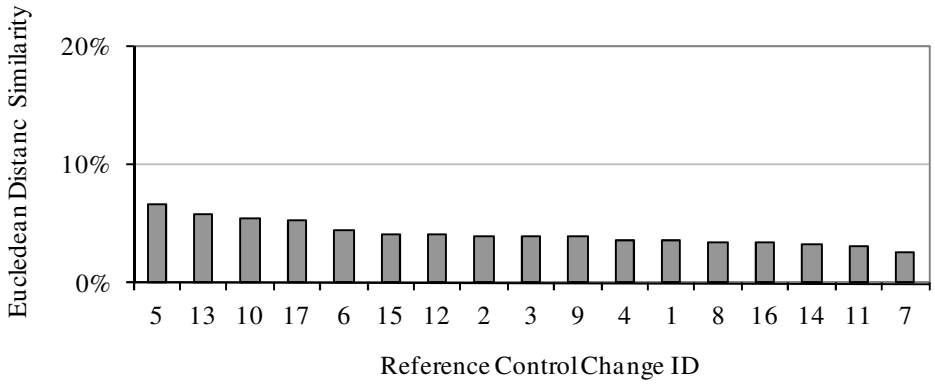

Figure 48 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building

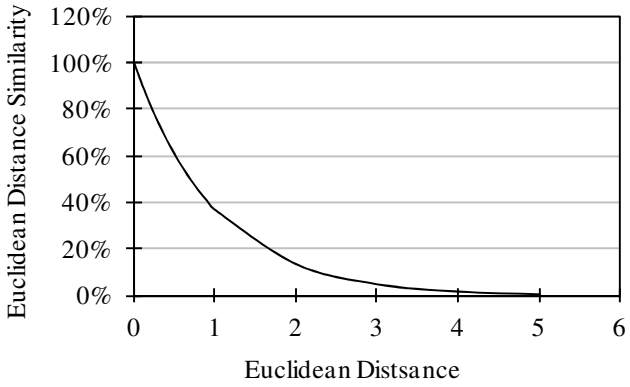

Figure 49 Euclidean distance similarity versus Euclidean distance

Both the Cosine Similarity and Euclidean Distance Similarity methods indicate the control change “Heat leakage of preheat coil increase” has the highest similarity and thus is considered to be the most probable reason for the observed abnormal energy consumption. The field inquiry indicates that there was a preheat valve leaking by on a pre-treat unit during the fault period (Claridge et al. 2009). The fault diagnosis results

with either the Cosine Similarity method or the Euclidean Distance Similarity method are consistent with the field inspection conclusion.

6.2.3 Summary

The Cosine Similarity and the Euclidean Distance Similarity methods were used to reveal the reasons for two abnormal consumption faults in the Sbisa Dining Hall and the Bush Academic Building respectively. In these two test cases, the fault diagnosis results for the Cosine Similarity methods match the field survey results in both cases, and the Euclidean Distance Similarity method only gives the correct answer in Bush Academic Building case.

7. INVESTIGATIONS ON SENSITIVITY OF THE WHOLE BUILDING FAULT DETECTION APPROACHES

The results from both the simulation tests and the field tests indicate that the DET-Toa method is superior to the DET-Date method for whole building fault detection. In this section, the sensitivity of the fault detection results of the DET-Toa method was examined using three factors: (1) the calibrated simulation model accuracy, (2) the fault severity, and (3) the time of fault occurrence.

Two representative HVAC systems, DDVAV and SDVAV were selected for the sensitivity analysis. The Eller Oceanography and Meteorology (EOM) building was chosen as the benchmark building for the DDVAV system and the Veterinary Research Building is chosen as the benchmark building for the SDVAV system. Basic information about these buildings was presented in Section 4.

7.1 Sensitivity to Calibrated Simulation Model Accuracy

The detection accuracy for different levels of calibrated simulation model accuracy is investigated in this section. When lower accuracy is found in the calibrated simulation model, more variation from the measured energy consumption will also be found in the predicted values. In the absence of real cases, white noise is added to the ideal simulation data sets to generate synthetic data against different levels of model accuracy. An increase in the white noise approximates the degradation of the model accuracy.

7.1.1 Simulated Data Sets

Each one of the following control changes is assumed to last for one year in a building. The ideal energy consumption under each specific control change is predicted by the calibrated simulation model in ABCAT. The parameter changes were chosen so the maximum monthly average cooling and heating consumption deviation caused by each synthetic control change is 15% of the yearly average cooling and heating consumption when there is no fault. The full description of the synthetic control changes is listed below.

DDVAV system (EOM Building)

It is assumed one of the following six synthetic control changes began on October 1, 1997 and lasted until September 30, 1998:

1. X_{oa} increase of 4.8%
2. T_{cl} decrease of 6.2°F
3. T_{hl} increase of 13.3°F
4. X_{min} increase of 23%
5. T_{rc} decrease of 2.2°F
6. T_{rh} increase of 8°F

SDVAV system (Veterinary Research Building)

It is assumed one of the following six synthetic control changes happened during the period from January 1 to December 31, 2000:

1. X_{oa} increase of 10%
2. T_{preh} increase of 5°F

3. T_{cl} decrease of 2.2°F
4. X_{min} increase of 6.7%
5. T_{rc} decrease of 2.3°F
6. T_{rh} increase of 3.5°F

The ideal simulated data sets were modified with normal white noise (WN(sd%)) at different levels, corresponding to equivalent levels of model accuracy. Abnormal energy use data modified samples with standard deviations of 3%, 5%, 10%, and 20% white noise were generated. The ideal abnormal data sets and the modified data sets were comparable to measured data. The test cases will be referred to as "control change description with WN %" in this section. For example, "T_{cl} decrease of 2°F with WN 3%" is the ideal abnormal data modified with 3% white noise. Obviously, for the ideal case, the white noise should be 0%.

The performance of the DET-Toa method was assessed for ideal fault cases and for the cases with variations on the ideal conditions by adding to the samples variation in the form of white noise to the samples. The daily model used to generate the synthetic data with variation was chosen as:

$$E=E\times(1+k\times\epsilon)$$

where E is the daily CHW or HW energy consumption in the ideal fault data sets; ϵ is a normally distributed random number with zero mean and unit standard deviation and k is a multiplier used to simulate varying levels of noise in the model with k = 3%, 5%, 10%, and 20% . Figure 50 shows an example of an ideal abnormal energy

data set and its modification by normal white noise (WN(10%)) for both CHW and HW energy use of a DDVAV system.

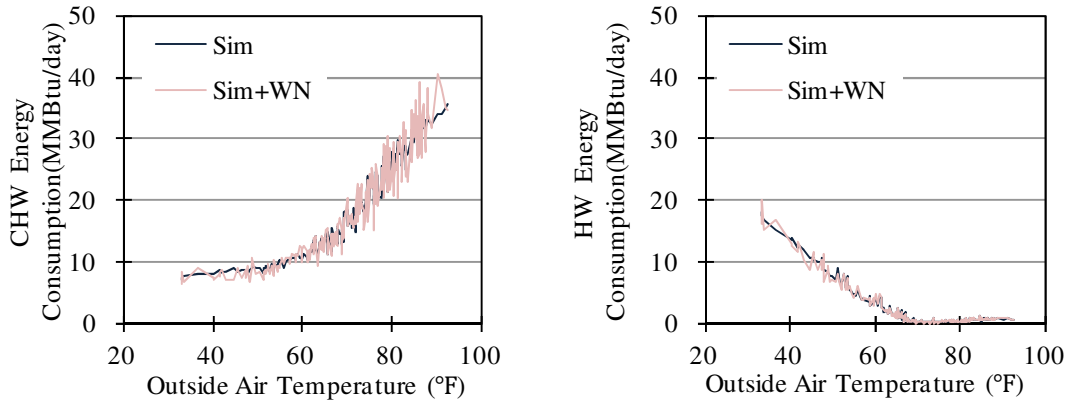

Figure 50 Synthetic data set and the same data modified by addition of 10% white noise (WN(10%)): samples of CHW and HW plotted as functions of outside air temperature

7.1.2 Detection Results

Table 31 shows the detection results obtained using the DET-Toa method in the EOM building. The results show that the DET-Toa method detected abnormal energy consumption in all six of the simulation cases when the added white noise level varied from 0% to 5%. It also detected abnormal energy consumption in five of the six simulation cases when the white noise level was 10% and in two of the six simulation cases when the white noise level was 20%. Similar results are seen for the Veterinary Research building in Table 32.

Detection accuracy is determined as the ratio of the number of cases where the DET-Toa method successfully identified the abnormal consumption to the total number of the cases. It was noted that the detection accuracy fell from 100% to 33% when the

white noise was increased from 0% to 20% in both buildings. This suggests that the involvement of the white noise weakens the capability of the DET-Toa method to detect faults. Tables 33 and 34 illustrate that with an identical control change, the increase of the white noise also delays the time when the abnormal consumption is identified. In the extreme condition, the DET-Toa method detected the abnormal consumption 10.5 months later when the white noise increased from 0% to 20% (Table 34).

Table 31 Detection results with the DET-Toa method on synthetic data modified with white noise for the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+4.8%	Y	Y	Y	Y	Y
T_{cl}-6.2°F	Y	Y	Y	Y	N
T_{hl}+13.3°F	Y	Y	Y	Y	N
X_{min}+23%	Y	Y	Y	Y	Y
T_{rc}-2.2°F	Y	Y	Y	N	N
T_{rh}+8°F	Y	Y	Y	Y	N

Table 32 Detection results with the DET-Toa method on synthetic data modified with white noise for the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+10%	Y	Y	Y	N	N
T_{cl}-2.2°F	Y	Y	Y	Y	N
T_{preh}+5°F	Y	Y	Y	Y	N
X_{min}+6.7%	Y	Y	Y	Y	N
T_{rc}-2.3°F	Y	Y	Y	Y	Y
T_{rh}+3.5°F	Y	Y	Y	Y	Y

Table 33 Fault identification day on synthetic data modified with white noise for the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+4.8%	12/12/1997	12/29/1997	12/29/1997	1/15/1998	1/15/1998
T_{cl}-6.2°F	11/12/1997	11/24/1997	11/24/1997	12/14/1997	
T_{hl}+13.3°F	11/27/1997	11/26/1997	11/28/1997	11/28/1997	
X_{min}+23%	11/9/1997	11/9/1997	11/9/1997	2/25/1998	2/26/1998
T_{rc}-2.2°F	3/29/1998	5/6/1998	7/11/1998		
T_{rh}+8°F	12/1/1997	12/1/1997	12/1/1997	12/26/1997	

Table 34 Fault identification day on synthetic data modified with white noise for the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+10%	5/29/2000	6/13/2000	6/25/2000		
T_{cl}-2.2°F	1/20/2000	1/30/2000	1/31/2000	10/18/2000	
T_{preh}+5°F	3/4/2000	3/4/2000	11/15/2000	12/25/2000	
X_{min}+6.7%	1/20/2000	1/30/2000	1/31/2000	10/18/2000	
T_{rc}-2.3°F	3/18/2000	3/18/2000	3/18/2000	3/18/2000	3/18/2000
T_{rh}+3.5°F	2/2/2000	2/2/2000	2/2/2000	2/4/2000	12/25/2000

The reason for the discrepancy in the detection results of the cases with different levels of white noise is that the white noise enlarges the variation of the energy consumption and thus intensifies the difficulty in meeting the fault detection metric of the DET-Toa method. For instance, Figure 51 compares the HW energy consumption changes between the cases "T_{cl}-2.2°F with WN 0%" and "T_{cl}-2.2°F with WN 3%" in the period from January 1-20. In Figure 7.2, the HW energy consumption changes remain constant at 1.16 sigma in the "WN 0%" case and scatter from 0.99 to 1.37 $SD\_{baseline}$ in the "WN 3%" case. In the "WN 0%" case, the DET-Toa method identified the

abnormally high HW energy consumption on January 20, as the condition that 20 days consecutive in outside air temperature in which the fault index is 1 was satisfied on that day. In the "WN 3%" case, the HW change decreased from 1.16 to 0.99 $SD_{\_baseline}$ on January 8; therefore, the sum of fault indexes only reached 19 on January 20. As a result, the DET-Toa method failed to identify the HW fault on January 20 when the white noise was 3%.

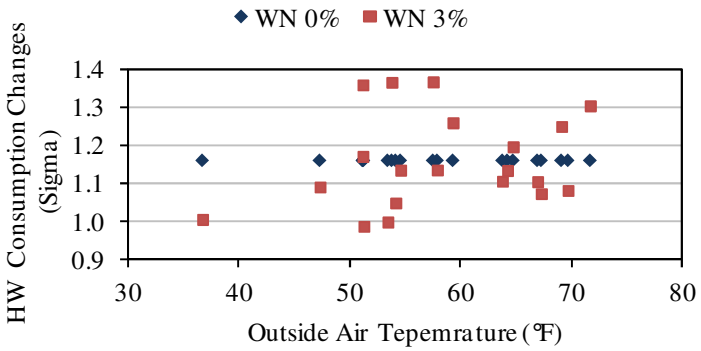

Figure 51 HW energy changes plotted as functions of outside air temperature for the period from 1/1 to 1/20 for the Veterinary Research Building

7.1.3 Summary

Overall, low model accuracy broadens the variation of the energy consumption and increases the difficulty for the DET-Toa method to detect abnormal consumption faults. The detection accuracy dropped from 100% to 33% when the white noise rose from 0% to 20% in the two investigated buildings. In addition, the time when abnormal consumption was discovered appeared to be postponed with the increase of the white noise.

7.2 Sensitivity to Fault Severity

7.2.1 Simulated Data Sets

The detection results of three fault groups with three different severity levels are compared in this section. The maximum monthly average energy use change index is 10%, 12%, and 15% respectively for groups one, two, and three. Table 35 summarizes the description of the three fault groups for the EOM Building and the Veterinary Research Building. Similar to the previous section, the performance of the DET-Toa method was evaluated for ideal fault cases and the modified cases in which the white noise was 3%, 5%, 10%, and 20% in the analysis for each fault group.

Table 35 Description of the faults assumed in the fault severity sensitivity study

EOM Building			Veterinary Research Building		
Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
X_{oa}+3.1%	X_{oa}+4%	X_{oa}+4.8%	X_{oa}+6.5%	X_{oa}+8%	X_{oa}+10%
T_{cl}-4.5°F	T_{cl}-5°F	T_{cl}-6.2°F	T_{cl}-1.5°F	T_{cl}-1.8°F	T_{cl}-2.2°F
T_{hl}+10°F	T_{hl}+11°F	T_{hl}+13.3°F	T_{preh}+3.3°F	T_{preh}+4°F	T_{preh}+5°F
X_{min}+16%	X_{min}+20%	X_{min}+23%	X_{min}+4.9%	X_{min}+5.5%	X_{min}+6.7%
T_{rc}-1.6°F	T_{rc}-1.8°F	T_{rc}-2.2°F	T_{rc}-1.7°F	T_{rc}-1.8°F	T_{rc}-2.3°F
T_{rh}+5.5°F	T_{rh}+6.5°F	T_{rh}+8°F	T_{rh}+2.3°F	T_{rh}+2.8°F	T_{rh}+3.5°F

7.2.2 Detection Results

The detection results for fault groups one and two are shown in Tables 36 through 39. The results for fault group three are presented in Tables 31 and 32 in the previous section. The results for the detection accuracy are presented in Figures 52 and 53. Under identical white noise levels, the detection accuracy of group three is higher

than the detection accuracy of group two, and the detection accuracy of group two is higher than the detection accuracy of group one. It is concluded that the detection accuracy increases with increasing fault severity level. By comparing the fault identification day within the ideal cases (WN0%) in the three fault groups, it is confirmed that the DET-Toa method identifies large-scale faults more quickly than identifies small-scale faults (Tables 40 and 41).

Table 36 Detection results of Fault Group 1 for the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+3.1%	Y	N	N	N	N
T_{cl}-4.5°F	Y	Y	Y	N	N
T_{hl}+10°F	N	N	N	N	N
X_{min}+16%	Y	N	N	N	N
T_{rc}-1.6°F	Y	N	N	N	N
T_{rh}+5.5°F	Y	Y	N	N	N

Table 37 Detection results of Fault Group 2 for the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+4%	Y	Y	N	N	N
T_{cl}-5°F	Y	Y	Y	N	N
T_{hl}+11°F	N	Y	Y	N	N
X_{min}+20%	Y	Y	Y	Y	Y
T_{rc}-1.8°F	Y	Y	Y	N	N
T_{rh}+6.5°F	Y	Y	Y	N	N

Table 38 Detection results of Fault Group 1 for the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+6.5%	Y	N	N	N	N
T_{cl}-1.5°F	Y	N	N	N	N
T_{preh}+3.3°F	Y	N	N	N	N
X_{min}+4.9%	Y	N	N	N	N
T_{rc}-1.7°F	Y	Y	Y	Y	N
T_{rh}+2.3°F	Y	Y	Y	N	N

Table 39 Detection results of Fault Group 2 for the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+8%	Y	Y	Y	N	N
T_{cl}-1.8°F	Y	N	N	N	N
T_{preh}+4°F	Y	Y	N	N	N
X_{min}+5.5%	Y	N	N	N	N
T_{rc}-1.8°F	Y	Y	Y	Y	Y
T_{rh}+2.8°F	Y	Y	Y	N	N

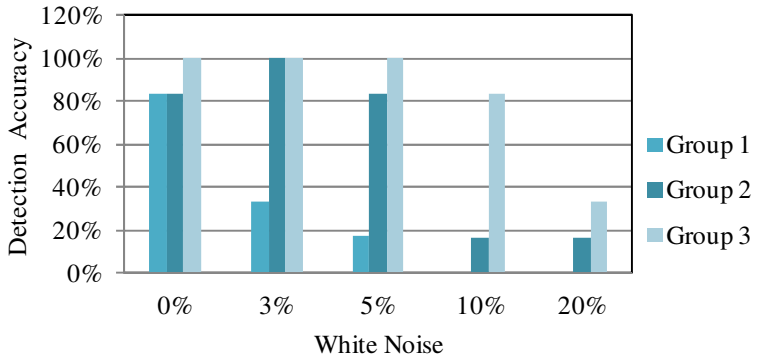

Figure 52 Comparison of detection accuracy of the three fault groups with different fault severity levels for the Bush Academic Building

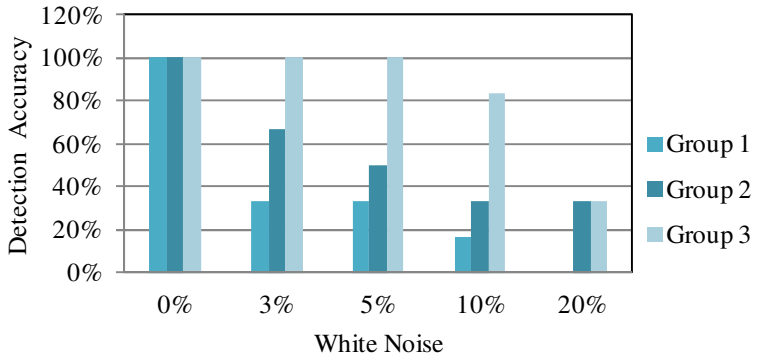

Figure 53 Comparison of detection accuracy of the three fault groups with different fault severity levels for the Veterinary Research Building

Table 40 Comparison of fault identification day within the ideal cases (WN 0%) in the three fault groups for the Bush Academic Building

	Group 1	Group 2	Group3
X_{oa}+	6/24/1998	6/12/1998	12/12/1997
T_{cl}-	12/6/1997	12/1/1997	11/12/1997
T_{hl}+			11/27/1997
X_{min}+	6/21/1998	11/9/1997	11/9/1997
T_{rc}-	6/21/1998	5/8/1998	3/29/1998
T_{rh}+	12/21/1997	12/1/1997	12/1/1997

Table 41 Comparison of fault identification day within the ideal cases (WN 0%) in the three fault groups for the Veterinary Research Building

	Group 1	Group 2	Group3
X_{oa}+	6/4/2000	5/30/2000	5/29/2000
T_{cl}-	3/15/2000	3/1/2000	1/20/2000
T_{preh}+	11/9/2000	10/8/2000	3/4/2000
X_{min}+	12/28/2000	3/3/2000	1/20/2000
T_{rc}-	3/18/2000	3/18/2000	3/18/2000
T_{rh}+	2/2/2000	2/2/2000	2/2/2000

7.2.3 Smallest Faults Detected by the DET-Toa Method

The minimum fault severity that the DET-Toa method can detect for each type of control change is summarized in Tables 42 and 43 respectively for the two analyzed buildings. The investigation interval is 1°F for the control changes "T_{preh}+", "T_{cl}-", "T_{hl}+", "T_{rc}-", and "T_{rh}+", and 1% for the control changes "X_{oa}+" and "X_{min}+". Three indexes - annual energy impact (AEI), energy impact when fault is identified (EIF), and the ratio of the energy impact when fault is identified to annual energy impact (REIF) - of these faults are displayed in Tables 44 through 49.

$$\text{Annual energy impact} = \frac{\text{One-year cumulative energy consumption variation caused by a fault}}{\text{One-year cumulative energy consumption when there is no fault}} \quad (11)$$

$$\text{Energy impact when fault is identified} = \frac{\text{Cumulative energy consumption variation from fault occurence to fault identification day}}{\text{One-year cumulative energy consumption when there is no fault}} \quad (12)$$

Table 42 Summary of the minimum fault severity for different control changes where the DET-Toa method can identify the abnormal consumption: EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	3%	4%	5%	5%	5%
T_{cl}-	4°F	4°F	4°F	6°F	9°F
T_{hl}+	12°F	11°F	11°F	12°F	15°F
X_{min}+	14%	16%	17%	18%	20%
T_{rc}-	2°F	2°F	2°F	3°F	4°F
T_{rh}+	5°F	6°F	6°F	8°F	11°F

Table 43 Summary of the minimum fault severity for different control changes where the DET-Toa method can identify the abnormal consumption: Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	6%	7%	7%	9%	12%
T_{cl}-	2°F	2°F	3°F	3°F	3°F
T_{preh}+	4°F	4°F	5°F	5°F	7°F
X_{min}+	5%	6%	6%	7%	7%
T_{rc}-	2°F	2°F	2°F	2°F	2°F
T_{rh}+	2°F	2°F	2°F	3°F	4°F

When the white noise is 0%, that is to say in the ideal condition, the minimum/ maximum/ median AEI, EIF, and REIF values are 2%/10.8%/5.2%, -0.2%/3.8%/1.7%, and -12.5%/66.1%/32.7% respectively for the smallest faults identified by the DET-Toa method in the EOM building. When the white noise increases to 20%, the minimum/ maximum/ median AEI, EIF, and REIF values rise to 3%/19.2%/10.2%, -1.1%/5.1%/3.3%, and -37%/76.1%/21.3% respectively for the smallest faults identified.

Table 44 AEI of the smallest faults that can be identified by the DET-Toa method in the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	2.0%	1.9%	1.8%	3.1%	3.0%
T_{cl}-	3.5%	3.4%	3.4%	5.7%	5.4%
T_{hl}+	10.8%	9.8%	9.8%	10.6%	13.3%
X_{min}+	5.8%	6.7%	7.2%	7.6%	8.4%
T_{rc}-	6.1%	6.0%	6.0%	12.0%	19.2%
T_{rh}+	4.7%	5.9%	5.9%	8.3%	12.0%
Max	10.8%	9.8%	9.8%	12.0%	19.2%
Median	5.2%	6.0%	5.9%	7.9%	10.2%
Min	2.0%	1.9%	1.8%	3.1%	3.0%

Table 45 AEI of the smallest faults that can be identified by the DET-Toa method in the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	3.3%	3.7%	3.7%	4.5%	5.8%
T_{cl}-	10.7%	10.6%	16.2%	15.9%	15.4%
T_{preh}+	2.5%	2.4%	3.1%	2.8%	4.1%
X_{min}+	9.1%	10.8%	10.7%	12.4%	11.9%
T_{rc}-	3.5%	3.4%	3.3%	3.3%	2.6%
T_{rh}+	2.2%	2.1%	2.0%	2.9%	3.6%
Max	10.7%	10.8%	16.2%	15.9%	15.4%
Median	3.4%	3.6%	3.5%	3.9%	4.9%
Min	2.2%	2.1%	2.0%	2.8%	2.6%

Table 46 EIF of the smallest faults that can be identified by the DET-Toa method in the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	-0.2%	0.2%	-0.5%	-1.1%	-1.1%
T_{cl}-	1.6%	1.5%	1.5%	2.4%	4.1%
T_{hl}+	1.8%	2.3%	4.8%	4.8%	2.6%
X_{min}+	3.8%	3.4%	3.6%	4.7%	5.1%
T_{rc}-	1.3%	3.4%	4.8%	2.3%	4.1%
T_{rh}+	2.3%	1.8%	2.1%	2.6%	2.6%
Max	3.8%	3.4%	4.8%	4.8%	5.1%
Median	1.7%	2.0%	2.8%	2.5%	3.3%
Min	-0.2%	0.2%	-0.5%	-1.1%	-1.1%

For the smallest faults identified in the Veterinary Research building, when the white noise is 0%, the minimum/maximum/median AEI, EIF, and REIF values are 2.2%/10.7%/3.4%, -0.6%/8.9%/0.8%, and -16.5%/98%/25.5% respectively. When WN

grows to 20%, the minimum/maximum/median AEI, EIF, and REIF values are 2.6%/15.4%/4.9%, -1.0%/3.6%/2.0% and -39.2%/88.3%/31.7% respectively.

Table 47 EIF of the smallest faults that can be identified by the DET-Toa method in the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	0.9%	1.4%	1.4%	1.9%	2.5%
T_{cl}-	0.8%	2.1%	1.2%	1.8%	3.3%
T_{preh}+	0.9%	0.8%	1.3%	2.5%	3.6%
X_{min}+	8.9%	0.8%	-0.5%	-0.5%	-1.0%
T_{rc}-	-0.6%	-0.6%	-0.5%	-0.5%	-1.0%
T_{rh}+	0.5%	0.6%	1.1%	1.0%	1.5%
Max	8.9%	2.1%	1.4%	2.5%	3.6%
Median	0.8%	0.8%	1.1%	1.4%	2.0%
Min	-0.6%	-0.6%	-0.5%	-0.5%	-1.0%

Table 48 REIF of the smallest faults that can be identified by the DET-Toa method in the EOM Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	-12.5%	10.5%	-29.0%	-33.6%	-37.0%
T_{cl}-	44.9%	42.8%	44.7%	42.6%	76.1%
T_{hl}+	16.3%	23.7%	49.6%	45.0%	19.1%
X_{min}+	66.1%	50.8%	49.8%	62.1%	60.4%
T_{rc}-	20.6%	55.5%	80.6%	19.3%	21.1%
T_{rh}+	48.3%	29.9%	35.6%	31.0%	21.5%
Max	66.1%	55.5%	80.6%	62.1%	76.1%
Median	32.7%	36.4%	47.2%	36.8%	21.3%
Min	-12.5%	10.5%	-29.0%	-33.6%	-37.0%

Table 49 REIF of the smallest faults that can be identified by the DET-Toa method in the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+	26.7%	37.9%	39.1%	40.8%	43.7%
T_{cl}-	7.0%	19.5%	7.2%	11.2%	21.2%
T_{preh}+	36.3%	34.1%	43.2%	87.6%	88.3%
X_{min}+	98.0%	7.6%	-5.1%	-4.4%	-8.5%
T_{rc}-	-16.5%	-16.5%	-16.5%	-16.5%	-39.2%
T_{rh}+	24.2%	29.7%	55.2%	33.3%	42.3%
Max	98.0%	37.9%	55.2%	87.6%	88.3%
Median	25.5%	24.6%	23.1%	22.3%	31.7%
Min	-16.5%	-16.5%	-16.5%	-16.5%	-39.2%

7.2.4 Summary

When comparing the detection results of three fault groups with three different fault severity levels, it is found that an increase in the fault scale improves the detection accuracy and let the fault to be detected earlier. The smallest faults that could be identified by the DET-Toa method in both buildings and their related energy consumption impact statistics are provided in this section.

7.3 Sensitivity to the Time of Fault Occurrence

Tables 50 and 51 compare the smallest faults identified by the DET-Toa method when each fault was assumed to occur on January 1, April 1, July 1, and December 1 and last for one year. The four selected start dates of the faults represent the start of spring, summer, fall and winter individually. The results indicate that the minimum fault magnitude noticed is the same with the four different times of fault occurrence for the

same control change. The detailed detection results of these faults are illustrated and compared below.

Table 50 Smallest faults identified by the DET-Toa method with different times of fault occurrence for the EOM Building

	The Time of Fault Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+	3%	3%	3%	3%
T_{cl}-	4°F	4°F	4°F	4°F
T_{hl}+	12°F	12°F	12°F	12°F
X_{min}+	14%	14%	14%	14%
T_{rc}-	2°F	2°F	2°F	2°F
T_{rh}+	5°F	5°F	5°F	5°F

Table 51 Smallest faults identified by the DET-Toa method with different times of fault occurrence for the Veterinary Research Building

	The Time of Fault Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+	6%	6%	6%	6%
T_{cl}-	2°F	2°F	2°F	2°F
T_{preh}+	4°F	4°F	4°F	4°F
X_{min}+	5%	5%	5%	5%
T_{rc}-	2°F	2°F	2°F	2°F
T_{rh}+	2°F	2°F	2°F	2°F

Tables 52 and 53 compare the fault identification day with four different times of fault occurrence. The results show that all the synthetic faults are detected whenever the fault occurs, whereas the fault identification day is not on the same day. A certain tendency for the fault identification day is discovered for the identical fault with different times of occurrence. Table 52 indicates that the DET-Toa method identified the

faults "X_{oa}+3%", "T_{hl}+12°F", and "T_{rc}-2°F" in summer, the faults "T_{cl}-2°F" and "T_{rh}+5°F" in winter, and the fault "X_{min}+14%" in spring or fall. According to the data in Table 52, the fault is most likely to be identified in the season when the fault has the largest impact on energy consumption variation. For instance, the fault "X_{oa}+3%" has the largest impact on the CHW energy consumption in the high outside air temperature range (Figure 54.1), so the fault identification days are concentrated in summer no matter when the fault occurs. Figures 54 and 55 show the energy consumption variations as functions of outside air temperature.

Table 52 The fault identification day of the smallest faults identified by the DET-Toa method with different times of fault occurrence for the EOM Building

	The Time of Fault Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+3%	6/24	6/24	7/24	6/24
T_{cl}-4°F	3/1	12/12	12/12	12/12
T_{hl}+12°F	7/17	7/17	8/26	11/28
X_{min}+14%	12/23	3/17	4/19	4/19
T_{rc}-2°F	5/28	5/28	7/20	5/8
T_{rh}+5°F	2/27	2/8	2/8	2/8

Table 53 The Fault identification day of the smallest faults identified by the DET-Toa method with different times of fault occurrence for the Veterinary Research Building

	The Fault Time of Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+6%	6/17	6/17	7/20	6/16
T_{cl}-2°F	1/20	9/25	10/19	11/5
T_{preh}+4°F	10/8	12/7	12/7	12/7
X_{min}+5%	12/22	1/27	1/27	1/27
T_{rc}-2°F	3/18	5/29	7/20	2/12
T_{rh}+2°F	2/2	11/27	11/30	11/30

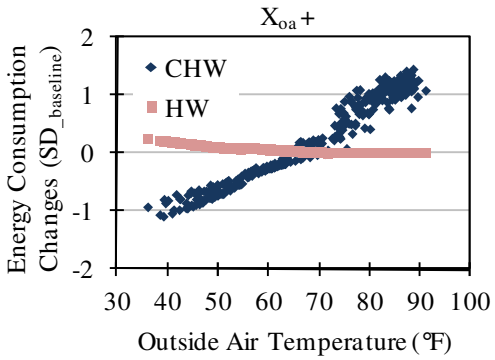

Figure 54.1

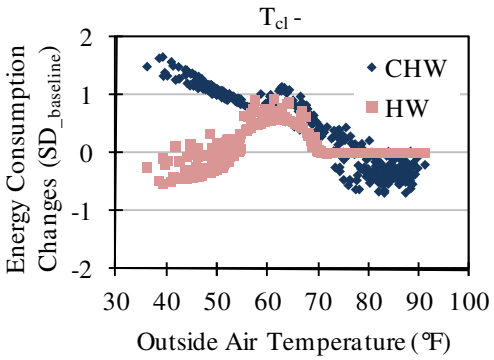

Figure 54.2

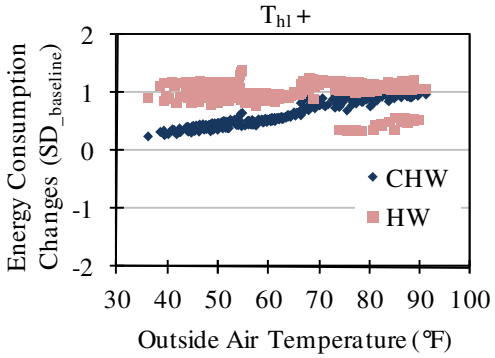

Figure 54.3

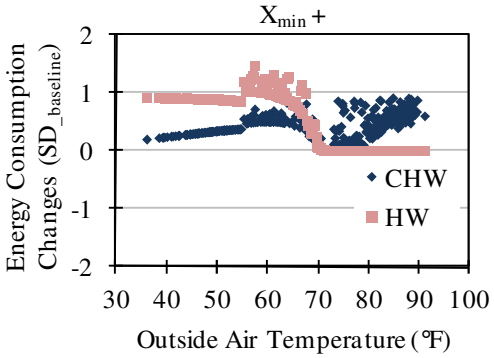

Figure 54.4

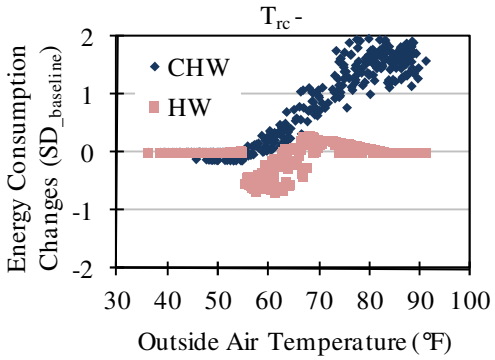

Figure 54.5

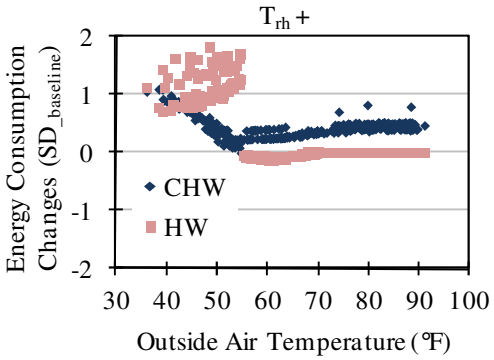

Figure 54.6

Figure 54 Cooling and heating energy changes plotted as functions of outside air temperature of the smallest faults identified by the DET-Toa method in the EOM Building. (54.1) X_{oa}+3%, (54.2) T_{cl}-4°F, (54.3) T_{hl}+12°F, (54.4) X_{min}+14%, (54.5) T_{rc}-2°F, (54.6) T_{rh}+5°F.

A similar situation is seen in Table 53. The DET-Toa method detected the fault "X_{oa}+6%" in summer, detected the fault "T_{cl}-2°F" in spring, fall, or winter, detected the faults "T_{preh}+4°F", "X_{min}+5%" and "T_{rh}+2°F" in winter, and detected fault "T_{rc}-2°F" in spring, summer or fall. Again, the faults tend to be identified during the season when the fault has the largest impact on energy consumption variation. It is natural that this tendency of the fault identification day would disappear if the fault scale is large enough that the energy variation is bigger than one $SD_{\_baseline}$ over the entire outside air temperature range.

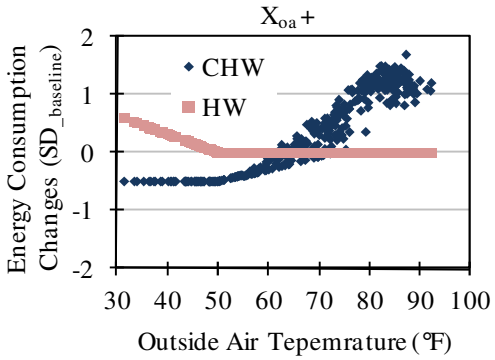

Figure 55.1

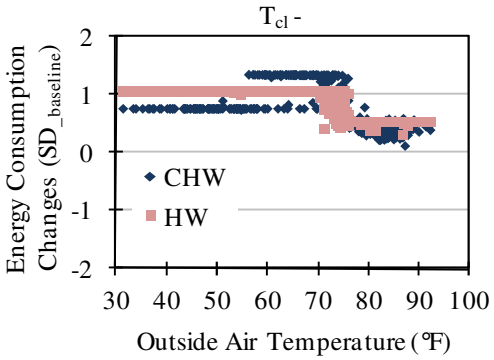

Figure 55.2

Figure 55 Cooling and heating energy changes plotted as functions of outside air temperature of the smallest Faults identified by the DET-Toa method in the Veterinary Research Building. (55.1) X_{oa}+6%, (55.2) T_{cl}-2°F, (55.3) T_{preh}+4°F, (55.4) X_{min}+5%, (55.5) T_{rc}-2°F, (55.6) T_{rh}+2°F.

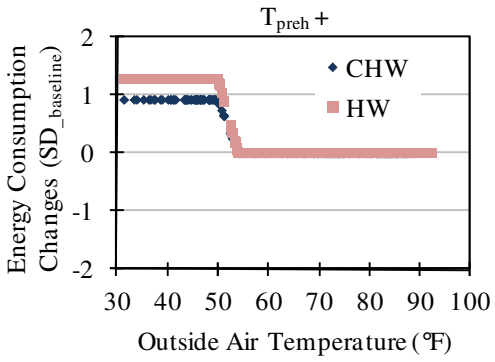

Figure 55.3

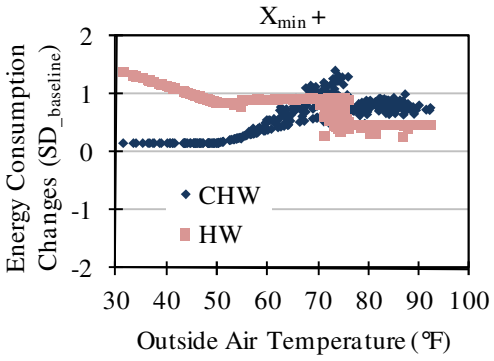

Figure 55.4

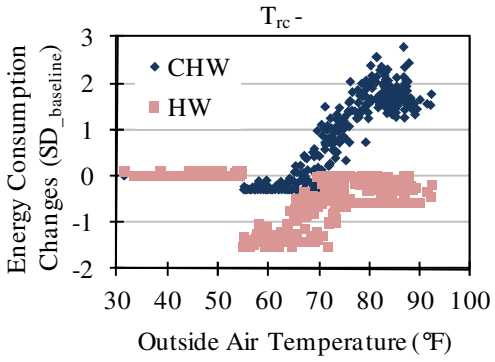

Figure 55.5

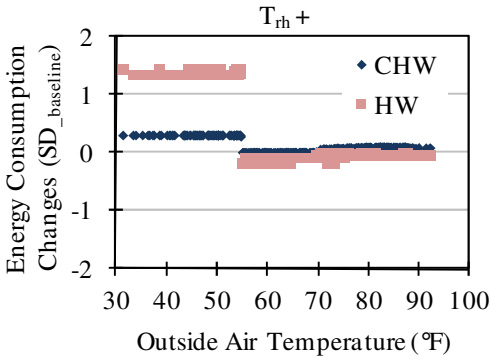

Figure 55.6

Figure 55 Continued.

Surveying the time between the start day of the fault and the fault identification day, the time of fault occurrence where the shortest time for identification of each type of control change is summarized in Table 54. The result varies for different control changes, because the outside air temperature range where the control change has the strongest influence on energy deviation is different for different types of control change. Tables 55 through 58 provide an overview of the indexes of EIF and REIF for the two buildings analyzed.

Table 54 The times of fault occurrence for which the minimum number of days of data is needed for the DET-Toa method to detect faults

EOM Building (DDVAV)		Veterinary Reasearch Building (SDVAV)	
X_{oa}+3%	1-Jul	X_{oa}+6%	1-Jul
T_{cl}-4°F	1-Jan	T_{cl}-2°F	1-Jan
T_{hl}+12°F	1-Jul	T_{preh}+4°F	1-Oct
X_{min}+14%	1-Oct	X_{min}+5%	1-Oct
T_{rc}-2°F	1-Jul	T_{rc}-2°F	1-Jul
T_{rh}+5°F	1-Jan	T_{rh}+2°F	1-Jan

Table 55 EIF of the smallest faults that can be identified by the DET-Toa method with different times of fault occurrence in the EOM Building

	The Time of Fault Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+3%	0.2%	1.0%	3.2%	-0.2%
T_{cl}-4°F	1.4%	1.0%	0.7%	1.6%
T_{hl}+12°F	5.7%	3.6%	4.9%	1.8%
X_{min}+14%	5.6%	5.6%	1.9%	3.8%
T_{rc}-2°F	1.2%	1.1%	0.7%	1.3%
T_{rh}+5°F	1.0%	3.9%	0.6%	2.3%

Table 56 REIF of the smallest faults that can be identified by the DET-Toa method with different times of fault occurrence in the EOM Building

	The Time of Fault Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+3%	12.4%	49.0%	67.3%	-12.5%
T_{cl}-4°F	41.5%	30.2%	11.6%	44.9%
T_{hl}+12°F	52.9%	33.0%	85.0%	16.3%
X_{min}+14%	97.4%	96.1%	17.7%	66.1%
T_{rc}-2°F	19.7%	17.6%	20.9%	20.6%
T_{rh}+5°F	22.1%	83.0%	31.8%	48.3%

Table 57 EIF of the smallest faults that can be identified by the DET-Toa method with different times of fault occurrence in the Veterinary Research Building

	The Fault Time of Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+6%	0.9%	1.1%	0.5%	0.8%
T_{cl}-2°F	0.8%	3.9%	2.4%	1.3%
T_{preh}+4°F	0.9%	0.8%	0.8%	0.8%
X_{min}+5%	8.9%	7.5%	5.1%	2.9%
T_{rc}-2°F	-0.6%	0.6%	0.7%	-0.5%
T_{rh}+2°F	0.5%	0.6%	0.6%	0.5%

Table 58 EIF of the smallest faults that can be identified by the DET-Toa method with different times of fault occurrence in the Veterinary Research Building

	The Fault Time of Occurrence			
	1-Jan	1-Apr	1-Jul	1-Oct
X_{oa}+6%	26.7%	32.6%	15.8%	25.0%
T_{cl}-2°F	7.0%	35.9%	21.9%	12.4%
T_{preh}+4°F	36.3%	33.5%	33.2%	33.2%
X_{min}+5%	98.0%	82.2%	56.6%	31.5%
T_{rc}-2°F	-16.5%	17.5%	19.1%	-13.5%
T_{rh}+2°F	24.2%	29.2%	27.2%	23.8%

To summarize, the time of fault occurrence has no effect on the scale of the smallest faults that can be identified by the DET-Toa method. The outside air temperature range where a control change has the largest influence on the energy change differs for different types of control change. For the low-scale control change, the abnormal consumption is more likely to be detected in the season when that control change significantly increases/decreases energy consumption. Hence, the abnormal symptom will be discovered faster if the control change coincidently occurs in that season.

8. INVESTIGATIONS ON SENSITIVITY OF THE WHOLE BUILDING FAULT DIAGNOSIS APPROACHES

The results of both the simulation tests and tests using site data from the two existing building have validated the effectiveness of the Cosine Similarity and the Euclidean Distance Similarity methods in isolating the causes of abnormal consumption faults. The diagnosis result may be influenced by many factors. This section studies the impact of reference control change magnitude settings, calibrated simulation model accuracy, fault severity, and fault period length on the robustness of the developed fault diagnosis approaches.

Similar to what was done in Section 7, the sensitivity analysis was conducted on two representative HVAC systems, DDVAV and SDVAV. The Bush Academic Building and the Veterinary Research Building are the two buildings investigated with DDVAV systems and SDVAV systems respectively. The building information is available in Sections 4 and 5.

8.1 Sensitivity to Reference Control Change Magnitude Setting

As presented in Section 5, the simulation test found that the reference control change magnitude setting has a significant impact on the diagnosis results for the Euclidean Distance Similarity method. This section details the influence of the reference control change magnitude settings on the Cosine Similarity and the Euclidean Distance Similarity methods.

8.1.2 DDVAV System (Bush Academic Building)

The proposed sensitivity analysis is intended to investigate the influence of the magnitude setting. This information is used for reducing the number of levels of magnitude settings in the reference library and making it more concise. Combined with the levels of magnitude, the reference control change library designed for the Bush Academic Building is illustrated in Table 59. Seventeen types of reference control change are included and each control change contains five levels of magnitude. There are five reference control change signature vectors for a single reference control change. The cosine similarity and the Euclidean distance similarity between the signature vector with smallest magnitude (magnitude I) and the signature vectors with magnitudes I, II, III, IV, and V were computed for each type of reference control change. The results are shown in Figure 56. The numbers 1, 2, 3, 4, and 5 stand for the magnitudes I, II, III, IV, and V respectively.

Figure 56.1 shows that the value of the cosine similarity of the signature vector "X_{oa} decreases 2%" is 100% with signature vectors "X_{oa} decreases 2%", "X_{oa} decreases 4%", "X_{oa} decreases 6%", "X_{oa} decreases 8%" and "X_{oa} decreases 10%." The value of the Euclidean distance similarity of the signature vector "X_{oa} decreases 2%" is 100% with the vector "X_{oa} deceases 2%", 57% with the vector "X_{oa} decreases 4%", 32% with the vector "X_{oa} decreases 6%", 18% with the vector "X_{oa} decreases 8%", and 10% with the vector "X_{oa} decrease 10%". We can infer that the cosine similarity result is not sensitive to the magnitude setting for the control change "X_{oa} decrease", because there is no difference among the results with different levels of magnitude. In contrast, the

Euclidean distance similarity result is very sensitive to the magnitude setting, because the sensitivity drops dramatically with the increase of the magnitude level.

Similar analysis was performed for the other 16 types of reference control change. The difference between the similarity of magnitude I and magnitude I and the similarity of magnitudes I and V are used to address the sensitivity, namely, the similarity difference index. The minimum/maximum/median similarity difference index of the 17 types of reference control changes were 0%/74%/6% for the Cosine Similarity method, and 31%/100%/93% for the Euclidean Distance Similarity method. In general, the magnitude setting has a much stronger effect on the Euclidean Distance Similarity method than on the Cosine Similarity method.

Table 59 Reference control change library for the Bush Academic Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%	-4%	-6%	-8%	-10%	
2	X_{oa} increase	2%	10%	30%	60%	80%	
3	T_{preh} decrease	-3	-6	-9	-12	-15	°F
4	T_{preh} increase	3	6	9	12	15	°F
5	PreHL increase	10	20	30	40	50	kBtu/hr
6	T_{cl} decrease	-2	-4	-6	-8	-10	°F
7	T_{cl} increase	2	4	6	8	10	°F
8	T_{hl} decrease	-2	-6	-12	-18	-20	°F
9	T_{hl} increase	2	10	20	30	40	°F
10	HL increase	20	40	60	80	100	kBtu/hr
11	X_{min} decrease	-2%	-4%	-6%	-10%	-20%	
12	X_{min} increase	2%	10%	30%	40%	50%	
13	T_{rc} decrease	-1	-2	-4	-6	-10	°F
14	T_{rc} increase	1	2	4	6	10	°F
15	T_{rh} decrease	-1	-2	-4	-6	-10	°F
16	T_{rh} increase	1	2	4	6	10	°F
17	TDL increase	2%	10%	30%	40%	50%	

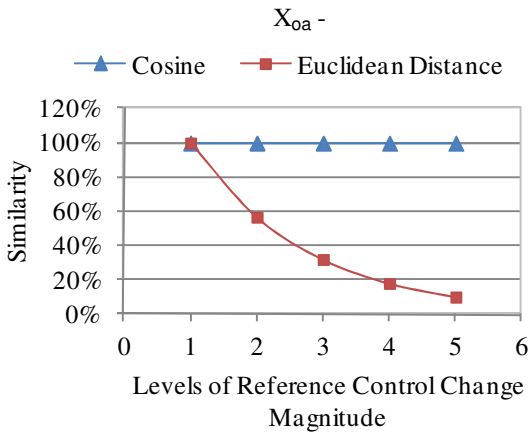

Figure 56.1

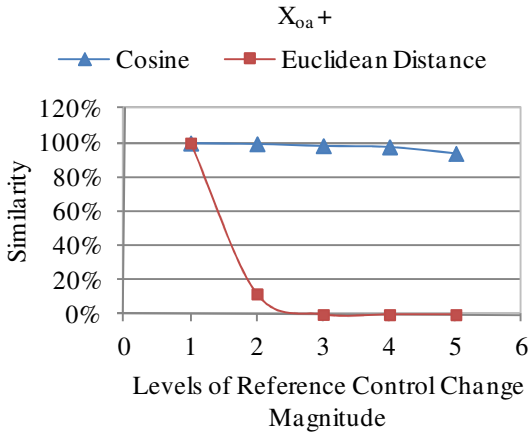

Figure 56.2

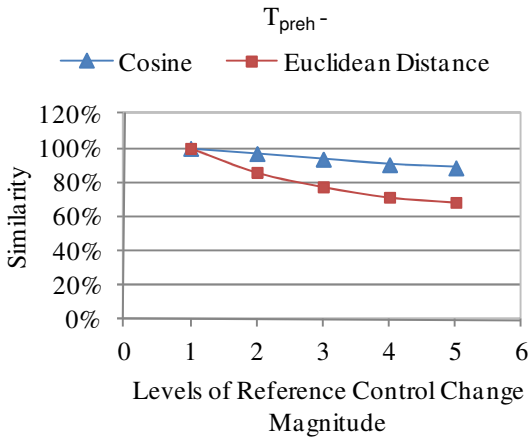

Figure 56.3

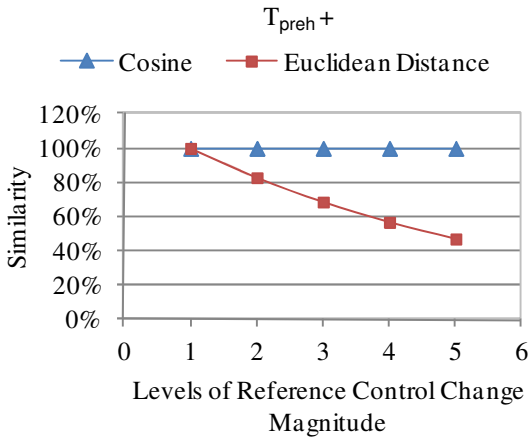

Figure 56.4

Figure 56 Cosine similarity and Euclidean distance similarity within signature vector (magnitude I) and signature vectors (magnitude I, II, III, IV, and V) for different types of reference control change for the Bush Academic Buildings. (56.1) X_{oa} decrease, (56.2) X_{oa} increase, (56.3) T_{preh} decrease, (56.4) T_{preh} increase, (56.5) The heat leakage of preheat Coil increase, (56.6) T_{cl} decrease, (56.7) T_{cl} increase, (56.8) T_{hl} decrease, (56.9) T_{hl} increase, (56.10) The heat leakage of heating coil increase, (56.11)X_{min} decrease, (56.12) X_{min} increase, (56.13) T_{rc} decrease, (56.14) T_{rc} increase, (56.15) T_{rh} decrease, (56.16) T_{rh} increase, (56.17) Terminal box damper leakage increase.

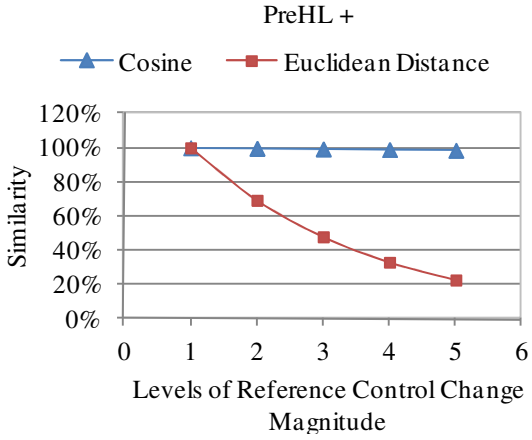

Figure 56.5

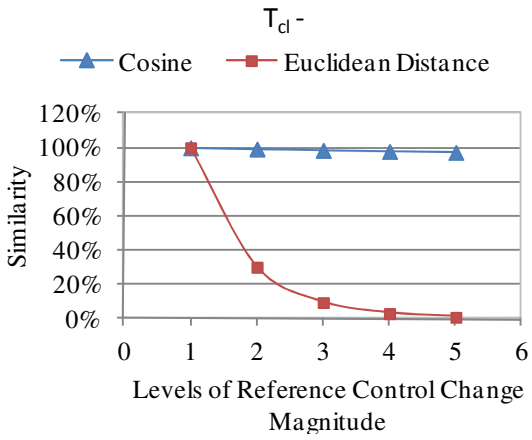

Figure 56.6

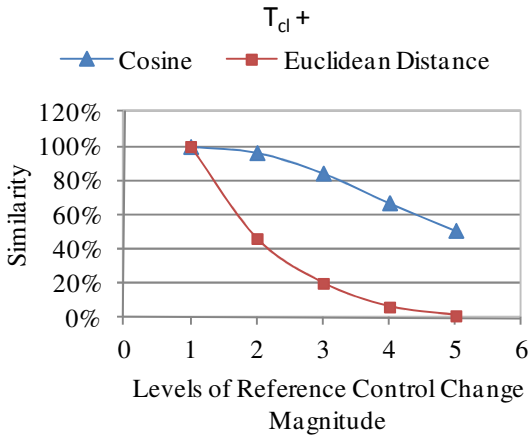

Figure 56.7

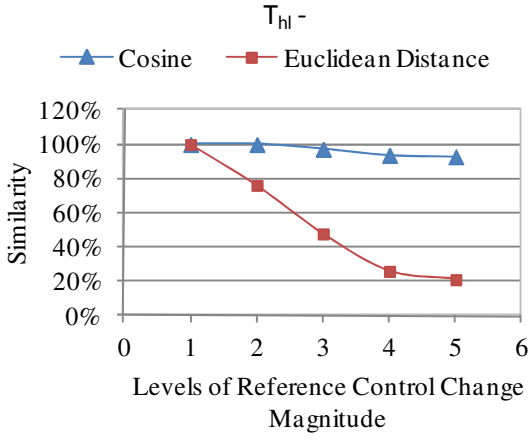

Figure 56.8

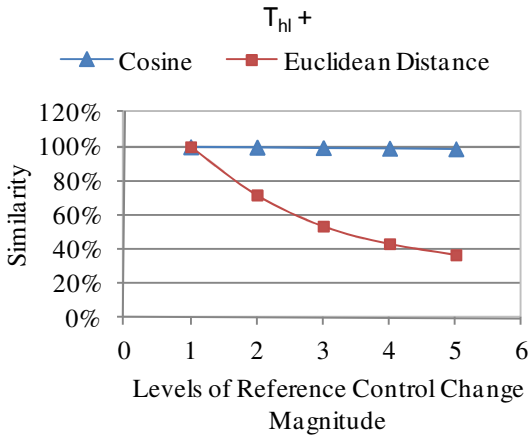

Figure 56.9

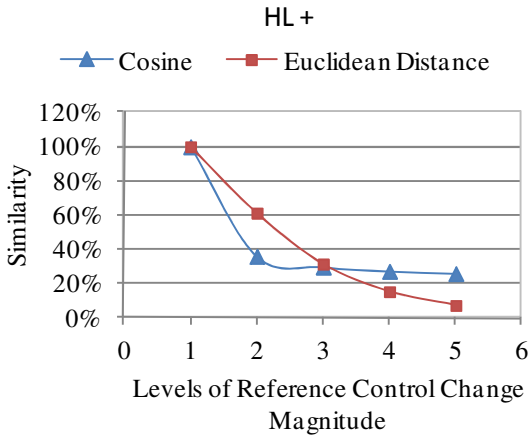

Figure 56.10

Figure 56 Continued.

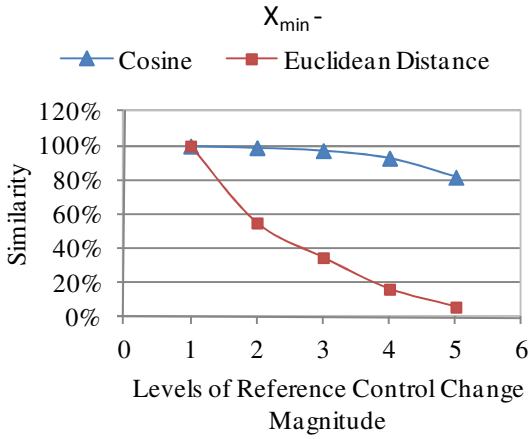

Figure 56.11

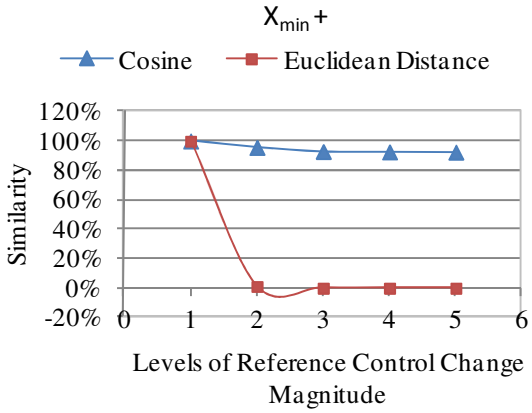

Figure 56.12

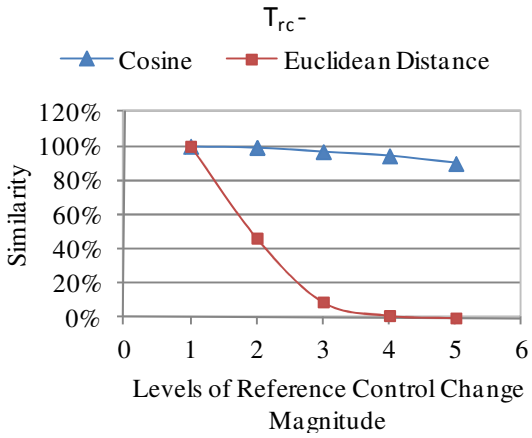

Figure 56.13

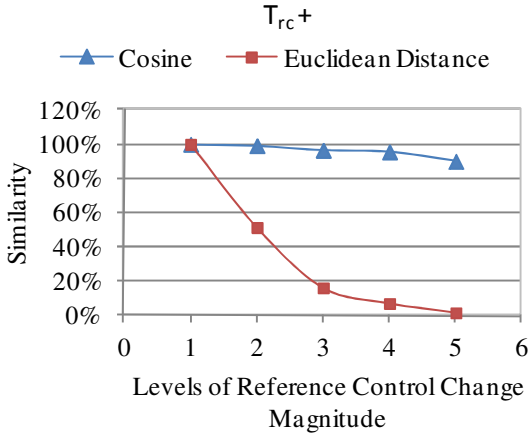

Figure 56.14

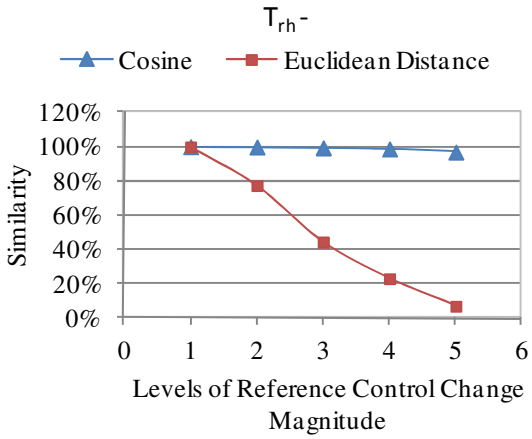

Figure 56.15

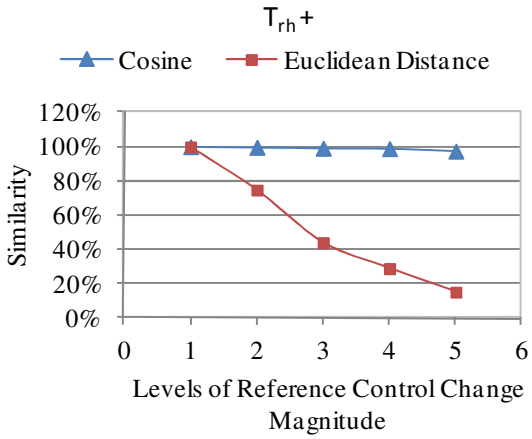

Figure 56.16

Figure 56 Continued

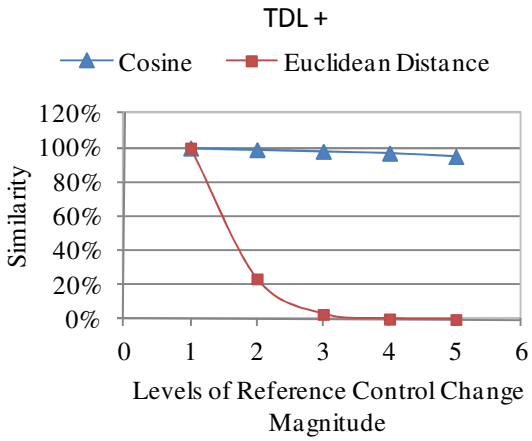

Figure 56.17

Figure 56 Continued.

8.1.3 SDVAV System (Veterinary Research Building)

The sensitivity analysis against the reference control change magnitude settings in the Veterinary Research Building was also studied. Table 60 presents the reference control change library for this building. The cosine similarity and the Euclidean distance similarity between the signature vector with smallest magnitude (I) and the signature vectors with magnitudes I, II, III, IV, and V are shown in Figure 57. The minimum/maximum/median similarity difference index of the 12 types of reference control change were 0%/42%/5% for the cosine similarity, and 61%/100%/100% for the Euclidean distance similarity. Again, the magnitude setting has a much greater influence on the Euclidean Distance Similarity method than on the Cosine Similarity method.

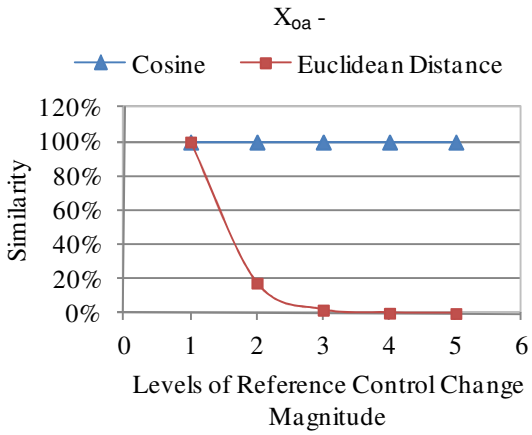

Figure 57.1

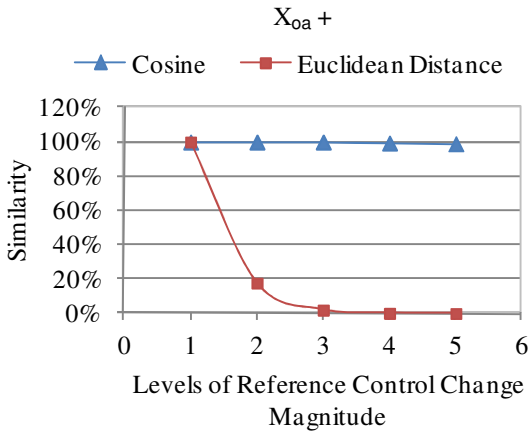

Figure 57.2

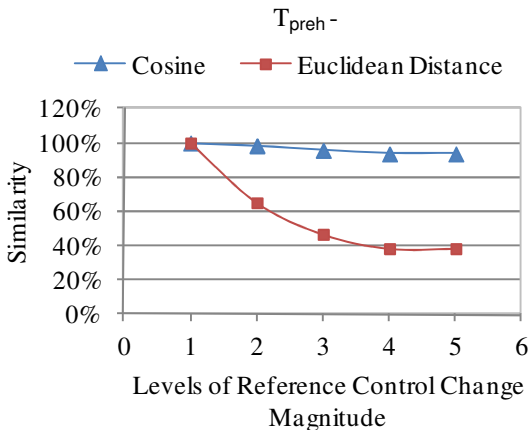

Figure 57.3

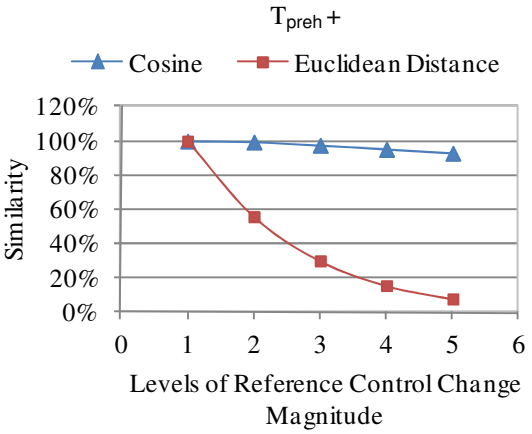

Figure 57.4

Figure 57 Cosine similarity and Euclidean distance similarity within signature vector (magnitude I) and signature vectors (magnitude I, II, III, IV, and V) for different types of reference control change for the Veterinary Research Building. (57.1) X_{oa} decrease, (57.2) X_{oa} increase, (57.3) T_{preh} decrease, (57.4) T_{preh} increase, (57.5) T_{cl} decrease, (57.6) T_{cl} increase, (57.7)X_{min} decrease, (57.8) X_{min} increase, (57.9) T_{rc} decrease, (57.10) T_{rc} increase, (57.11) T_{rh} decrease, (57.12) T_{rh} increase.

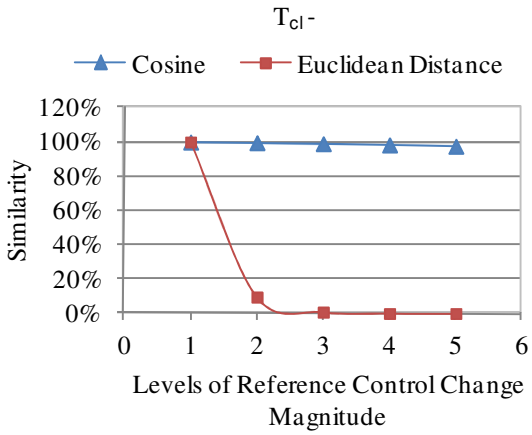

Figure 57.5

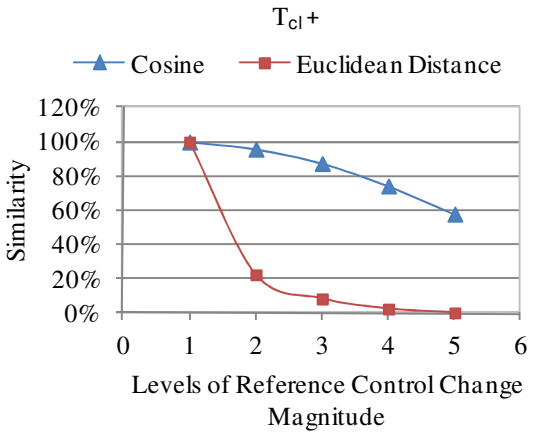

Figure 57.6

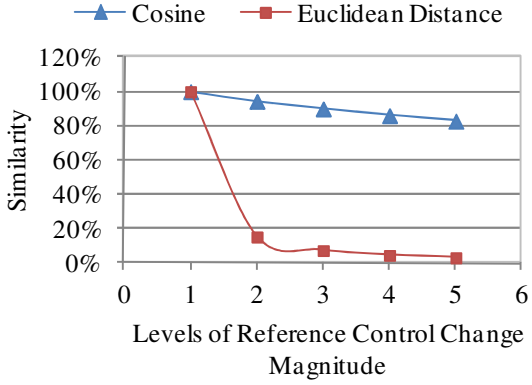

Figure 57.7

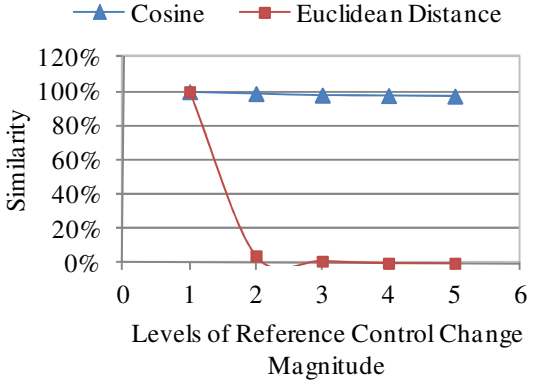

Figure 57.8

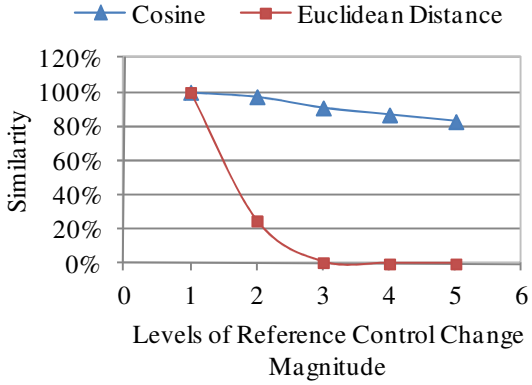

Figure 57.9

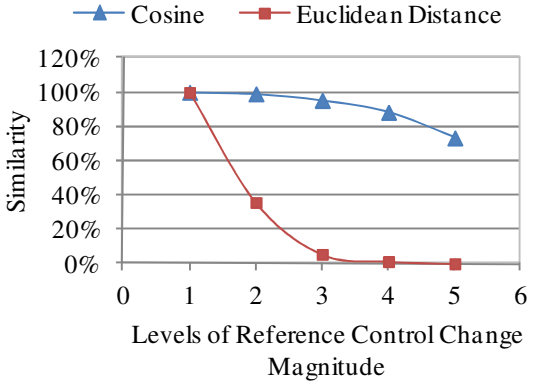

Figure 57.10

Figure 57 Continued

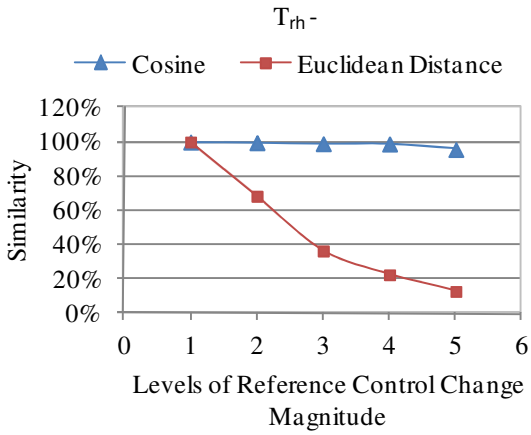

Figure 57.11

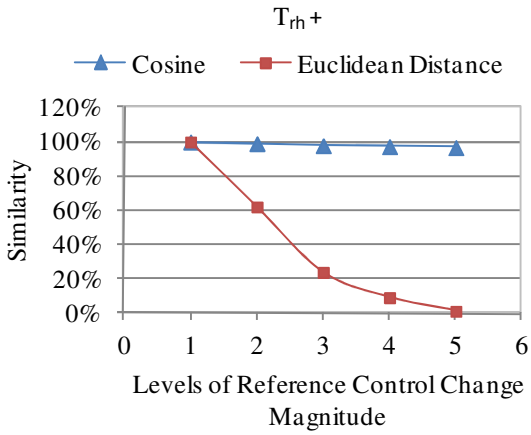

Figure 57.12

Figure 57 Continued.

Table 60 Reference control change library for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%	-10%	-20%	-30%	-40%	
2	X_{oa} increase	2%	10%	20%	30%	35%	
3	T_{preh} decrease	-2	-4	-6	-8	-10	°F
4	T_{preh} increase	2	4	6	8	10	°F
5	T_{cl} decrease	-2	-4	-6	-8	-10	°F
6	T_{cl} increase	2	4	6	8	10	°F
7	X_{min} decrease	-2%	-10%	-20%	-30%	-40%	
8	X_{min} increase	2%	10%	20%	30%	40%	
9	T_{rc} decrease	-1	-2	-4	-6	-10	°F
10	T_{rc} increase	1	2	4	6	10	°F
11	T_{rh} decrease	-1	-2	-4	-6	-10	°F
12	T_{rh} increase	1	2	4	6	10	°F

8.1.4 Discussion

The results for the Euclidean Distance Similarity method and the Cosine Similarity method are quite different because that they are based on essentially different

concepts. Euclidean distance similarity is a distance-based similarity measure. It considers only the impact of the distance between vectors. Conversely, cosine similarity is a direction-based similarity measure. It considers only the impact of the direction of vectors. The alteration of the control change magnitude dramatically varies the values of energy consumption dramatically but only changes its direction slightly. Consequently, a proper magnitude setting is much more important to the Euclidean Distance Similarity method than it is to the Cosine Similarity method.

In Section 5, the Euclidean Distance Similarity method isolates the incorrect fault cause in the simulation case "T_{hl}+30°F" because a too narrow magnitude setting. Another example of a false diagnosis resulting from inappropriate magnitude setting follows.

It is assumed that the minimum airflow ratio increased 1.2% during the weekdays from July 2008 to June 2009 for the Bush Academic Building. If we use the reference library in Table 59, the Euclidean Distance Similarity method will infer an incorrect reason for the fault, while the Cosine Similarity method will provide the correct reason for the fault. The distance (0.31) of the observed fault vector from the vector "X_{min}+2%" is the minimum distance in the reference control change X_{min} increase category. This distance is longer than the distance (0.22) to the vector "T_{hl}+2°F". As a result, T_{hl} increase becomes the final conclusion, while the right answer X_{min} increase only occupies the fourth place in the ranking list. This false diagnosis could be avoided if a smaller magnitude of X_{min} is adopted in the library.

The above example shows that the diagnosis results with the Euclidean Distance Similarity method could easily be incorrect if the magnitude setting is unsuitable. That is to say that a proper magnitude setting is essential for the precision of fault diagnosis with the Euclidean Distance Similarity method. However, determining an appropriate magnitude setting for the Euclidean Distance Similarity method is complicated. Because the consumption varies greatly under different levels of fault severity, numerous levels of magnitude should be stored in the library to cover all the possibilities. This work would be redundant and very time-consuming.

Table 61 Cosine similarity between the vectors with the adjacent magnitudes for different types of reference control change for the Bush Academic Building

	I vs II	II vs III	III vs VI	VI vs V
X_{oa} decrease	1.000	1.000	0.999	0.999
X_{oa} increase	0.995	0.995	1.000	0.985
T_{preh} decrease	0.969	0.983	0.990	0.996
T_{preh} increase	1.000	1.000	1.000	1.000
PreHL increase	0.997	0.999	0.999	0.999
T_{cl} decrease	0.993	0.997	0.998	0.998
T_{cl} increase	0.964	0.932	0.952	0.975
T_{hl} decrease	0.999	0.968	0.981	0.996
T_{hl} increase	0.998	1.000	1.000	1.000
HL increase	0.360	0.945	0.992	0.996
X_{min} decrease	0.990	0.991	0.980	0.927
X_{min} increase	0.957	0.975	1.000	1.000
T_{rc} decrease	0.993	0.984	0.992	0.987
T_{rc} increase	0.992	0.985	0.992	0.969
T_{rh} decrease	0.999	0.997	0.998	0.982
T_{rh} increase	0.997	1.000	1.000	0.992
TDL increase	0.990	0.987	0.987	0.996

In contrast, determining suitable magnitude settings for the Cosine Similarity method is simpler. Extensive coverage of the full range of magnitudes is not necessary for every type of reference control change. The levels of magnitude can be condensed, because the direction of the vectors seldom changes for different magnitudes for a particular control change. The data gathered about the cosine similarity between the vectors with the adjacent magnitudes (I vs. II, II vs. III, III vs. IV, and IV vs. V), shown in Table 59, was used to condense the levels of magnitude. The results are presented in Table 61. If the similarity was more than 99%, we determined that the pattern of the two vectors has no difference, and then the larger magnitude would be deducted from the library. In the opposite situation, the larger magnitude was left in the library. For example, for the reference control change "X_{oa} decrease", the cosine similarity values were all higher than 99% in the first row of Table 61, and hence the magnitudes II-V were removed from the library. For the reference control change "X_{oa} increase", the cosine similarity of magnitude V was less than 99% with magnitude VI, and thus the magnitude V remained in the library. The resulting compaction dropped half of the reference control change signature vectors, and the compacted reference control change library is shown on Table 62.

A similar study was made for the Veterinary Research Building. The compacted reference control change library for the Cosine Similarity method application is displayed in Table 63.

Table 62 Compact reference control change library for the cosine similarity method application for the Bush Academic Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%					
2	X_{oa} increase	2%				80%	
3	T_{preh} decrease	-3	-6	-9	-12		°F
4	T_{preh} increase	3					°F
5	PreHL increase	10					kBtu/hr
6	T_{cl} decrease	-2					°F
7	T_{cl} increase	2	4	6	8	10	°F
8	T_{hl} decrease	-2		-12	-18		°F
9	T_{hl} increase	2					°F
10	HL increase	20	40	60			kBtu/hr
11	X_{min} decrease	-2%			-10%	-20%	
12	X_{min} increase	2%	10%	30%			
13	T_{rc} decrease	-1		-4		-10	°F
14	T_{rc} increase	1		4		10	°F
15	T_{rh} decrease	-1				-10	°F
16	T_{rh} increase	1					°F
17	TDL increase	2%		30%	40%		

Table 63 Compact reference control change library for the cosine similarity method application for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Units
		I	II	III	IV	V	
1	X_{oa} decrease	-2%					
2	X_{oa} increase	2%					
3	T_{preh} decrease	-2	-4	-6			°F
4	T_{preh} increase	2		6			°F
5	T_{cl} decrease	-2					°F
6	T_{cl} increase	2	4	6	8	10	°F
7	X_{min} decrease	-2%	-10%	-20%	-30%		
8	X_{min} increase	2%	10%				
9	T_{rc} decrease	-1	-2	-4			°F
10	T_{rc} increase	1		4	6	10	°F
11	T_{rh} decrease	-1				-10	°F
12	T_{rh} increase	1	2				°F

8.1.5 Summary

The previous analysis indicates that the reference control change magnitude setting has an impact on the diagnosis result of both the Euclidean Distance Similarity method and the Cosine Similarity method. This impact is much more significant on the Euclidean Distance Similarity method than on the Cosine Similarity method. If the magnitude setting in the Euclidean distance similarity implementation is unsuitable, it is easy to make the wrong conclusion. Determining an appropriate magnitude setting is a difficult task for the Euclidean Distance Similarity method. In contrast, developing a compact magnitude setting for the Cosine Similarity method is comparatively simple and has been presented in this section. Therefore, the Cosine Similarity method is considered to be superior and more applicable for whole building fault diagnosis. The following sensitivity analysis will only treat the Cosine Similarity method.

8.2 Sensitivity to Calibrated Simulation Model Accuracy

In this section, the performance of the Cosine Similarity method was evaluated for the ideal fault cases and the cases modified with white noise of 3%, 5%, 10%, and 20%. Different levels of white noise stand for different levels of model accuracy. The higher level of white noise approximates lower model accuracy.

8.2.1 Simulated Data Sets

One of the following control changes is assumed to last for one year in the building. The parameter changes are chosen so the maximum monthly average cooling and heating consumption deviation caused by each synthetic control change is 5% of the

yearly average cooling plus heating consumption when there is no fault. The full description of the synthetic control changes is given below.

DDVAV System (Bush Academic Building)

It is assumed one of the following ten synthetic control changes happened on July 1 and continued during the weekdays from July 2008 to June 2009.

1. X_{oa} increase of 2%
2. T_{preh} increase of 7°F
3. The amount of heat leakage from preheat coil (PreHL) is 20kBtu/hr
4. T_{cl} decrease of 1°F
5. T_{hl} increase of 13°F
6. The amount of heat leakage from heating coil (HL) is 40kBtu/hr
7. X_{min} increase of 1.2%
8. T_{rh} increase of 1.6°F
9. T_{rc} decrease of 1.2°F
10. Terminal box damper leakage (TDL) increase of 4%

SDVAV System (Veterinary Research Building)

It is assumed one of the following six synthetic control changes happened in January 1 and continued during the period from January 1 to December 31, 2000:

1. X_{oa} increase of 3.5%
2. T_{preh} increase of 1.8°F
3. T_{cl} decrease of 0.8°F
4. X_{min} increase of 2.5%

5. T_{rh} increase of 1.2°F

6. T_{rc} decrease of 1°F

The ideal abnormal data were produced by the calibrated simulation model of ABCAT. The ideal abnormal data sets and the modified data sets were comparable to observed fault data samples. The reference control change libraries used for the two studied buildings are shown in Tables 8.1 and 8.2.

8.2.2 Diagnostic Results

The diagnosis results with all the data sets for the two buildings analyzed are presented in Tables 64 and 65. The column "0%" shows the diagnosis results for the ideal cases. Character "Y" indicates that the diagnosis result agrees with the synthetic setting, and thus the performance of the Cosine Similarity method is successful for that case. Character "N" indicates that the diagnosis result disagrees with the synthetic setting. The number in the parenthesis indicates the position of the correct answer in the ranking of all the reference control changes. For example, number "2" in Table 64 suggests the real fault reason ranks second on the ranking list. In other words, the correct fault reason is given the second highest probability as the fault cause among the 17 different types of reference control change.

Table 64 shows that the Cosine Similarity method produces excellent results for the ideal cases or for data sets modified with 3%/5% white noise in the study of the Bush Academic Building. In three out of seven unsuccessful cases at the higher noise levels, the correct answer ranks second among all types of reference control change. As shown in Table 65, for the Veterinary Research building, in all ideal cases and all the cases

where data sets modified with 3%/5%/10% white noise, the fault reason identified by the Cosine Similarity method is consistent with the synthetic setting. The correct answer ranks second among all types of reference control change in the two failed test cases with 20% white noise.

Table 64 Diagnosis results with cosine similarity method on synthetic data modified with white noise for the Bush Academic Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+2%	Y	Y	Y	Y	Y
T_{preh}+7°F	Y	Y	Y	N (2)	N (9)
T_{cl}-1°F	Y	Y	Y	Y	N (4)
T_{hl}+13°F	Y	Y	Y	Y	N (2)
PreHL+20kBtu/hr	Y	Y	Y	Y	Y
X_{min}+1.2%	Y	Y	Y	Y	N (4)
HL+40kBtu/hr	Y	Y	Y	Y	N (2)
T_{rc}-1.2°F	Y	Y	Y	Y	Y
T_{rh}+1.6°F	Y	Y	Y	Y	N (4)
TDL+4%	Y	Y	Y	Y	Y

Table 65 Diagnosis results with cosine similarity method on synthetic data modified with white noise for the Veterinary Research Building

	0%	3%	5%	10%	20%
X_{oa}+3.5%	Y	Y	Y	Y	N (2)
T_{preh}+1.8°F	Y	Y	Y	Y	Y
T_{cl}-0.8°F	Y	Y	Y	Y	Y
X_{min}+2.5%	Y	Y	Y	Y	Y
T_{rc}-1°F	Y	Y	Y	Y	N (2)
T_{rh}+1.2°F	Y	Y	Y	Y	Y

8.2.3 Discussion

The test case is considered successful where the Cosine Similarity method isolates the correct fault reason. The term “diagnostic accuracy” for a particular white noise level is defined as the ratio of the number of successful cases to the total number of cases. There are ten test cases and six test cases respectively for the studies of the Bush Academic Building and Veterinary Research Building at each white noise level. The diagnostic accuracy is shown as bars for easy comparison. There is an overall trend of the accuracy getting worse at higher white noise levels in Figures 58 and 59. Higher levels of white noise, corresponding to lower model accuracy, reduce the precision of the results of fault diagnosis.

The more white noise found in the sample, the more variation from the ideal observed fault signature found in the modified fault signature. The variation in the signature pattern widens as the white noise increases. Under extreme conditions, the characteristics of the modified observed fault signature would only show the characteristics of the white noise, and no longer represent the characteristics of the ideal signature. Thus, it is most likely that the Cosine Similarity method induces the wrong conclusion.

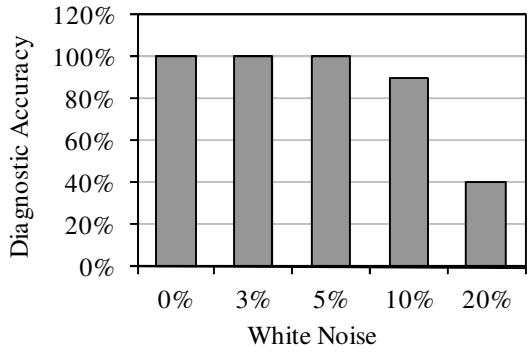

Figure 58 Diagnostic accuracy of the cosine similarity method on synthetic data modified with different levels of white noise for the Bush Academic Building

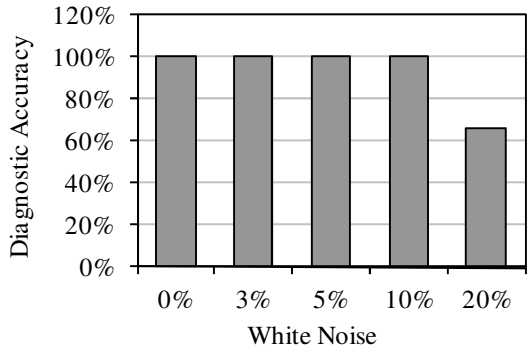

Figure 59 Diagnostic accuracy of the cosine similarity method on synthetic data modified with different levels of white noise for the Veterinary Research Building

Figure 60 compares the patterns of the observed fault signature of the ideal case "T_{preh}+7F" and the modified case "T_{preh}+7F with WN 20%". It is obvious that the pattern of the CHW part of the observed fault signature components is completely different after adding the white noise.

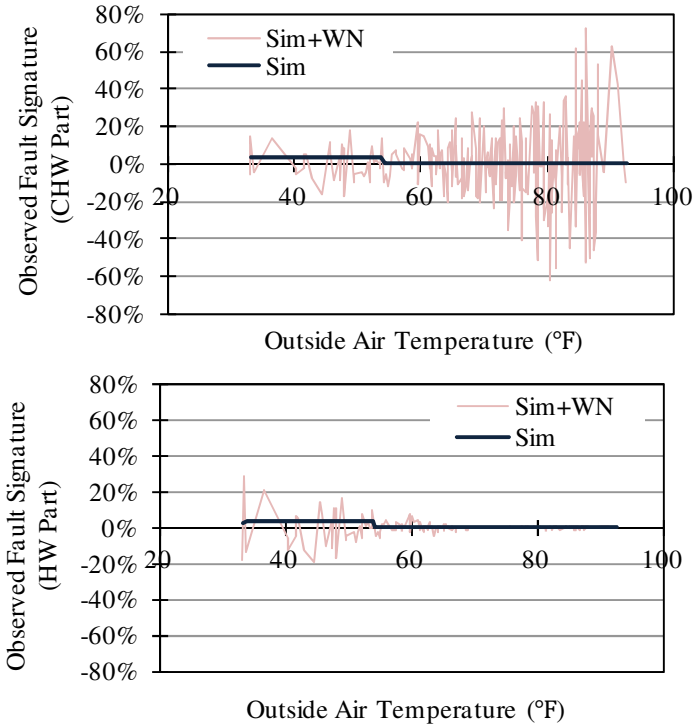

Figure 60 The observed fault signature plotted as a function of outside air temperature in the test case "T_{preh}+7°F with WN 20%" for the Bush Academic Building

8.2.4 Summary

The diagnostic accuracy of the Cosine Similarity method is influenced by the calibration simulation model accuracy. For the identical control change, low model accuracy reduces the precision of the fault diagnosis result, as the low model accuracy modifies the characteristics of the fault pattern. The diagnostic accuracy decreased from 100% to 40% when the white noise level was increased from 0% to 20% in the Bush Academic Building, and declined from 100% to 67% when the white noise level was increased from 0% to 20% in the Veterinary Research Building.

8.3 Sensitivity to Fault Severity

8.3.1 Simulated Data Sets

The diagnosis results of three fault groups with different levels of fault severity are assessed in this section. The maximum monthly average energy use change index is 2%, 5%, and 10% respectively for groups one, two, and three. Table 66 summarizes the description of the three fault groups for the Bush Academic Building and the Veterinary Research Building. The performance of the Cosine Similarity method was evaluated for the ideal fault cases and the modified cases with white noise levels of 3%, 5%, 10%, and 20% during the analysis of each fault group.

Table 66 Description of faults assumed in the fault severity sensitivity study

Bush Aademic Building			Veterinary Research Building		
Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
X_{oa}+0.8%	X_{oa}+2%	X_{oa}+3.9%	X_{oa}+1.5%	X_{oa}+3.5%	X_{oa}+6.5%
T_{preh}+3°F	T_{preh}+7°F	T_{preh}+14°F	T_{preh}+0.8°F	T_{preh}+1.8°F	T_{cl}-1.5°F
T_{cl}-0.5°F	T_{cl}-1°F	T_{cl}-2°F	T_{cl}-0.4°F	T_{cl}-0.8°F	T_{preh}+3.3°F
T_{hl}+4°F	T_{hl}+13°F	T_{hl}+30°F	X_{min}+1%	X_{min}+2.5%	X_{min}+4.9%
PreHL+10kBtu/hr	PreHL+20kBtu/hr	PreHL+40kBtu/hr	T_{rc}-0.4°F	T_{rc}-1°F	T_{rc}-1.7°F
X_{min}+0.5%	X_{min}+1.2%	X_{min}+2.5%	T_{rh}+0.7°F	T_{rh}+1.2°F	T_{rh}+2.3°F
HL+30kBtu/hr	HL+40kBtu/hr	HL+60kBtu/hr			
T_{rc}-0.5°F	T_{rc}-1.2°F	T_{rc}-2.2°F			
T_{rh}+0.7°F	T_{rh}+1.6°F	T_{rh}+3°F			
TDL+1.5%	TDL+4%	TDL+7.5%			

8.3.2 Diagnostic Results

The diagnosis results of the fault groups one and three are presented in Tables 67 and 68 for the Bush Academic Building, and in Tables 69 and 70 for the Veterinary Research Building. The results of fault group two are presented in Tables 64 and 65.

Table 67 Diagnosis results of Fault Group 1 for the Bush Academic Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+0.8%	Y	Y	Y	Y	Y
T_{preh}+3°F	Y	Y	N (3)	N (9)	N (4)
T_{cl}-0.5°F	Y	Y	Y	N(4)	N(10)
T_{hl}+4°F	Y	Y	Y	N(5)	N(10)
PreHL+10kBtu/hr	Y	Y	Y	Y	Y
X_{min}+0.5%	Y	Y	Y	N(4)	N(8)
HL+30kBtu/hr	Y	Y	Y	N(4)	N(5)
T_{rc}-0.5°F	Y	Y	Y	Y	Y
T_{rh}+0.7°F	Y	Y	Y	N (7)	N (4)
TDL+1.5%	Y	Y	Y	N (2)	N (6)

Table 68 Diagnosis results of Fault Group 3 for the Bush Academic Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+3.9%	Y	Y	Y	Y	Y
T_{preh}+14°F	Y	Y	Y	Y	N (2)
T_{cl}-2°F	Y	Y	Y	Y	Y
T_{hl}+30°F	Y	Y	Y	Y	Y
PreHL+40kBtu/hr	Y	Y	Y	Y	Y
X_{min}+2.5%	Y	Y	Y	Y	Y
HL+60kBtu/hr	Y	Y	Y	Y	Y
T_{rc}-2.2°F	Y	Y	Y	Y	Y
T_{rh}+3°F	Y	Y	Y	Y	Y
TDL+7.5%	Y	Y	Y	Y	Y

Table 69 Diagnosis results of Fault Group 1 for the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+1.5%	Y	Y	Y	N(2)	N(2)
T_{preh}+0.8°F	Y	Y	Y	Y	N(2)
T_{cl}-0.4°F	Y	Y	Y	N(2)	N(4)
X_{min}+1%	Y	Y	Y	Y	N(6)
T_{rc}-0.4°F	Y	Y	Y	N(2)	N(2)
T_{rh}+0.7°F	Y	Y	Y	Y	N(2)

Table 70 Diagnosis results of Fault Group 3 for the Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa}+6.5%	Y	Y	Y	Y	Y
T_{cl}-1.5°F	Y	Y	Y	Y	Y
T_{preh}+3.3°F	Y	Y	Y	Y	Y
X_{min}+4.9%	Y	Y	Y	Y	Y
T_{rc}-1.7°F	Y	Y	Y	Y	Y
T_{rh}+2.3°F	Y	Y	Y	Y	Y

Figures 61 and 62 show that in the identical fault group, the increase in the white noise decreased the diagnostic accuracy, which is consistent with the conclusion in the previous section. If the white noise level is uniform, the diagnostic accuracy would be improved with an increase of fault severity. For instance, when the white noise level is set at 20%, the diagnostic accuracy is 0% for group one, 67% for group two, and 100% for group three in the Veterinary Research Building. The increase in fault severity makes the pattern of the observed fault signature more prominent when compared to the pattern of the white noise. In other words, the distinguishing characteristics of the signature are harder to be covered by the characteristics of white noise when the fault severity

increases. A large-scale fault helps the Cosine Similarity method draw the right diagnostic conclusion.

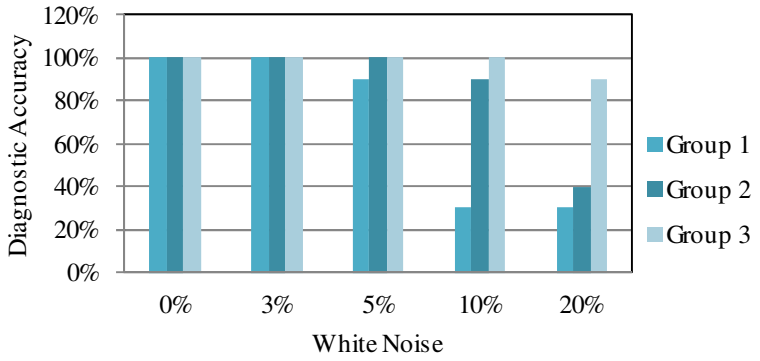

Figure 61 Comparison of diagnostic accuracy of the three fault groups with different fault severity levels for the Bush Academic Building

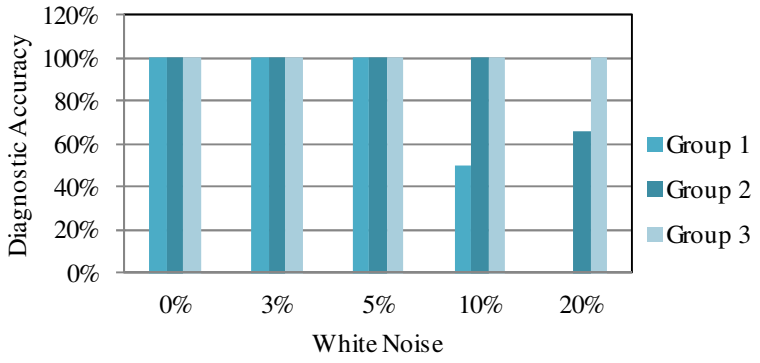

Figure 62 Comparison of diagnostic accuracy of the three fault groups different fault severity levels for the Veterinary Research Building

8.3.3 Small Faults That the Cosine Similarity Method can Successfully Diagnose

Because high fault severity lessens the impact of white noise, the minimum fault severity for which the Cosine Similarity method can give the correct answer for each type of control change for the two buildings analyzed is summarized in Tables 71 and 72 for the two buildings analyzed. The investigation interval is 1°F for the control changes "T_{preh}+", "T_{cl}-", "T_{hl}+", "T_{rc}-", and "T_{rh}+", 1% for the control changes "X_{oa}+", "X_{min}+" and "TDL+", and 10kBtu/hr for the control changes "PreHL+" and "HL+".

The indexes of annual energy impact (AEI) of the faults in Table 71 are listed in Table 73, while the indexes for Table 72 are listed in Table 74. For the Bush Academic Building, the minimum/maximum/median AEI values are 0%/2.6%/0.8% when WN is 0%, and 0.5%/6.7%/2.5% when WN is 20%. For the Veterinary Research Building, the minimum/maximum/median AEI values are 0.5%/5.0%/0.9% when WN is 0%, and are 0.5%/5.0%/2.8% when WN is 20%.

Table 71 Summary of the minimum fault severity for different control changes where the cosine similarity method can isolate the correct fault reason: Bush Academic Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa} +	1%	1%	1%	1%	1%
T_{preh} +	1°F	3°F	4°F	8°F	15°F
T_{cl} -	1°F	1°F	1°F	1°F	2°F
T_{hl} +	1°F	2°F	3°F	7°F	15°F
PreHL+	10kBtu/hr	10kBtu/hr	10kBtu/hr	10kBtu/hr	10kBtu/hr
X_{min} +	1%	1%	1%	2%	3%
HL+	20kBtu/hr	30kBtu/hr	30kBtu/hr	40kBtu/hr	50kBtu/hr
T_{rc} -	1°F	1°F	1°F	1°F	1°F
T_{rh} +	1°F	1°F	1°F	2°F	3°F
TDL+	1%	1%	1%	2%	4%

Table 72 Summary of the minimum fault severity for different control changes where the cosine similarity method can isolate the correct fault reason: Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa} +	1%	1%	1%	2%	4%
T_{preh} +	1°F	1°F	1°F	1°F	1°F
T_{cl} -	1°F	1°F	1°F	1°F	1°F
X_{min} +	1%	1%	1%	1%	2%
T_{rc} -	1°F	1°F	1°F	1°F	2°F
T_{rh} +	1°F	1°F	1°F	1°F	1°F

Table 73 Summary of the AEI for the smallest faults where the cosine similarity method can isolate the correct fault reason: Bush Academic Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa} +	0.5%	0.5%	0.5%	0.5%	0.5%
T_{preh} +	0.1%	0.4%	0.6%	1.2%	2.2%
T_{cl} -	2.6%	2.6%	2.6%	2.6%	5.6%
T_{hl} +	0.3%	0.5%	0.7%	1.5%	2.8%
PreHL+	2.0%	2.0%	2.0%	2.0%	2.0%
X_{min} +	2.0%	2.0%	2.0%	4.2%	6.7%
HL+	0.0%	0.7%	0.7%	2.0%	3.7%
T_{rc} -	1.2%	1.2%	1.2%	1.2%	1.2%
T_{rh} +	0.6%	0.6%	0.6%	1.3%	2.0%
TDL+	0.9%	0.9%	0.9%	1.8%	3.7%
Max	2.6%	2.6%	2.6%	4.2%	6.7%
Median	0.8%	0.8%	0.8%	1.7%	2.5%
Min	0.0%	0.4%	0.5%	0.5%	0.5%

Table 74 Summary of the AEI for the smallest faults where the cosine similarity method can isolate the correct fault reason: Veterinary Research Building

	White Noise				
	0%	3%	5%	10%	20%
X_{oa} +	0.5%	0.5%	0.5%	1.1%	2.1%
T_{preh} +	0.5%	0.5%	0.5%	0.5%	0.5%
T_{cl} -	5.0%	5.0%	5.0%	5.0%	5.0%
X_{min} +	1.7%	1.7%	1.7%	1.7%	3.4%
T_{rc} -	0.7%	0.7%	0.7%	0.7%	3.4%
T_{rh} +	1.0%	1.0%	1.0%	1.0%	1.0%
Max	5.0%	5.0%	5.0%	5.0%	5.0%
Median	0.9%	0.9%	0.9%	1.0%	2.8%
Min	0.5%	0.5%	0.5%	0.5%	0.5%

8.3.4 Summary

The diagnostic accuracy of the Cosine Similarity method under three levels of fault severity is compared in this section. The results show that the diagnostic accuracy increases when the level of fault severity increases. The smallest faults for which the Cosine Similarity method could classify the correct fault reason in the two buildings and their related energy consumption impact statistics were also summarized in this section.

8.4 Sensitivity to the Fault Period Length

The implementation of the Cosine Similarity method required that some fault detection mechanism had already determined that an abnormal consumption fault was present and had persisted for a certain time. The fault period could have lasted from several days to several months. This section deals with the influence of the length of the fault period on the diagnostic accuracy of the Cosine Similarity method.

The Cosine Similarity method is applied to two fault groups with different fault period lengths of one month and three months. The fault description of the two fault

groups are the same, as listed in the column “Group 3” in Table 66. The information of the fault periods was summarized in Table 75. The performance of the Cosine Similarity method was evaluated for the ideal fault cases and the modified cases with the white noise level at 10%.

Table 75 Fault periods of the two fault groups analyzed

Group 1 (One Month)	Group 2 (Three Months)
January	January-March
April	April-June
July	July-September
October	October-December

Figures 63 and 64 illustrate the results of the diagnostic accuracy. The results show that the Cosine Similarity method predicted the correct conclusion in all of the ideal cases, whereas the diagnostic accuracy of fault group two was generally higher than that of fault group one when the white noise level was raised to 10%. This suggests that the diagnostic accuracy of the Cosine Similarity method is greater with the longer fault period. In summary, the longer fault period length helps for the Cosine Similarity method isolate the correct fault reason.

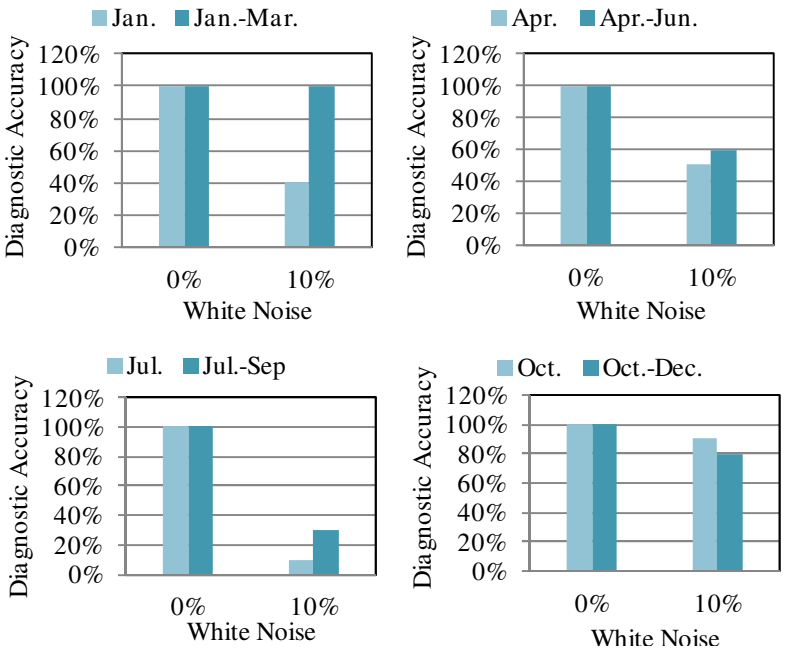

Figure 63 Comparison of the diagnostic accuracy of two fault groups with different fault period lengths for the Bush Academic Building

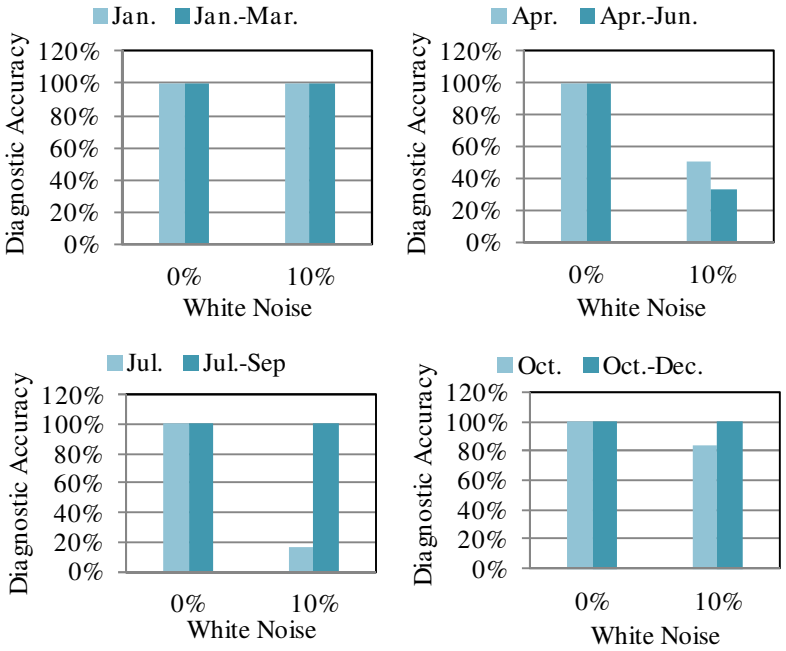

Figure 64 Comparison of the diagnostic accuracy of two fault groups with different fault period lengths for the Veterinary Research Building

9. CONCLUSIONS

This dissertation develops new whole building HVAC fault detection and fault diagnosis approaches to aid commissioning persistence. It also demonstrates the effectiveness and assesses the sensitivity of these new approaches through tests using simulated faults and in real buildings.

9.1 Fault Detection Approaches

Two statistical approaches called the Days Exceeding Threshold – Date (DET-Date) method and the Days Exceeding Threshold – Outside Air Temperature (DET-Toa) method are proposed to detect abnormal whole building cooling or heating energy consumption. Both the DET-Date and DET-Toa methods define a fault as identified when the absolute deviation between the measured and simulated consumption is greater than $SD_{\_baseline}$ (as determined from the statistics of the calibrated simulation) for 20 days. The difference is that in the DET-Date method the fault-flag days must be consecutive in time, while in the DET-Toa method the fault-flag days must be consecutive in outside air temperature.

In the simulation test, the fault detection accuracy for the DET-Date and DET-Toa methods are 55% and 95% respectively in the 20 test cases. The number of the synthetic faults detected by the DET-Toa method is greater than the number of the faults detected by the DET-Date method. In addition, in the cases where both methods have detected faults, the DET-Toa method detected the faults earlier. Hence, the DET-Toa method is considered more reliable and is recommended for future use. The results of the field tests verified this conclusion.

The results of the sensitivity analysis on the DET-Toa method show that the detection accuracy increases with high calibrated simulation model accuracy and large-scale control changes. The impact of the time of fault occurrence depends on the operating conditions that maximize the impact of the fault. The abnormal consumption is more likely to be detected quickly if the control change coincidently happens in the season when that control change significantly increases/decreases the energy consumption.

9.2 Fault Diagnosis Approaches

Two approached called the Cosine Similarity method and the Euclidean Distance Similarity method proposed to diagnose abnormal whole building cooling or heating energy consumption faults were described in Section 3. In these two approaches, a reference control change library collection of known whole building faults is determined in advance. The cosine similarity/Euclidean distance similarity within the observed fault signature vectors and reference control change signature vectors are calculated. A larger similarity value suggested a greater probability that the corresponding reference control change is the cause of the observed fault.

In the simulation test, for the DDVAV system building, the fault diagnosis results using the Cosine Similarity method are consistent with the synthetic control changes in all of the ten simulation cases, and the diagnosis results using the Euclidean Distance Similarity method matches the synthetic control changes in nine of the ten simulation cases. The false diagnosis given by the Euclidean Distance Similarity method is due to the improper setting of the magnitude level in the reference control change

library. For the SDVAV system building, the fault diagnosis results of both methods matched the stated synthetic control changes perfectly. The capabilities of the two methods are also proved in the two field test cases.

Large differences in the sensitivity to the reference control change magnitude setting were observed for the Euclidean Distance Similarity method and the Cosine Similarity method. The Euclidean distance similarity varies greatly when the control change magnitude was changed; therefore, it is easy to reach the wrong conclusion if the magnitude setting is unsuitable in the Euclidean distance similarity implementation. Furthermore, figuring out an appropriate magnitude setting is a difficult task for the Euclidean Distance Similarity method. Therefore, the Cosine Similarity method is considered to be superior and more useful for the whole building fault diagnosis. The results of the sensitivity analysis on the Cosine Similarity method also suggest that the diagnostic accuracy would be improved by the high calibrated simulation model accuracy, large-scale control changes, and long fault periods.

9.3 Summary of Recommended Whole Building FDD Procedure

The following procedure is recommended to use for the detection of abnormal energy consumption and the identification of possible causes.

Step 1: Establish calibrated simulation model.

a. Choose a baseline period from a post-commissioning time period when the building's operation is considered to be optimal.
b. Establish and calibrate a building energy simulation model based on the building cooling and heating consumption in the baseline period.

Step 2: Perform fault detection using the DET-Toa method.

a. Collect building energy consumption data in an investigated period.
b. Use DET-Toa method to detect is there abnormal cooling or heating energy consumption in the investigated period.
c. If the abnormal cooling or heating energy consumption is detected in the investigated period, use step 3 to identify the possible causes.

Step 3: Perform fault diagnosis using the Cosine Similarity method.

a. Determine a reference control change library including the known whole building level faults.
b. Determine the fault period. The first day in the fault period is the the earliest day in the 20 fault-flag days of the DET-Toa method. The last day in the fault period is the last day in the investigated period.
c. Use the calibrated simulation model to produce the expected cooling and heating consumption when there is a known control change from the reference library persisting during the fault period.
d. Generate the observed fault signature vector and reference control change signature vectors.
e. Calculate the cosine similarity values between the observed fault signature vector and each of the reference control change signature vectors.
f. Sort different types of reference control changes by representative similarity in descending order. The ranking indicated the probability that the reference control change is the cause of the observed fault. The reference control

change having the largest similarity value is the most probable reason of the observed abnormal energy consumption.

9.4 Future Work

While this study developed good methods for whole building fault detection and diagnosis, more research is needed in the future to investigate the FDD capabilities of the proposed approaches. The recommendations for the future study include:

1. The use of 20 days is a conservative time limit for the DET-Toa method to determine if the energy consumption is abnormal. Further testing is necessary to examine the suitability of this time limit.
2. The use of the proposed methods on other HVAC systems should be investigated.
3. It is valuable to apply the proposed FDD methods in a large number of buildings to further validate their effectiveness.

REFERENCES

Borsjie, H.J. 1999. Fault Detection in Boilers Using Canonical Variate Analysis, Proceeding of the American Control Conference 2: 1167-1170.

Candan, K. S. and M. L. Sapino. 2010. Data Management for Multimedia Retrieval, Cambridge University Press.

Claridge, D.E., M. Liu, and W.D. Turner. 1999. Whole Building Diagnostics Building Diagnostics. Proceedings of the Workshop on Diagnostics for Commercial Buildings: Research to Practice, San Francisco, CA, June 16-17: 3.2.1-3.2.17.

Claridge, D.E., M. Liu, G. Wei, S. U. Lee, and N. Bensouda. 2001. Manual of Procedures for Calibrating Simplified Simulations of Building Systems, Energy Systems Laboratory, Texas A&M University, College Station, TX, July.

Claridge, D. E., B. John, and G. Lin. 2009. Final Report on Retrospective Testing and Application of an Automated Building Commissioning Analysis Tool, Texas A&M University, College Station, TX.

Comstock, M. C. and J. E. Braun, 1999b, Development of Analysis Tools for the Evaluation of Fault Detection and Diagnostics in Chillers, Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN.

Curtin. J. M. 2007. Development and Testing of an Automated Building Commissioning Analysis Tool (ABCAT), MS Thesis, Texas A&M University, College Station, TX.

Elyas, R., S. Iman, and R. Kumars. 2009. Application of Data Mining on Fault Detection and Prediction in Boiler of Power Plant Using Artificial Neural Network, 2nd

International Conference on Power Engineering, Energy and Electrical Drives Proceeding: 473-478.

Energy Information Administration (EIA). 2009. Annual Energy Review 2009, Report DOE/EIA-0384(2009), U.S. Department of Energy, Washington, DC.

Fasolo, P. and D.E. Seborg. 1994. An SQC Approach to Monitoring and Fault Detection in HVAC Control Systems, Proceeding of the American Control Conference, Baltimore, Maryland.

Friedman, H. and M. A. Piette. 2001. Comparison of Emerging Diagnostic Tools for Large Commercial HVAC Systems. Proceedings of the 9th National Conference on Building Commissioning, Cherry Hill, NJ.

Friedman, H., A. Potter, T. Haasl, and D. Claridge. 2003. Report on Strategies for Improving Persistence of Commissioning Benefits, Final Report.

Grimmelius, H. T., J. K. Woud and G. Been. 1995. On-Line Failure Diagnosis for Compression Refrigeration Plants, International Journal of Refrigeration 18(1): 31-41.

Haberl, J. S. and S. Cho. 2004. Literature Review of Uncertainty of Analysis Methods (Inverse Model Toolkit), Energy Systems Lab ESL-TR-04/10-03, October.

Haberl, J. S. and D.E. Claridge. 1987. An Expert System for Building Energy Consumption Analysis: Prototype Results, ASHRAE Transactions 93(1):979-998.

Haberl, J. S. and E. J. Vajda. 1988. Use of Metered Data Analysis to Improve Building Operation and Maintenance: Early Results from Two Federal Complexes, Proceedings of the ACEEE Summer Study on Energy Efficiency in Building, Asilomar, California.

Haberl J. S., L.K. Norford, and J. V. Spadaro. 1989. Diagnosing building operating problems, ASHRAE Journal, June: 20-30.

Haberl, J. S. 1992. The Use of a Monthly Whole-Campus Energy Analysis for Evaluating a Third Party Energy Service Agreement, Proceedings ACEEE Summer Study on Energy Efficiency in Buildings 3(1): 3.95-3.110.

Han, C.Y., Y. Xiao, and C.J. Ruther. 1999. Fault Detection and Diagnosis of HVAC Systems. ASHRAE Transactions 105(1): 568-578.

Haves, P., and S.K. Khalsa. 2000. Model-Based Performance Monitoring: Review of Diagnostics Methods and Chiller Case Study, Proceeding of ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA, August.

Hoaglin. I. 1993. How to Detect and Handle Outliers, American Society for Quality, Milwaukee.

House, J. M., H. Vaezi-Nejad and J. M. Whitcomb. 2001. An Expert Rule Set for Fault Detection in Air-Handling Units. ASHRAE Transactions 107(1): 858-871.

IEA. 1996. IEA Annex 25, Building optimization and fault diagnosis source book, Eds. J. Hyvärinen and S. Kärki. Espoo, Finland: Technical Research Centre of Finland.

IEA. 2001. IEA Annex 34, Demonstrating automated fault detection and diagnosis methods in real buildings. Ed. L. Ukskoski, Espoo, Finland: Technical Research Centre of Finland.

Huang H.,C. Li, and J. Jeng. 2007. Multiple Multiplicative Fault Diagnosis for Dynamic Processes via Parameter Similarity Measures, Ind. Eng. Chem. Res. 46:4517-4570.

Kabir. M. 2009. Similarity Matching Techniques for Fault Diagnosis in Automotive Infotainment Electronics. International Journal of Computer Science Issues 3:14-19.

Katipamula, S., N. Bauman, R.G. Pratt, and M.R. Brambley. 2003. Demonstration of the Whole-Building Diagnostician in a Single-Building Operator Environment, Final Report, Prepared for the U.S. Department of Energy.

Katipamula, S. and M.R. Brambley. 2005a. Methods for Fault Detection, Diagnostics and Prognostics for Building Systems- A Review, Part I, IJHVAC&R Research 11(1):3-25.

Katipamula, S. and M.R. Brambley. 2005b. Methods for Fault Detection, Diagnostics and Prognostics for Building Systems- A Review, Part II, IJHVAC&R Research 12(2):169-187.

Knebel, D.E. 1983. Simplified Energy Analysis Using the Modified Bin Method. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta.

Lee. S. U., F. L. Painter, D. E. Claridge. 2007. Whole-Building Commercial HVAC System Simulation for Use in Energy Consumption Fault Detection, ASHRAE Transactions 113(2): 52-61.

Lee W., H.W. Shen, and G. Zhang. 2009. Research on Fault Diagnosis of Turbine Based on Similarity Measures between Interval-valued Intuitionistic Fuzzy Sets, International Conference on Measuring Technology and Mechatronics Automation, China.

Lee, W.-Y., C. Park, J. M. House and G. E. Kelly. 1996. Fault Diagnosis of an Air Handling Unit Using Artificial Neural Networks, ASHRAE Transactions, 102(1): 540-549.

Li, J. and W. Dai. 2005. Research on Fault Diagnosis for Oil-immersed Transformer Based on Included Angle Cosine, China Academic Journal Electronic Publishing House 26(12):1302-1304.

Liu, M., D. E. Claridge, and W. D. Turner. 2003. Continuous Commissioning^{SM} of Building Energy Systems, Transactions of the ASME. Journal of Solar Energy Engineering 125(3): 275-281.

Liu, M., L. Song, G. Wei, and D. E. Claridge. 2004. Simplified Building and Air-handling Unit Model Calibration and Applications. Transactions of the ASME. Journal of Solar Energy Engineering 126(1); 601-609.

McGill. M., M. Koll, and T. Noreault. 1979. An Evaluation of Factors Affecting Document Ranking By Information Retrieval Systems, Syracuse, NY: School of Information Studies, Syracuse University

Mills, E., N. Bourassa, M. A. Piette, H. Friedman, T.Haasl, T. Powell and D. E. Claridge. 2005. The Cost- Effectiveness of Commissioning, HPAC Heating, Piping, Air Conditioning Engineering 77(10): 20-24.

Montgomery,D. C. 1985. Introduction to Statistical Quality Control, New York, Wiley

NIST/SEMATECH. 2003. E-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/.

Norford, L.K., J.A. Wright, R.A. Buswell, D. Luo, C.J. Klaassen, and A. Suby. 2002. Demonstration of fault detection and diagnosis methods for air-handling units (ASHRAE 1020-RP), International Journal of Heating, Ventilating, Air Conditioning and Refrigerating Research, 8(1):41-71.

Peterson, J. 2005. Evaluation of Retrocommissioning Results after Four Years: A Case Study, Proceedings of the National Conference on Building Commissioning, Portland Energy Conservation, Inc., New York.

Portland Energy Conservation, Inc. 2003. Methods for Automated and Continuous Commissioning of Building Systems, Final Report.

Reddy, T. A., D. Niebur, K. Anderson, P. Pericolo and G. Cabrera. 2003. Evaluation of the Suitability of Different Chiller Performance Models for On-Line Training

Applied to Automated Fault Detection and Diagnosis (RP-1139), HVAC&R Research 9(4): 385-414.

Rossi, T. M. and J. E. Braun. 1997. Statistical, Rule-Based Fault Detection and Diagnostic Method for Vapor Compression Air Conditioners, HVAC&R Research 3(1): 19-37.

Seem, J.E. 2005. Pattern Recognition Algorithm for Determining Days of the Week with Similar Energy Consumption Profiles, Energy and Buildings 37:127-139.

Seem, J.E. 2007. Using Intelligent Data Analysis to Detect Abnormal Energy Consumption in Buildings, Energy and Buildings 39:52-58.

Shepard, R. N., 1987, Toward a Universal Law of Generalization for Psychological Science, Science237: 1317-1323.

Song,Y., Y. Akashi, J. Yee. 2008. A Development of Easy-to-use Tool for Fault Detection and Diagnosis in Building Air-conditioning Systems, Energy and Buildings 40: 71-82.

TIAX LLC. 2005. Energy Impact of Commercial Building Controls and Performance Diagnostics: Market Characterization, Energy Impact of Building Faults and Energy Savings Potential, Report D0180, Prepared for the DOE Building Technologies Program.

Toole, C.D. and Claridge, D., "The Persistence of Retro-Commissioning Savings in Ten University Buildings," Proc. 11th International Conference for Enhanced Building Operation, October 18-20, 2011, New York, NY.

Wang, S. and J. Qin. 2005. Sensor Fault Detection and Validation of VAV Terminals in Air Conditioning Systems, Energy Conversion and Management 46(15-16): 2482-2500.

Xu, P. and P. Haves. 2002. Field Testing of Component Level Model Based Fault Detection Methods for Mixing Boxes and VAV Fan Systems, California Energy Commission PIER Program.

Yoon , S. and J. F. MacGregor. 2001. Fault Diagnosis with Multivariate Statistical Models Part I: Using Steady State Fault Signatures, Journal of Process Control 11:387-400.

Yu, B., and H. C. Dolf. 2003. Fuzzy Neural Networks Model for Building Energy Diagnosis, The eighth International IBSPA Conference, Eindhoven, Netherlans, August 11-14.

APPENDIX A

SUBROUTINES OF THE DET-TOA PROGRAM

The VBA codes used in the DET-Toa program are given in this appendix.

Module mMain

```
Option Explicit
Public CHWMean As Variant
Public HWMean As Variant
Public CHWSigma As Variant
Public HWSigma As Variant
Public BDdata(1000, 10) As Variant 'Ddata sort by Toa ascending order
Public Ddata(1000, 6) As Variant ' include date, Toa, CHW and HW fault index,
CHW and HW residuals
Public TDdata(1000, 6) As Variant 'include date, Toa, CHW and HW fault index,
CHW and HW residuals
Public StartRow As Integer
Public EndRow As Integer
Public Timelimit As Integer ' Time limit in the criterion (days)
Public RenewTime As Integer 'Transition matrix renew frequency (days)
Public NumberofDays As Integer ' Number of days in transition matrix
Public RunTime As Integer
Public OldNumberofdays As Integer
Public TotalNumberofDays As Integer 'The total number of days in sheet
"data"(Number of days in fault detection matrix)

Sub Main()
Dim i As Integer
Dim j As Integer
Dim w As Integer
Dim z As Integer
Dim Tdata(1, 10) As Variant 'transition variable

Erase Ddata
Erase BDdata

'-----------------------------------------------------------------------------------
'Indicate calculation beginning and end rows, database renew time, and time limit
'-----------------------------------------------------------------------------------
Sheets("Main").Activate
Range("c2").Activate
RenewTime = ActiveCell.Offset(0, 0)
Timelimit = ActiveCell.Offset(1, 0)
```

```
        StartRow = ActiveCell.Offset(2, 0)
        EndRow = ActiveCell.Offset(3, 0)
        CHWMean = ActiveCell.Offset(4, 0)
        HWMean = ActiveCell.Offset(5, 0)
        CHWSigma = ActiveCell.Offset(6, 0)
        HWSigma = ActiveCell.Offset(7, 0)

        Range("G3:J4").Activate
        Selection.ClearContents

        TotalNumberofDays = EndRow - StartRow + 1 'calculated total number of days
in sheet "data"

        i = 0
        j = 0
        w = 0
        z = 0
        Sheets("Data").Activate
        Range("a6").Activate  'Selects first cell in Column A
        'Input date, Toa, CHW and HW Res Labels
        For i = 0 To TotalNumberofDays - 1
           TDdata(i, 0) = ActiveCell.Offset(i, 0) 'date
           TDdata(i, 1) = ActiveCell.Offset(i, 1) 'Toa
           TDdata(i, 2) = ActiveCell.Offset(i, 10) 'CHW fault index
           TDdata(i, 3) = ActiveCell.Offset(i, 11) 'HW fault index
           TDdata(i, 4) = ActiveCell.Offset(i, 2) 'CHW residual
           TDdata(i, 5) = ActiveCell.Offset(i, 6) 'HW residual
        Next i

        OldNumberofdays = 0
        NumberofDays = Timelimit
        Do While OldNumberofdays < TotalNumberofDays
           'Determine the period length running "CHW Detection"(if transition
matrix=fault detection matrix)
           If NumberofDays < TotalNumberofDays Then
              RunTime = NumberofDays
           Else
              RunTime = TotalNumberofDays
           End If
              For i = OldNumberofdays To RunTime - 1
                 Ddata(i, 0) = TDdata(i, 0)
                 Ddata(i, 1) = TDdata(i, 1)
                 Ddata(i, 2) = TDdata(i, 2)
                 Ddata(i, 3) = TDdata(i, 3)
```

```
                Ddata(i, 4) = TDdata(i, 4)
                Ddata(i, 5) = TDdata(i, 5)
                BDdata(i, 0) = Ddata(i, 0)
                BDdata(i, 1) = Ddata(i, 1)
                BDdata(i, 2) = Ddata(i, 2)
                BDdata(i, 3) = Ddata(i, 3)
                BDdata(i, 4) = 0 'sum of CHW fault index
                BDdata(i, 5) = 0 'sum of HW fault index
                BDdata(i, 6) = 0 'CHW Date Index, to determine whether the row's date
larger than date i
                BDdata(i, 7) = 0 'HW Date Index
                BDdata(i, 8) = Ddata(i, 4) 'CHW residual
                BDdata(i, 9) = Ddata(i, 5) 'HW residual
            Next i
            OldNumberofdays = RunTime

            'Sort BDdata by Toa ascending order
            For i = 0 To RunTime - 2
                For j = i + 1 To RunTime - 1
                    If BDdata(i, 1) > BDdata(j, 1) Then
                        Tdata(0, 0) = BDdata(i, 0)
                        Tdata(0, 1) = BDdata(i, 1)
                        Tdata(0, 2) = BDdata(i, 2)
                        Tdata(0, 3) = BDdata(i, 3)
                        Tdata(0, 8) = BDdata(i, 8)
                        Tdata(0, 9) = BDdata(i, 9)

                        BDdata(i, 0) = BDdata(j, 0)
                        BDdata(i, 1) = BDdata(j, 1)
                        BDdata(i, 2) = BDdata(j, 2)
                        BDdata(i, 3) = BDdata(j, 3)
                        BDdata(i, 8) = BDdata(j, 8)
                        BDdata(i, 9) = BDdata(j, 9)

                        BDdata(j, 0) = Tdata(0, 0)
                        BDdata(j, 1) = Tdata(0, 1)
                        BDdata(j, 2) = Tdata(0, 2)
                        BDdata(j, 3) = Tdata(0, 3)
                        BDdata(j, 8) = Tdata(0, 8)
                        BDdata(j, 9) = Tdata(0, 9)
                    End If
                Next j
            Next i
```

```
            Call CHWDetection 'identify CHW fault

          NumberofDays = NumberofDays + RenewTime 'update the database
        Loop

        OldNumberofdays = 0
        NumberofDays = Timelimit
        Erase Ddata
        Erase BDdata
        Do While OldNumberofdays < TotalNumberofDays
          'Determine the period length running "HWDetection"(if transition matrix=fault
detection matrix)
          If NumberofDays < TotalNumberofDays Then
            RunTime = NumberofDays
          Else
            RunTime = TotalNumberofDays
          End If
            For i = OldNumberofdays To RunTime - 1
              Ddata(i, 0) = TDdata(i, 0)
              Ddata(i, 1) = TDdata(i, 1)
              Ddata(i, 2) = TDdata(i, 2)
              Ddata(i, 3) = TDdata(i, 3)
              Ddata(i, 4) = TDdata(i, 4)
              Ddata(i, 5) = TDdata(i, 5)
              BDdata(i, 0) = Ddata(i, 0)
              BDdata(i, 1) = Ddata(i, 1)
              BDdata(i, 2) = Ddata(i, 2)
              BDdata(i, 3) = Ddata(i, 3)
              BDdata(i, 4) = 0 'sum of CHW fault index
              BDdata(i, 5) = 0 'sum of HW fault index
              BDdata(i, 6) = 0 'CHW Date Index, to determine whether the row's date
larger than date i
              BDdata(i, 7) = 0 'HW Date Index
              BDdata(i, 8) = Ddata(i, 4) 'CHW residual
              BDdata(i, 9) = Ddata(i, 5) 'HW residual
            Next i
            OldNumberofdays = RunTime

            'Sort BDdata by Toa ascending order
            For i = 0 To RunTime - 2
              For j = i + 1 To RunTime - 1
                If BDdata(i, 1) > BDdata(j, 1) Then
                  Tdata(0, 0) = BDdata(i, 0)
                  Tdata(0, 1) = BDdata(i, 1)
```

```
            Tdata(0, 2) = BDdata(i, 2)
            Tdata(0, 3) = BDdata(i, 3)
            Tdata(0, 8) = BDdata(i, 8)
            Tdata(0, 9) = BDdata(i, 9)

            BDdata(i, 0) = BDdata(j, 0)
            BDdata(i, 1) = BDdata(j, 1)
            BDdata(i, 2) = BDdata(j, 2)
            BDdata(i, 3) = BDdata(j, 3)
            BDdata(i, 8) = BDdata(j, 8)
            BDdata(i, 9) = BDdata(j, 9)

            BDdata(j, 0) = Tdata(0, 0)
            BDdata(j, 1) = Tdata(0, 1)
            BDdata(j, 2) = Tdata(0, 2)
            BDdata(j, 3) = Tdata(0, 3)
            BDdata(j, 8) = Tdata(0, 8)
            BDdata(j, 9) = Tdata(0, 9)
          End If
        Next j
      Next i

      Call HWDetection 'identify HW fault

    NumberofDays = NumberofDays + RenewTime 'update the database
  Loop
  End Sub
```

Module mCHWDetection

```
  Option Explicit
  Sub CHWDetection()
  Dim i As Integer
  Dim j As Integer
  Dim k As Integer
  Dim x As Integer
  Dim y As Integer
  Dim w As Integer
  Dim z As Integer
  Dim FSDay As Variant ' Fault staring day
  Dim FIDay As Variant ' Fault Identified day
  Dim SumCFlag As Variant ' Sum of CHW residuals during 20 fault-flag days
  Dim CZm As Variant 'Modified z-score
```

```
Dim CBackupBD(1000, 10) As Variant 'Output of CHW results


Erase CBackupBD

i = 0
j = 0
w = 0
z = 0
'clear CHW Date Index
For j = 0 To RunTime - 1
   BDdata(j, 6) = 0
Next j

'Determine whether there is a CHW fault by calculating the sum of consecutive fault indexes(without the -99 rows)
For i = 0 To RunTime - 1 ' i stands for the date in Ddata
   If WorksheetFunction.Or(Ddata(i, 2) = 1, Ddata(i, 2) = -1) Then
      'Cumulative the concecutive "timelimit" CHW fault indexes in BDdata during the period when the date is equal or later than the i date
      For j = 0 To RunTime - (Timelimit - 1)
         w = 0
         z = 0
         BDdata(j, 4) = 0
         BDdata(j, 5) = 0
         Do While z < Timelimit
            If BDdata(j + w + z, 6) = -99 Then ' z is the number of days whose residual is larger than a sigma
               w = w + 1                    ' w is the number of rows containing -99; skip the rows whose date is earlier than i date
            Else
            BDdata(j, 4) = BDdata(j, 4) + BDdata(j + w + z, 2) 'Cumulative the concecutive "timelimit" CHW fault indexes in BDdata during the period when the date is equal or later than the i date
            BDdata(j, 5) = BDdata(j, 5) + BDdata(j + w + z, 3)
            z = z + 1
           End If
         Loop
      Next j

      For w = 0 To RunTime – 1
      'Determin if there is the cumulative sum of fault indexes meet +/-"timelimit"
         If WorksheetFunction.Or(BDdata(w, 4) = Timelimit, BDdata(w, 4) = -Timelimit) Then
```

```
                    FIDay = Ddata(RunTime - 1, 0)
                    j = 0
                    y = 0
                    Do While j < RunTime 'Delete rows from BDdata whose date is early
than i date when +/-"timelimit" is meet
                    If BDdata(j, 0) < Ddata(i, 0) Then
                       j = j + 1
                    Else
                       CBackupBD(y, 0) = BDdata(j, 0)
                       CBackupBD(y, 1) = BDdata(j, 1)
                       CBackupBD(y, 2) = BDdata(j, 2)
                       CBackupBD(y, 3) = BDdata(j, 3)
                       CBackupBD(y, 4) = BDdata(j, 4)
                       CBackupBD(y, 5) = BDdata(j, 5)
                       CBackupBD(y, 6) = BDdata(j, 6) 'CHW time index =-99(date <i
date), otherwise =0 used only when determine CHW fault
                       CBackupBD(y, 7) = BDdata(j, 7) 'HW time index =-99(date <i date),
otherwise =0 used only when determine HW fault
                       CBackupBD(y, 8) = BDdata(j, 8)
                       CBackupBD(y, 9) = BDdata(j, 9)
                       y = y + 1
                       j = j + 1
                    End If
                    Loop

                    'Caculate modified z-score and determine fault starting day
                    FSDay = Ddata(i, 0) 'Fault start day is the earlist day in the sub
transition matrix
                    For k = 0 To y
                        If WorksheetFunction.Or(CBackupBD(k, 4) = Timelimit,
CBackupBD(k, 4) = -Timelimit) Then
                           SumCFlag = 0
                           x = 0
                           For x = 0 To Timelimit - 1
                           SumCFlag = SumCFlag + CBackupBD(k + x, 8) 'Sum of CHW
residuals in the 20 fault-flag days
                           Next x
                       End If
                    Next k

                    CZm = (SumCFlag - Timelimit * CHWMean) / Timelimit /
CHWSigma 'modified z-score

                    'erase the content of sheet2
```

```
                    Application.Calculation = xlCalculationManual
                    Sheets("CHW").Activate
                    Range("A6:H400").Select
                    Selection.ClearContents

                    'Output the data soring by Toa ascending order of the period later than i
date
                    Range("a6").Select
                    For z = 0 To RunTime - 1
                       ActiveCell.Offset(z, 0) = CBackupBD(z, 0) 'date
                       ActiveCell.Offset(z, 1) = CBackupBD(z, 1) 'Toa
                       ActiveCell.Offset(z, 2) = CBackupBD(z, 2) 'CHW fault index
                       ActiveCell.Offset(z, 3) = CBackupBD(z, 3) 'HW fault index
                       ActiveCell.Offset(z, 4) = CBackupBD(z, 4) 'cumulative "time limit"
CHW fault indexes
                       ActiveCell.Offset(z, 5) = CBackupBD(z, 5) 'cumulative "time limit"
HW fault indexes
                       ActiveCell.Offset(z, 6) = CBackupBD(z, 8) 'CHW Residuals
                       ActiveCell.Offset(z, 7) = CBackupBD(z, 9) 'HW Residuals
                    Next z

                    Sheets("CHW").Activate
                    Range("A6:H400").Select
                    Selection.Sort Key1:=Range("B6"), Order1:=xlAscending,
Header:=xlGuess, _
                    OrderCustom:=1, MatchCase:=False, Orientation:=xlTopToBottom, _
                    DataOption1:=xlSortNormal

                    Sheets("Main").Activate
                    Range("G3").Select
                    ActiveCell.Offset(0, 0) = "Y"
                    ActiveCell.Offset(0, 1) = FIDay
                    ActiveCell.Offset(0, 2) = FSDay
                    ActiveCell.Offset(0, 3) = CZm
                    Application.Calculation = xlCalculationAutomatic
                    OldNumberofdays = TotalNumberofDays
                    Exit Sub
                 End If
              Next w

           Else
              For j = 0 To RunTime - 1 'Label -99 to the rows of BDdata whose date is
early than the i date
```

```
                If WorksheetFunction.Or(BDdata(j, 0) < Ddata(i, 0), BDdata(j, 0) =
Ddata(i, 0)) Then
                BDdata(j, 6) = -99
                End If
            Next j
        End If
    Next i

    End Sub
```

Module mHWDetection

```
    Option Explicit
    Sub HWDetection()
    Dim i As Integer
    Dim j As Integer
    Dim k As Integer
    Dim x As Integer
    Dim y As Integer
    Dim w As Integer
    Dim z As Integer
    Dim FSDay As Variant ' Fault staring day
    Dim FIDay As Variant ' Fault Identified day
    Dim SumHFlag As Variant ' Sum of HW residuals during 20 fault-flag days
    Dim HZm As Variant 'Modified z-score

    Dim HBackupBD(1000, 10) As Variant 'Output of HW results

    '-------------------------------------------------------------------------------
    'Indicate calculation begining and end rows
    '-------------------------------------------------------------------------------
    Erase HBackupBD

    i = 0
    j = 0
    w = 0
    z = 0
    'clear HW Date Index
    For j = 0 To RunTime - 1
        BDdata(j, 7) = 0
    Next j

    'Determine wether there is a HW fault by calculating the sum of consecutive Res
Labels(without the -99 rows)
```

```
        For i = 0 To RunTime - 1
           If WorksheetFunction.Or(Ddata(i, 3) = 1, Ddata(i, 3) = -1) Then
              For j = 0 To RunTime - (Timelimit - 1)
                 w = 0
                 z = 0
                 BDdata(j, 4) = 0
                 BDdata(j, 5) = 0
                 Do While z < Timelimit
                    If BDdata(j + w + z, 7) = -99 Then ' z is the number of days whose
residual is larger than a sigma
                       w = w + 1                    ' w is the number of rows containing -99
                    Else
                    BDdata(j, 4) = BDdata(j, 4) + BDdata(j + w + z, 2)
                    BDdata(j, 5) = BDdata(j, 5) + BDdata(j + w + z, 3)
                    z = z + 1
                   End If
                 Loop
              Next j

              For w = 0 To RunTime - 1
                 If WorksheetFunction.Or(BDdata(w, 5) = Timelimit, BDdata(w, 5) = -
Timelimit) Then

                  ' Sheets("TEST").Activate
                  'Range("A6:H400").Select
                  'Selection.ClearContents
                  'Application.Calculation = xlCalculationManual
                  ' Output the data soring by Toa ascending order of the period later than i
date
                   'Range("a6").Select
                   'For z = 0 To RunTime - 1
                    '  ActiveCell.Offset(z, 0) = BDdata(z, 0) 'date
                    '   ActiveCell.Offset(z, 1) = BDdata(z, 1) ' Toa
                    '   ActiveCell.Offset(z, 2) = BDdata(z, 2) 'CHW fault index
                    '   ActiveCell.Offset(z, 3) = BDdata(z, 3) 'HW fault index
                    '   ActiveCell.Offset(z, 4) = BDdata(z, 4) 'date
                    '   ActiveCell.Offset(z, 5) = BDdata(z, 5) ' Toa
                    '   ActiveCell.Offset(z, 6) = BDdata(z, 6) 'CHW fault index
                    '   ActiveCell.Offset(z, 7) = BDdata(z, 7) 'HW fault index
                  ' Next z
                  ' Application.Calculation = xlCalculationAutomatic

                   FIDay = Ddata(RunTime - 1, 0)
                   j = 0
```

```
                y = 0
                Do While j < RunTime 'Delete rows from BDdata whose date is
smaller than i date
                If BDdata(j, 0) < Ddata(i, 0) Then
                   j = j + 1
                Else
                   HBackupBD(y, 0) = BDdata(j, 0)
                   HBackupBD(y, 1) = BDdata(j, 1)
                   HBackupBD(y, 2) = BDdata(j, 2)
                   HBackupBD(y, 3) = BDdata(j, 3)
                   HBackupBD(y, 4) = BDdata(j, 4)
                   HBackupBD(y, 5) = BDdata(j, 5)
                   HBackupBD(y, 6) = BDdata(j, 6)
                   HBackupBD(y, 7) = BDdata(j, 7)
                   HBackupBD(y, 8) = BDdata(j, 8)
                   HBackupBD(y, 9) = BDdata(j, 9)
                   y = y + 1
                   j = j + 1
                End If
                Loop

                'Caculate modified z-score and determine fault starting day
                FSDay = Ddata(i, 0) 'Fault start day is the earlist day in the sub
transition matrix
                For k = 0 To y
                   If WorksheetFunction.Or(HBackupBD(k, 5) = Timelimit,
HBackupBD(k, 5) = -Timelimit) Then
                      SumHFlag = 0
                      x = 0
                      For x = 0 To Timelimit - 1
                      SumHFlag = SumHFlag + HBackupBD(k + x, 9) 'Sum of HW
residuals in the 20 fault-flag days

                      Next x
                   End If
                Next k

                HZm = (SumHFlag - Timelimit * HWMean) / Timelimit / HWSigma
'modified z-score

                'erase the content of sheet2
                Application.Calculation = xlCalculationManual
                Sheets("HW").Activate
                Range("A6:H400").Select
```

```
                    Selection.ClearContents
                    'Output the result
                    Range("a6").Select
                    For z = 0 To RunTime - 1
                       ActiveCell.Offset(z, 0) = HBackupBD(z, 0)
                       ActiveCell.Offset(z, 1) = HBackupBD(z, 1)
                       ActiveCell.Offset(z, 2) = HBackupBD(z, 2)
                       ActiveCell.Offset(z, 3) = HBackupBD(z, 3)
                       ActiveCell.Offset(z, 4) = HBackupBD(z, 4)
                       ActiveCell.Offset(z, 5) = HBackupBD(z, 5)
                       ActiveCell.Offset(z, 6) = HBackupBD(z, 8)
                       ActiveCell.Offset(z, 7) = HBackupBD(z, 9)
                    Next z

                    Sheets("HW").Activate
                    Range("A6:H400").Select
                    Selection.Sort Key1:=Range("B6"), Order1:=xlAscending,
Header:=xlGuess, _
                    OrderCustom:=1, MatchCase:=False, Orientation:=xlTopToBottom, _
                    DataOption1:=xlSortNormal

                    Sheets("Main").Activate
                    Range("G4").Select
                    ActiveCell.Offset(0, 0) = "Y"
                    ActiveCell.Offset(0, 1) = FIDay
                    ActiveCell.Offset(0, 2) = FSDay
                    ActiveCell.Offset(0, 3) = HZm

                    Application.Calculation = xlCalculationAutomatic
                    OldNumberofdays = TotalNumberofDays
                    Exit Sub
                 End If
              Next w

           Else
              For j = 0 To RunTime - 1 'Label -99 to the rows of BDdata whose date is
smaller than the investigating date
                 If WorksheetFunction.Or(BDdata(j, 0) < Ddata(i, 0), BDdata(j, 0) =
Ddata(i, 0)) Then
                 BDdata(j, 7) = -99
                 End If
              Next j
           End If
        Next i
```

```
End Sub
```

APPENDIX B

RESULTS FIGURES OF SIMULATION TEST FOR WHOLE BUILDING FAULT DETECTION APPROACHES

Figure B.1 shows CHW and HW energy consumption changes during the period of October, 1997 to September, 1998 for the 10 simulation test cases in the EOM Building.

Figure B.2 shows CHW and HW energy consumption changes in 2000 for the 10 simulation test cases in the Veterinary Research Building.

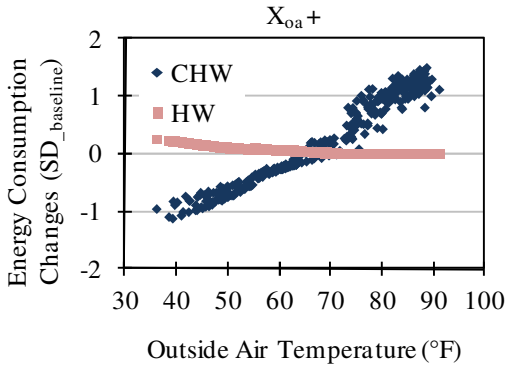

Figure B.1.1

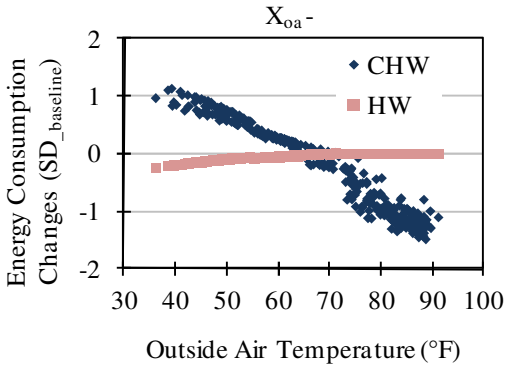

Figure B.1.2

Figure B.1 Cooling and heating energy changes plotted as functions of outside air temperature for the period from 10/1997 to 09/1998 for the EOM Building. (B.1.1) Outside airflow ratio increase of 3.1%, (B.1.2) Outside airflow ratio decrease of 3.1%, (B.1.3) Cold deck leaving air temperature increase of 4°F, (B.1.4) Cold deck leaving air temperature decrease of 4.5°F, (B.1.5) Hot deck leaving air temperature increase of 10°F, (B.1.6) Hot deck leaving air temperature decrease of 21°F, (B.1.7) Minimum airflow ratio increase of 17%, (B.1.8) Minimum airflow ratio decrease of 20%, (B.1.9) Room cooling set-point temperature increase of 1.8°F, (B.1.10) Room cooling set-point temperature decrease of 1.7°F

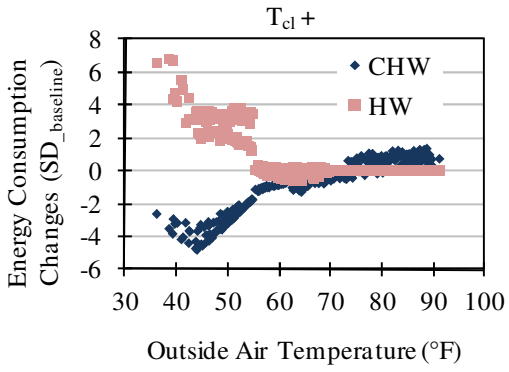

Figure B.1.3

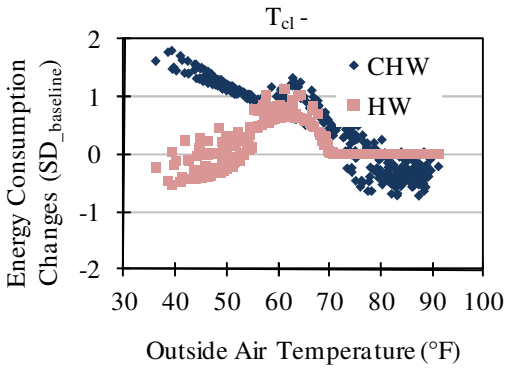

Figure B.1.4

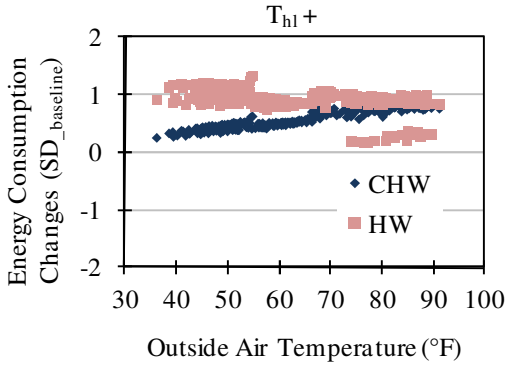

Figure B.1.5

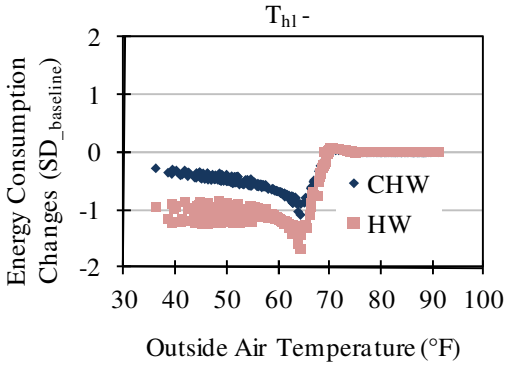

Figure B.1.6

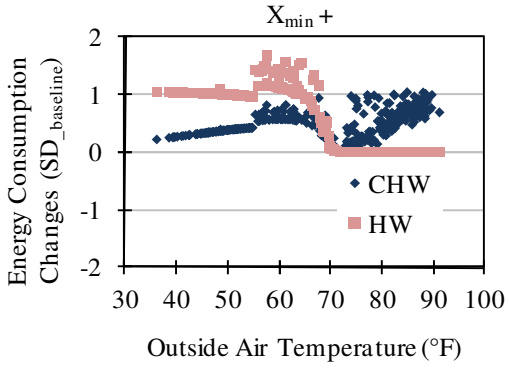

Figure B.1.7

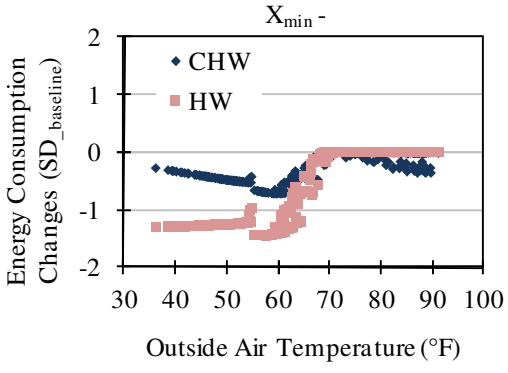

Figure B.1.8

Figure B.1 Continued

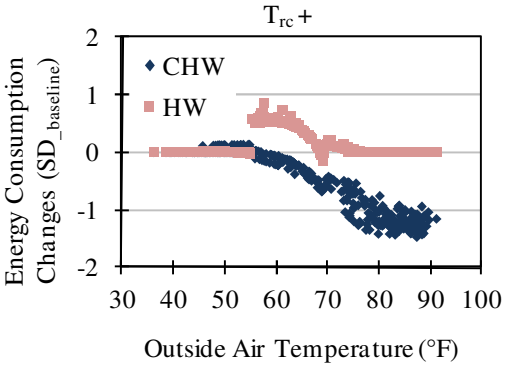

Figure B.1.9

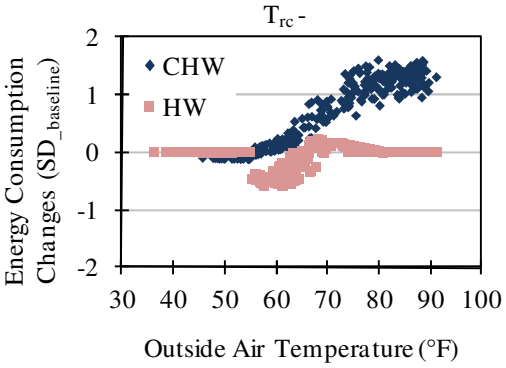

Figure B.1.10

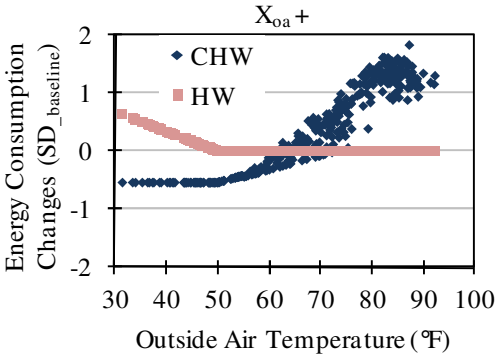

Figure B.2.1

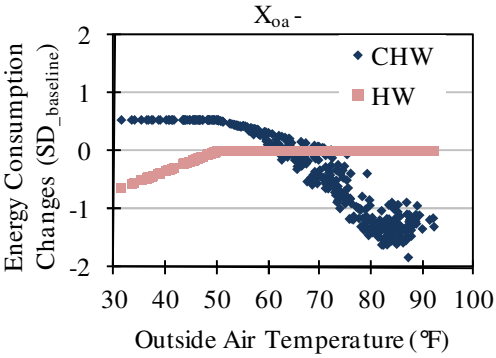

Figure B.2.2

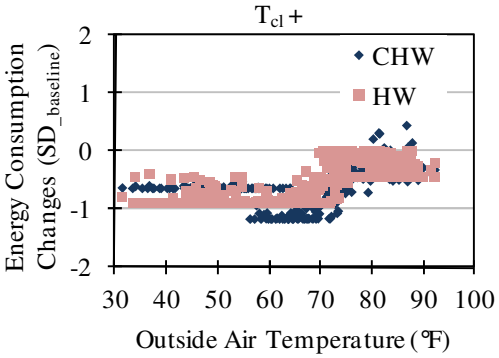

Figure B.2.3

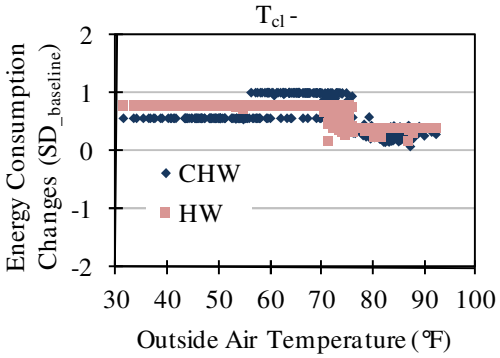

Figure B.2.4

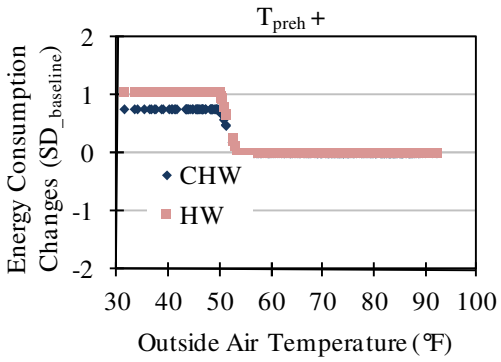

Figure B.2.5

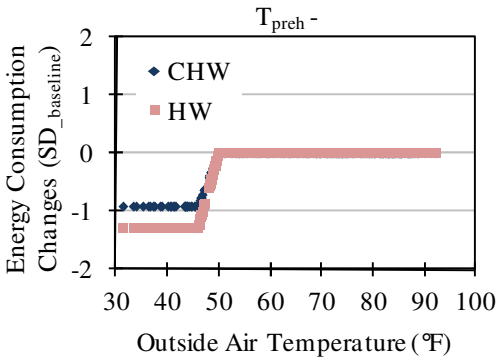

Figure B.2.6

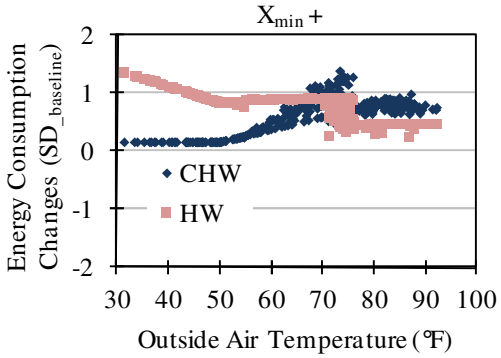

Figure B.2.7

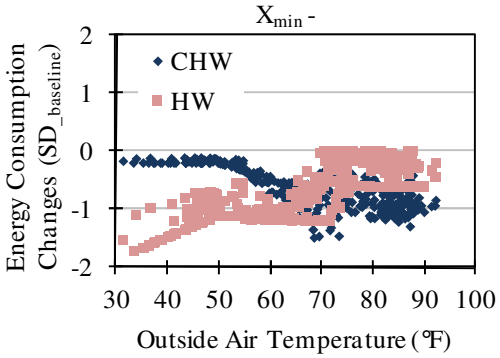

Figure B.2.8

Figure B.2 Cooling and heating energy changes plotted as functions of outside air temperature for the period from 1/1/2000 to12/31/2000 for the Veterinary Research Building. (B.2.1) Outside airflow ratio increase of 6.5%, (B.2.2) Outside airflow ratio decrease of 6.5%, (B.2.3) Cooling coil leaving air temperature increase of 1.7°F, (B.2.4) Cooling coil leaving air temperature decrease of 1.5°F, (B.2.5) Preheat temperature increase of 3.3°F, (B.2.6) Preheat leaving temperature decrease of 4°F, (B.2.7) Minimum airflow ratio increase of 4.9%, (B.2.8) Minimum airflow ratio decrease of 6.5%, (B.2.9) Room cooling set-point temperature increase of 2.8°F, (B.2.10) Room cooling set-point temperature decrease of 1.7°F

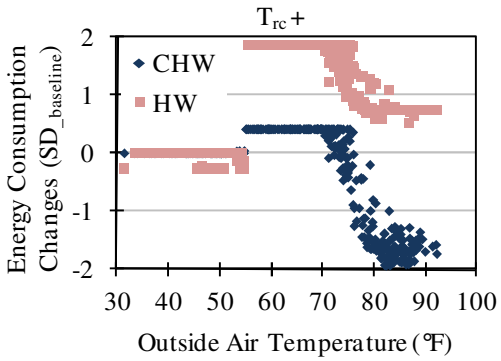

Figure B.2.9

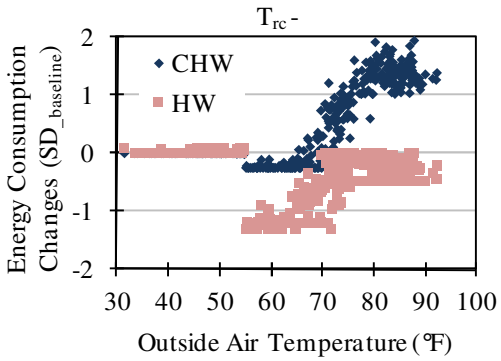

Figure B.2.10

Figure B.2 Continued

APPENDIX C

RESULTS FIGURES AND TABLES OF SIMULATION TEST FOR WHOLE BUILDING FAULT DIAGNOSIS APPROACHES

Tables C.1 to C.18 present the values of cosine similarity and Euclidean distance similarity in nine simulation test cases for the Bush Academic Building. Figures C.1 to C.18 show the reference control changes ordered in descending order of their representative similarity in nine simulation test cases for the Bush Academic Building.

Tables C.19 to C.28 present the values of cosine similarity and Euclidean distance similarity in five simulation test cases for the Veterinary Research Building. Figures C.19 to C.28 show the reference control changes ordered in descending order of their representative similarity in five simulation test cases for the Veterinary Research Building.

Table C.1 Cosine similarity results in Case 2 “Outside air preheat temperature increase of 14°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.12	0.13	0.14	0.14	0.15	0.15
2	X_{oa} increase	-0.06	-0.05	-0.05	-0.04	-0.04	-0.04
3	T_{preh} decrease	-0.97	-0.92	-0.89	-0.86	-0.84	-0.84
4	T_{preh} increase	1.00	1.00	1.00	1.00	1.00	1.00
5	PreHL increase	0.04	0.08	0.10	0.13	0.16	0.16
6	T_{cl} decrease	0.56	0.56	0.55	0.55	0.54	0.56
7	T_{cl} increase	-0.46	-0.53	-0.57	-0.55	-0.49	-0.46
8	T_{hl} decrease	-0.62	-0.61	-0.60	-0.61	-0.62	-0.60
9	T_{hl} increase	0.64	0.66	0.66	0.67	0.68	0.68
10	HL increase	0.00	0.00	0.00	0.00	0.00	0.00
11	X_{min} decrease	-0.56	-0.58	-0.62	-0.67	-0.70	-0.56
12	X_{min} increase	0.59	0.59	0.58	0.56	0.54	0.59
13	T_{rc} decrease	0.00	0.00	0.00	0.00	0.00	0.00
14	T_{rc} increase	0.00	0.00	0.00	0.00	0.00	0.00
15	T_{rh} decrease	-0.77	-0.78	-0.79	-0.80	-0.81	-0.77
16	T_{rh} increase	0.77	0.79	0.79	0.78	0.77	0.79
17	TDL increase	0.36	0.38	0.38	0.37	0.36	0.38

Table C.2 Euclidean distance similarity results in Case 2 “Outside air preheat temperature increase of 14°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.35	0.22	0.13	0.08	0.04	0.35
2	X_{oa} increase	0.34	0.24	0.17	0.11	0.07	0.34
3	T_{preh} decrease	0.35	0.31	0.28	0.27	0.26	0.35
4	T_{preh} increase	0.50	0.61	0.73	0.88	0.94	0.94
5	PreHL increase	0.39	0.34	0.27	0.20	0.15	0.39
6	T_{cl} decrease	0.39	0.15	0.05	0.02	0.01	0.39
7	T_{cl} increase	0.21	0.11	0.05	0.02	0.00	0.21
8	T_{hl} decrease	0.39	0.35	0.32	0.28	0.25	0.39
9	T_{hl} increase	0.44	0.47	0.49	0.50	0.51	0.51
10	HL increase	0.42	0.37	0.23	0.13	0.06	0.42
11	X_{min} decrease	0.25	0.15	0.10	0.06	0.04	0.25
12	X_{min} increase	0.47	0.24	0.09	0.03	0.01	0.47
13	T_{rc} decrease	0.31	0.17	0.09	0.04	0.02	0.31
14	T_{rc} increase	0.32	0.20	0.12	0.07	0.05	0.32
15	T_{rh} decrease	0.33	0.26	0.20	0.15	0.11	0.33
16	T_{rh} increase	0.53	0.62	0.54	0.45	0.37	0.62
17	TDL increase	0.41	0.27	0.14	0.06	0.03	0.41

Table C.3 Cosine similarity results in Case 3 “Heat leakage from preheat coil is 40kBtu/hr” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	-0.29	-0.28	-0.28	-0.28	-0.28	-0.28
2	X_{oa} increase	0.31	0.32	0.33	0.34	0.36	0.36
3	T_{preh} decrease	-0.11	-0.08	-0.06	-0.05	-0.05	-0.05
4	T_{preh} increase	0.13	0.13	0.13	0.13	0.13	0.13
5	PreHL increase	0.99	1.00	1.00	1.00	1.00	1.00
6	T_{cl} decrease	0.74	0.75	0.77	0.79	0.81	0.81
7	T_{cl} increase	-0.74	-0.66	-0.43	-0.20	-0.02	-0.02
8	T_{hl} decrease	-0.68	-0.69	-0.69	-0.67	-0.63	-0.63
9	T_{hl} increase	0.66	0.66	0.65	0.64	0.64	0.66
10	HL increase	0.20	0.71	0.83	0.86	0.89	0.89
11	X_{min} decrease	-0.71	-0.68	-0.64	-0.59	-0.54	-0.54
12	X_{min} increase	0.74	0.77	0.79	0.82	0.84	0.84
13	T_{rc} decrease	0.31	0.34	0.38	0.42	0.46	0.46
14	T_{rc} increase	-0.27	-0.23	-0.19	-0.18	-0.18	-0.18
15	T_{rh} decrease	-0.12	-0.12	-0.12	-0.13	-0.13	-0.12
16	T_{rh} increase	0.10	0.10	0.11	0.11	0.10	0.11
17	TDL increase	0.83	0.85	0.86	0.87	0.88	0.88

Table C.4 Euclidean distance similarity results in Case 3 “Heat leakage from preheat coil is 40kBtu/hr” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	X_{oa} decrease	0.17	0.11	0.07	0.04	0.03	0.17
2	X_{oa} increase	0.25	0.20	0.14	0.10	0.06	0.25
3	T_{preh} decrease	0.23	0.22	0.22	0.21	0.21	0.23
4	T_{preh} increase	0.24	0.23	0.23	0.21	0.20	0.24
5	PreHL increase	0.33	0.48	0.69	1.00	0.69	1.00
6	T_{cl} decrease	0.37	0.22	0.09	0.03	0.01	0.37
7	T_{cl} increase	0.10	0.06	0.04	0.02	0.01	0.10
8	T_{hl} decrease	0.22	0.20	0.17	0.16	0.14	0.22
9	T_{hl} increase	0.25	0.26	0.28	0.29	0.29	0.29
10	HL increase	0.23	0.31	0.44	0.38	0.23	0.44
11	X_{min} decrease	0.14	0.08	0.06	0.04	0.03	0.14
12	X_{min} increase	0.36	0.33	0.16	0.06	0.02	0.36
13	T_{rc} decrease	0.25	0.18	0.10	0.05	0.02	0.25
14	T_{rc} increase	0.18	0.14	0.10	0.06	0.04	0.18
15	T_{rh} decrease	0.22	0.19	0.16	0.13	0.11	0.22
16	T_{rh} increase	0.23	0.22	0.19	0.16	0.14	0.23
17	TDL increase	0.38	0.45	0.31	0.16	0.07	0.45

Table C.5 Cosine similarity results in Case 4 “Cold deck leaving air temperature decrease of 2°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.12	0.13	0.13	0.13	0.13	0.13
2	Xoa increase	-0.07	-0.06	-0.05	-0.03	-0.01	-0.01
3	Tpreh decrease	-0.55	-0.53	-0.51	-0.50	-0.49	-0.49
4	Tpreh increase	0.56	0.56	0.56	0.56	0.56	0.56
5	The heat leakage of preheat coil increase	0.69	0.71	0.72	0.74	0.75	0.75
6	Tcl decrease	1.00	0.99	0.98	0.98	0.97	1.00
7	Tcl increase	-0.97	-0.93	-0.83	-0.67	-0.51	-0.51
8	Thl decrease	-0.86	-0.86	-0.86	-0.88	-0.88	-0.86
9	Thl increase	0.86	0.86	0.86	0.86	0.86	0.86
10	The heat leakage of heating coil increase	0.18	0.40	0.44	0.47	0.50	0.50
11	Xmin decrease	-0.81	-0.80	-0.80	-0.79	-0.78	-0.78
12	Xmin increase	0.87	0.89	0.89	0.88	0.87	0.89
13	Trc decrease	0.05	0.09	0.13	0.17	0.21	0.21
14	Trc increase	-0.02	0.00	0.01	0.01	-0.01	0.01
15	Trh decrease	-0.40	-0.41	-0.42	-0.43	-0.44	-0.40
16	Trh increase	0.41	0.41	0.41	0.41	0.40	0.41
17	Terminal box damper leakage increase	0.61	0.64	0.66	0.67	0.68	0.68

Table C.6 Euclidean distance similarity results in Case 4 “Cold deck leaving air temperature decrease of 2°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.30	0.20	0.13	0.08	0.04	0.30
2	Xoa increase	0.29	0.21	0.15	0.10	0.07	0.29
3	Tpreh decrease	0.30	0.28	0.26	0.25	0.24	0.30
4	Tpreh increase	0.37	0.40	0.41	0.40	0.38	0.41
5	The heat leakage of preheat coil increase	0.42	0.46	0.44	0.37	0.29	0.46
6	Tcl decrease	1.00	0.30	0.10	0.03	0.01	1.00
7	Tcl increase	0.13	0.07	0.04	0.02	0.00	0.13
8	Thl decrease	0.31	0.27	0.24	0.21	0.18	0.31
9	Thl increase	0.37	0.40	0.43	0.46	0.48	0.48
10	The heat leakage of heating coil increase	0.34	0.37	0.30	0.19	0.10	0.37
11	Xmin decrease	0.18	0.11	0.07	0.05	0.03	0.18
12	Xmin increase	0.62	0.40	0.16	0.05	0.01	0.62
13	Trc decrease	0.27	0.16	0.08	0.04	0.02	0.27
14	Trc increase	0.31	0.21	0.13	0.08	0.05	0.31
15	Trh decrease	0.28	0.23	0.18	0.14	0.11	0.28
16	Trh increase	0.38	0.37	0.32	0.27	0.22	0.38
17	Terminal box damper leakage increase	0.43	0.34	0.19	0.09	0.04	0.43

Table C.7 Cosine similarity results in Case 5 “Hot deck leaving air temperature increase of 30°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.04	0.05	0.06	0.06	0.06	0.06
2	Xoa increase	0.01	0.03	0.04	0.05	0.06	0.06
3	Tpreh decrease	-0.71	-0.68	-0.66	-0.64	-0.63	-0.63
4	Tpreh increase	0.72	0.72	0.72	0.72	0.72	0.72
5	The heat leakage of preheat coil increase	0.54	0.56	0.58	0.60	0.62	0.62
6	Tcl decrease	0.86	0.85	0.83	0.83	0.83	0.86
7	Tcl increase	-0.84	-0.90	-0.86	-0.72	-0.58	-0.58
8	Thl decrease	-0.99	-0.98	-0.98	-0.98	-0.98	-0.98
9	Thl increase	0.99	0.99	0.99	1.00	1.00	1.00
10	The heat leakage of heating coil increase	0.02	0.12	0.22	0.26	0.29	0.29
11	Xmin decrease	-0.88	-0.90	-0.91	-0.92	-0.92	-0.88
12	Xmin increase	0.88	0.87	0.86	0.84	0.83	0.88
13	Trc decrease	-0.05	-0.04	-0.02	0.01	0.04	0.04
14	Trc increase	0.04	0.04	0.04	0.04	0.04	0.04
15	Trh decrease	-0.56	-0.57	-0.58	-0.59	-0.60	-0.56
16	Trh increase	0.56	0.57	0.57	0.57	0.56	0.57
17	Terminal box damper leakage increase	0.58	0.60	0.61	0.62	0.62	0.62

Table C.8 Euclidean distance similarity results in Case 5 “Hot deck leaving air temperature increase of 30°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.32	0.21	0.12	0.07	0.04	0.32
2	Xoa increase	0.33	0.24	0.17	0.11	0.07	0.33
3	Tpreh decrease	0.34	0.31	0.29	0.27	0.26	0.34
4	Tpreh increase	0.44	0.49	0.52	0.52	0.50	0.52
5	The heat leakage of preheat coil increase	0.45	0.45	0.39	0.31	0.24	0.45
6	Tcl decrease	0.58	0.22	0.07	0.02	0.01	0.58
7	Tcl increase	0.16	0.08	0.04	0.02	0.00	0.16
8	Thl decrease	0.35	0.31	0.27	0.23	0.20	0.35
9	Thl increase	0.43	0.48	0.52	0.56	0.60	0.60
10	The heat leakage of heating coil increase	0.39	0.37	0.27	0.16	0.08	0.39
11	Xmin decrease	0.21	0.12	0.07	0.05	0.04	0.21
12	Xmin increase	0.65	0.35	0.13	0.04	0.01	0.65
13	Trc decrease	0.29	0.16	0.08	0.04	0.01	0.29
14	Trc increase	0.33	0.22	0.13	0.08	0.05	0.33
15	Trh decrease	0.32	0.26	0.20	0.15	0.11	0.32
16	Trh increase	0.45	0.46	0.41	0.35	0.29	0.46
17	Terminal box damper leakage increase	0.47	0.34	0.18	0.08	0.04	0.47

Table C.9 Cosine similarity results in Case 7 “Heat leakage from heating coil is 70kBtu/hr” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.49	-0.49	-0.49	-0.48	-0.48	-0.48
2	Xoa increase	0.52	0.52	0.53	0.54	0.56	0.56
3	Tpreh decrease	-0.16	-0.17	-0.17	-0.16	-0.16	-0.16
4	Tpreh increase	0.17	0.17	0.17	0.17	0.17	0.17
5	The heat leakage of preheat coil increase	0.88	0.88	0.88	0.88	0.88	0.88
6	Tcl decrease	0.55	0.58	0.61	0.64	0.67	0.67
7	Tcl increase	-0.49	-0.36	-0.10	0.12	0.28	0.28
8	Thl decrease	-0.44	-0.45	-0.46	-0.42	-0.37	-0.37
9	Thl increase	0.44	0.43	0.43	0.42	0.42	0.44
10	The heat leakage of heating coil increase	0.24	0.87	0.96	0.97	0.96	0.97
11	Xmin decrease	-0.57	-0.53	-0.47	-0.43	-0.40	-0.40
12	Xmin increase	0.64	0.67	0.70	0.73	0.77	0.77
13	Trc decrease	0.58	0.61	0.64	0.66	0.69	0.69
14	Trc increase	-0.52	-0.47	-0.43	-0.40	-0.40	-0.40
15	Trh decrease	-0.15	-0.15	-0.16	-0.16	-0.16	-0.15
16	Trh increase	0.16	0.16	0.16	0.16	0.16	0.16
17	Terminal box damper leakage increase	0.93	0.93	0.94	0.94	0.93	0.94

Table C.10 Euclidean distance similarity results in Case 7 “Heat leakage from heating coil is 70kBtu/hr” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.10	0.07	0.04	0.02	0.01	0.10
2	Xoa increase	0.19	0.19	0.15	0.11	0.08	0.19
3	Tpreh decrease	0.14	0.14	0.13	0.13	0.13	0.14
4	Tpreh increase	0.15	0.15	0.15	0.14	0.14	0.15
5	The heat leakage of preheat coil increase	0.20	0.27	0.34	0.39	0.40	0.40
6	Tcl decrease	0.20	0.14	0.07	0.03	0.01	0.20
7	Tcl increase	0.08	0.05	0.04	0.02	0.01	0.08
8	Thl decrease	0.14	0.13	0.12	0.11	0.10	0.14
9	Thl increase	0.15	0.16	0.16	0.17	0.17	0.17
10	The heat leakage of heating coil increase	0.15	0.22	0.42	0.61	0.40	0.61
11	Xmin decrease	0.09	0.06	0.04	0.03	0.03	0.09
12	Xmin increase	0.21	0.22	0.14	0.06	0.02	0.22
13	Trc decrease	0.20	0.25	0.16	0.08	0.03	0.25
14	Trc increase	0.10	0.07	0.05	0.03	0.02	0.10
15	Trh decrease	0.14	0.12	0.10	0.09	0.07	0.14
16	Trh increase	0.15	0.15	0.13	0.12	0.10	0.15
17	Terminal box damper leakage increase	0.27	0.46	0.47	0.26	0.11	0.47

Table C.11 Cosine similarity results in Case 7 “Minimum airflow ratio increase of 2.5%” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.27	-0.27	-0.26	-0.26	-0.26	-0.26
2	Xoa increase	0.32	0.32	0.33	0.34	0.36	0.36
3	Tpreh decrease	-0.58	-0.58	-0.57	-0.55	-0.55	-0.55
4	Tpreh increase	0.60	0.60	0.60	0.60	0.60	0.60
5	The heat leakage of preheat coil increase	0.71	0.72	0.73	0.75	0.76	0.76
6	Tcl decrease	0.88	0.87	0.86	0.87	0.88	0.88
7	Tcl increase	-0.83	-0.81	-0.65	-0.45	-0.26	-0.26
8	Thl decrease	-0.89	-0.89	-0.89	-0.89	-0.88	-0.88
9	Thl increase	0.89	0.89	0.88	0.88	0.88	0.89
10	The heat leakage of heating coil increase	0.14	0.37	0.48	0.52	0.55	0.55
11	Xmin decrease	-0.97	-0.96	-0.95	-0.93	-0.92	-0.92
12	Xmin increase	1.00	0.99	0.98	0.97	0.96	1.00
13	Trc decrease	0.15	0.18	0.22	0.27	0.30	0.30
14	Trc increase	-0.11	-0.10	-0.10	-0.09	-0.10	-0.09
15	Trh decrease	-0.57	-0.58	-0.58	-0.58	-0.59	-0.57
16	Trh increase	0.59	0.59	0.59	0.58	0.58	0.59
17	Terminal box damper leakage increase	0.77	0.80	0.82	0.82	0.83	0.83

Table C.12 Euclidean distance similarity results in Case 7 “Minimum airflow ratio increase of 2.5%” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.25	0.15	0.09	0.05	0.03	0.25
2	Xoa increase	0.36	0.27	0.19	0.13	0.09	0.36
3	Tpreh decrease	0.32	0.29	0.27	0.26	0.25	0.32
4	Tpreh increase	0.39	0.42	0.44	0.43	0.41	0.44
5	The heat leakage of preheat coil increase	0.45	0.50	0.47	0.38	0.29	0.50
6	Tcl decrease	0.61	0.25	0.08	0.03	0.01	0.61
7	Tcl increase	0.15	0.08	0.04	0.02	0.00	0.15
8	Thl decrease	0.32	0.29	0.25	0.22	0.19	0.32
9	Thl increase	0.39	0.42	0.45	0.48	0.51	0.51
10	The heat leakage of heating coil increase	0.36	0.38	0.33	0.20	0.11	0.38
11	Xmin decrease	0.18	0.10	0.07	0.05	0.03	0.18
12	Xmin increase	0.79	0.48	0.17	0.05	0.01	0.79
13	Trc decrease	0.32	0.18	0.09	0.04	0.02	0.32
14	Trc increase	0.28	0.19	0.12	0.07	0.05	0.28
15	Trh decrease	0.30	0.24	0.19	0.15	0.11	0.30
16	Trh increase	0.41	0.44	0.40	0.34	0.29	0.44
17	Terminal box damper leakage increase	0.52	0.43	0.24	0.11	0.05	0.52

Table C.13 Cosine similarity results in Case 8 “Room heating set-point temperature increase of 3°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.09	-0.09	-0.09	-0.08	-0.08	-0.08
2	Xoa increase	0.11	0.11	0.12	0.12	0.12	0.12
3	Tpreh decrease	-0.74	-0.73	-0.72	-0.70	-0.69	-0.69
4	Tpreh increase	0.79	0.79	0.79	0.79	0.79	0.79
5	The heat leakage of preheat coil increase	0.04	0.07	0.08	0.11	0.13	0.13
6	Tcl decrease	0.41	0.42	0.42	0.41	0.41	0.42
7	Tcl increase	-0.34	-0.39	-0.41	-0.39	-0.33	-0.33
8	Thl decrease	-0.51	-0.50	-0.49	-0.51	-0.51	-0.49
9	Thl increase	0.52	0.53	0.54	0.54	0.54	0.54
10	The heat leakage of heating coil increase	0.00	0.00	0.00	0.00	0.00	0.00
11	Xmin decrease	-0.56	-0.59	-0.64	-0.69	-0.73	-0.56
12	Xmin increase	0.59	0.58	0.56	0.54	0.52	0.59
13	Trc decrease	0.00	0.00	0.00	0.00	0.00	0.00
14	Trc increase	0.00	0.00	0.00	0.00	0.00	0.00
15	Trh decrease	-0.99	-0.99	-0.99	-0.99	-0.99	-0.99
16	Trh increase	0.99	1.00	1.00	1.00	1.00	1.00
17	Terminal box damper leakage increase	0.35	0.37	0.37	0.36	0.35	0.37

Table C.14 Euclidean distance similarity results in Case 8 “Room heating set-point temperature increase of 3°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.29	0.17	0.10	0.06	0.03	0.29
2	Xoa increase	0.32	0.22	0.15	0.10	0.07	0.32
3	Tpreh decrease	0.37	0.32	0.30	0.28	0.27	0.37
4	Tpreh increase	0.52	0.55	0.55	0.54	0.53	0.55
5	The heat leakage of preheat coil increase	0.39	0.32	0.25	0.19	0.14	0.39
6	Tcl decrease	0.32	0.13	0.05	0.02	0.01	0.32
7	Tcl increase	0.20	0.10	0.05	0.02	0.00	0.20
8	Thl decrease	0.41	0.37	0.33	0.29	0.25	0.41
9	Thl increase	0.46	0.48	0.48	0.48	0.48	0.48
10	The heat leakage of heating coil increase	0.44	0.36	0.22	0.12	0.06	0.44
11	Xmin decrease	0.25	0.15	0.09	0.06	0.04	0.25
12	Xmin increase	0.44	0.24	0.09	0.03	0.01	0.44
13	Trc decrease	0.29	0.16	0.08	0.04	0.01	0.29
14	Trc increase	0.30	0.18	0.11	0.07	0.04	0.30
15	Trh decrease	0.34	0.27	0.21	0.15	0.11	0.34
16	Trh increase	0.57	0.76	1.00	0.78	0.62	1.00
17	Terminal box damper leakage increase	0.38	0.25	0.14	0.07	0.03	0.38

Table C.15 Cosine similarity results in Case 9 “Room cooling set-point temperature decrease of 2.2°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.68	-0.68	-0.68	-0.68	-0.68	-0.68
2	Xoa increase	0.70	0.71	0.71	0.72	0.74	0.74
3	Tpreh decrease	0.00	0.00	0.00	0.00	0.00	0.00
4	Tpreh increase	0.00	0.00	0.00	0.00	0.00	0.00
5	The heat leakage of preheat coil increase	0.35	0.35	0.35	0.35	0.34	0.35
6	Tcl decrease	0.09	0.10	0.13	0.16	0.19	0.19
7	Tcl increase	-0.08	0.07	0.35	0.55	0.67	0.67
8	Thl decrease	0.03	0.03	0.03	0.07	0.11	0.11
9	Thl increase	-0.03	-0.03	-0.03	-0.04	-0.04	-0.03
10	The heat leakage of heating coil increase	0.10	0.50	0.58	0.58	0.56	0.58
11	Xmin decrease	-0.16	-0.14	-0.08	-0.03	0.00	0.00
12	Xmin increase	0.19	0.19	0.21	0.24	0.29	0.29
13	Trc decrease	0.99	1.00	1.00	0.99	0.98	1.00
14	Trc increase	-0.97	-0.96	-0.94	-0.93	-0.93	-0.93
15	Trh decrease	0.00	0.00	0.00	0.00	0.00	0.00
16	Trh increase	0.00	0.00	0.00	0.00	0.00	0.00
17	Terminal box damper leakage increase	0.59	0.58	0.58	0.57	0.56	0.59

Table C.16 Euclidean distance similarity results in Case 9 “Room cooling set-point temperature decrease of 2.2°F” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.13	0.08	0.05	0.03	0.02	0.13
2	Xoa increase	0.26	0.29	0.24	0.17	0.11	0.29
3	Tpreh decrease	0.19	0.19	0.18	0.18	0.18	0.19
4	Tpreh increase	0.19	0.19	0.17	0.16	0.15	0.19
5	The heat leakage of preheat coil increase	0.20	0.19	0.18	0.16	0.14	0.20
6	Tcl decrease	0.14	0.06	0.03	0.01	0.00	0.14
7	Tcl increase	0.17	0.13	0.11	0.06	0.02	0.17
8	Thl decrease	0.20	0.20	0.20	0.20	0.21	0.21
9	Thl increase	0.19	0.19	0.18	0.18	0.17	0.19
10	The heat leakage of heating coil increase	0.19	0.22	0.24	0.22	0.13	0.24
11	Xmin decrease	0.18	0.13	0.09	0.07	0.05	0.18
12	Xmin increase	0.17	0.11	0.06	0.02	0.01	0.17
13	Trc decrease	0.40	0.85	0.53	0.23	0.09	0.85
14	Trc increase	0.10	0.05	0.03	0.02	0.01	0.10
15	Trh decrease	0.19	0.17	0.14	0.12	0.09	0.19
16	Trh increase	0.19	0.16	0.14	0.11	0.09	0.19
17	Terminal box damper leakage increase	0.23	0.25	0.19	0.10	0.04	0.25

Table C.17 Cosine similarity results in Case 10 “Terminal box damper leakage increase of 7.5%” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.55	-0.55	-0.54	-0.54	-0.54	-0.54
2	Xoa increase	0.59	0.60	0.61	0.62	0.63	0.63
3	Tpreh decrease	-0.37	-0.37	-0.37	-0.37	-0.37	-0.37
4	Tpreh increase	0.37	0.37	0.37	0.37	0.37	0.37
5	The heat leakage of preheat coil increase	0.83	0.83	0.84	0.84	0.85	0.85
6	Tcl decrease	0.64	0.66	0.68	0.71	0.74	0.74
7	Tcl increase	-0.57	-0.51	-0.26	-0.01	0.17	0.17
8	Thl decrease	-0.62	-0.63	-0.63	-0.59	-0.56	-0.56
9	Thl increase	0.62	0.62	0.62	0.62	0.61	0.62
10	The heat leakage of heating coil increase	0.14	0.68	0.80	0.82	0.83	0.83
11	Xmin decrease	-0.75	-0.74	-0.70	-0.66	-0.64	-0.64
12	Xmin increase	0.79	0.81	0.83	0.85	0.88	0.88
13	Trc decrease	0.56	0.58	0.61	0.63	0.66	0.66
14	Trc increase	-0.52	-0.49	-0.45	-0.43	-0.42	-0.42
15	Trh decrease	-0.35	-0.35	-0.36	-0.36	-0.36	-0.35
16	Trh increase	0.36	0.36	0.37	0.37	0.36	0.37
17	Terminal box damper leakage increase	1.00	1.00	1.00	0.99	0.99	1.00

Table C.18 Euclidean distance similarity results in Case 10 “Terminal box damper leakage increase of 7.5%” for the Bush Academic Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.18	0.11	0.06	0.04	0.02	0.18
2	Xoa increase	0.34	0.31	0.23	0.16	0.11	0.34
3	Tpreh decrease	0.25	0.23	0.22	0.21	0.20	0.25
4	Tpreh increase	0.29	0.30	0.30	0.29	0.28	0.30
5	The heat leakage of preheat coil increase	0.35	0.44	0.48	0.45	0.37	0.48
6	Tcl decrease	0.36	0.19	0.07	0.03	0.01	0.36
7	Tcl increase	0.13	0.07	0.05	0.02	0.01	0.13
8	Thl decrease	0.25	0.23	0.20	0.18	0.17	0.25
9	Thl increase	0.29	0.30	0.32	0.33	0.34	0.34
10	The heat leakage of heating coil increase	0.27	0.35	0.45	0.34	0.19	0.45
11	Xmin decrease	0.15	0.09	0.06	0.04	0.03	0.15
12	Xmin increase	0.43	0.36	0.16	0.06	0.02	0.43
13	Trc decrease	0.35	0.28	0.14	0.07	0.03	0.35
14	Trc increase	0.19	0.12	0.08	0.05	0.03	0.19
15	Trh decrease	0.23	0.19	0.16	0.12	0.10	0.23
16	Trh increase	0.30	0.29	0.26	0.23	0.19	0.30
17	Terminal box damper leakage increase	0.53	0.91	0.43	0.19	0.08	0.91

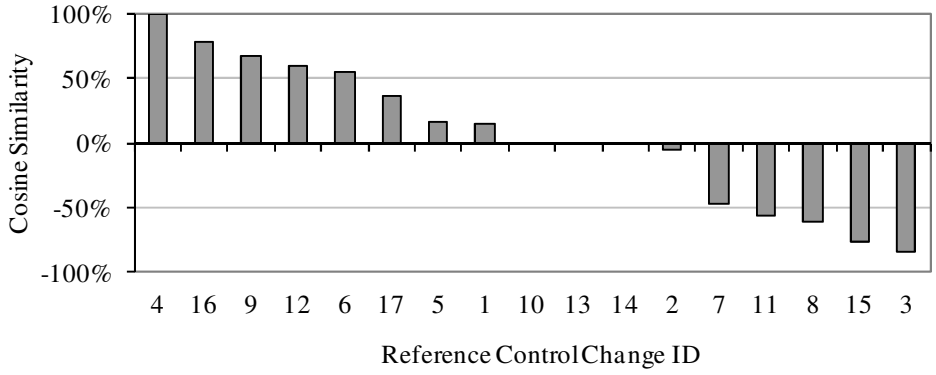

Figure C.1 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 2)

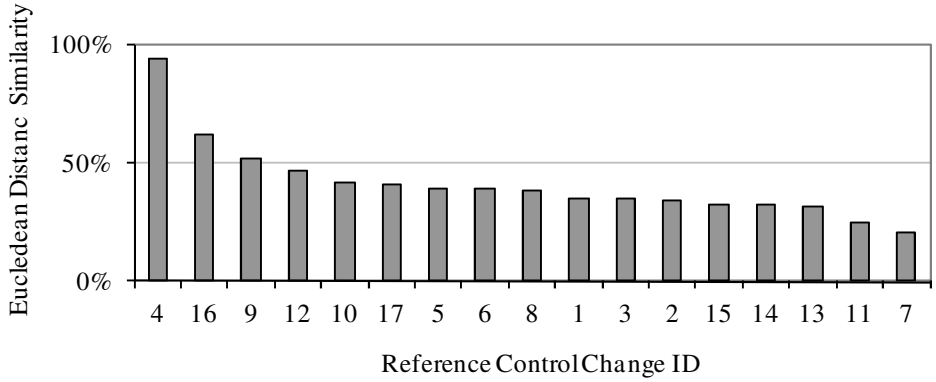

Figure C.2 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 2)

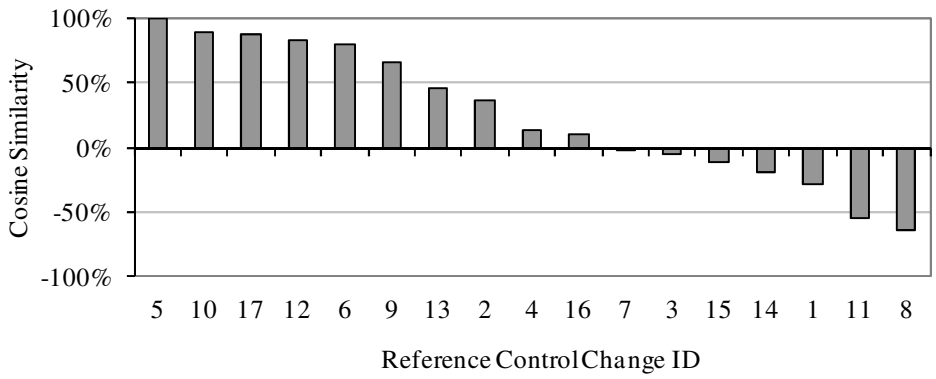

Figure C.3 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 3)

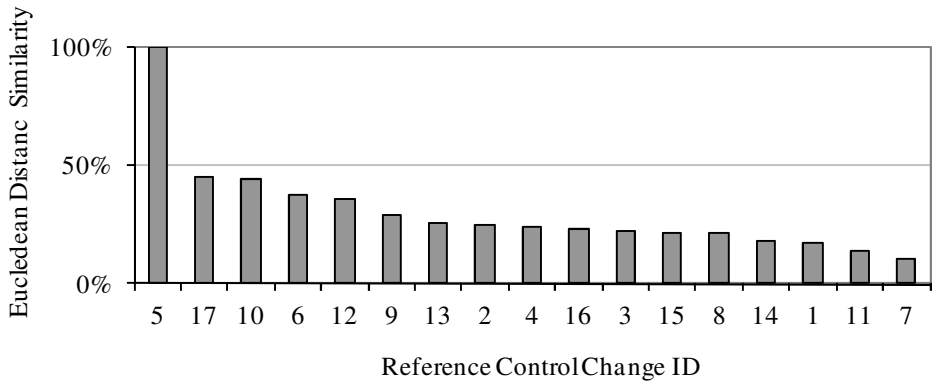

Figure C.4 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 3)

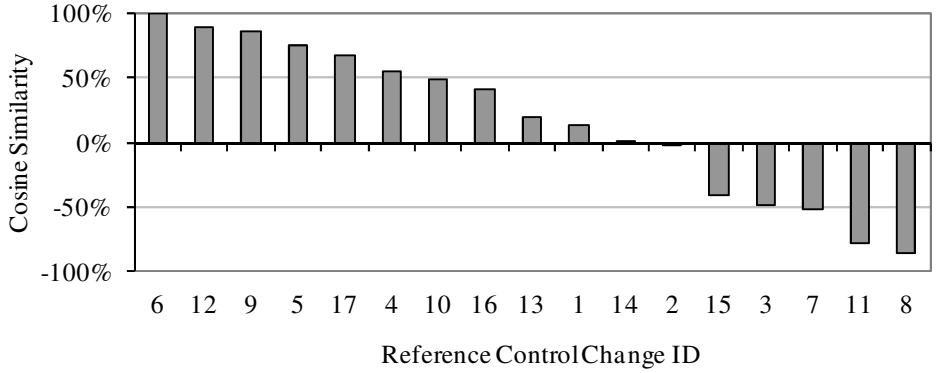

Figure C.5 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 4)

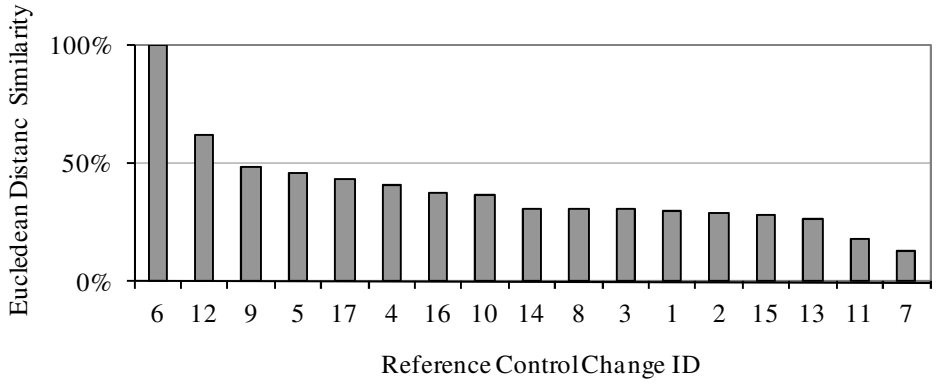

Figure C.6 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 4)

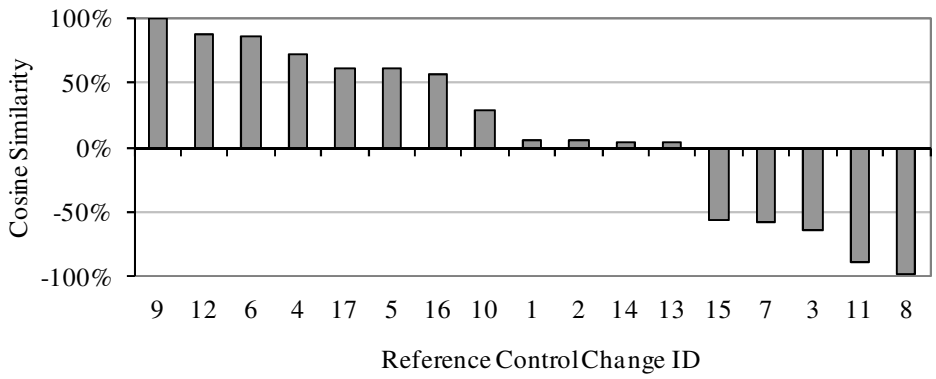

Figure C.7 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 5)

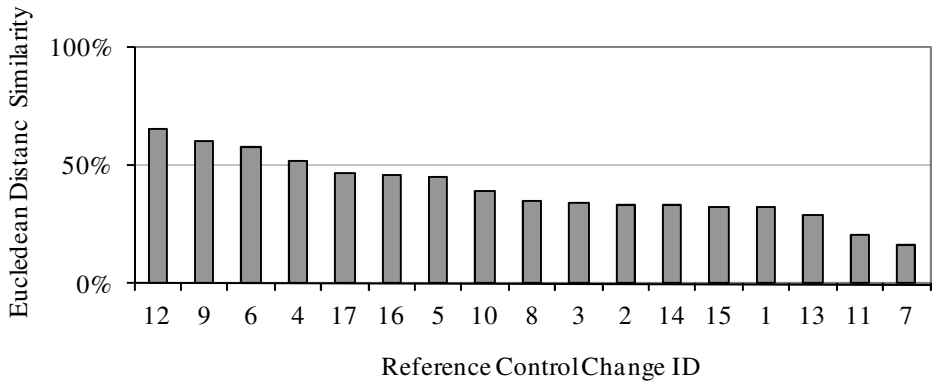

Figure C.8 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 5)

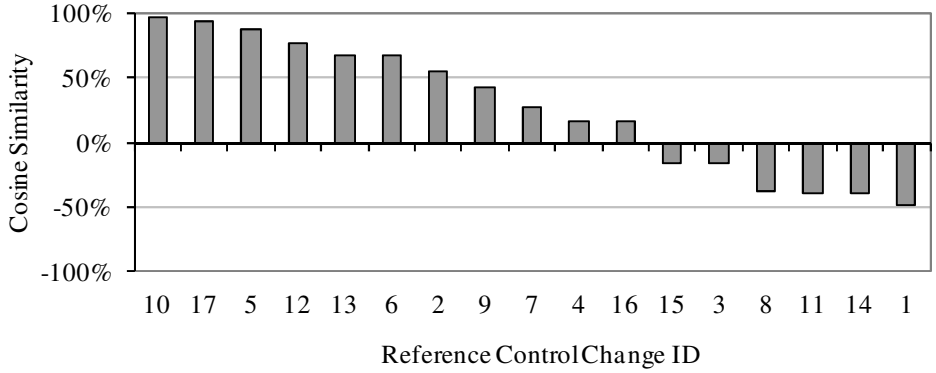

Figure C.9 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 7)

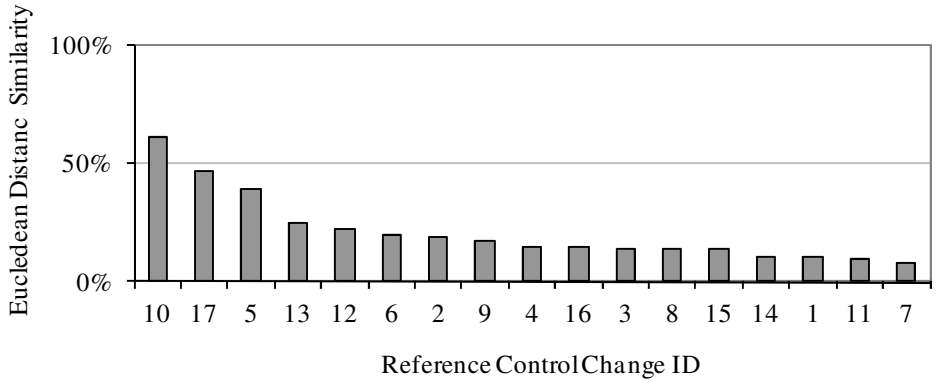

Figure C.10 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 7)

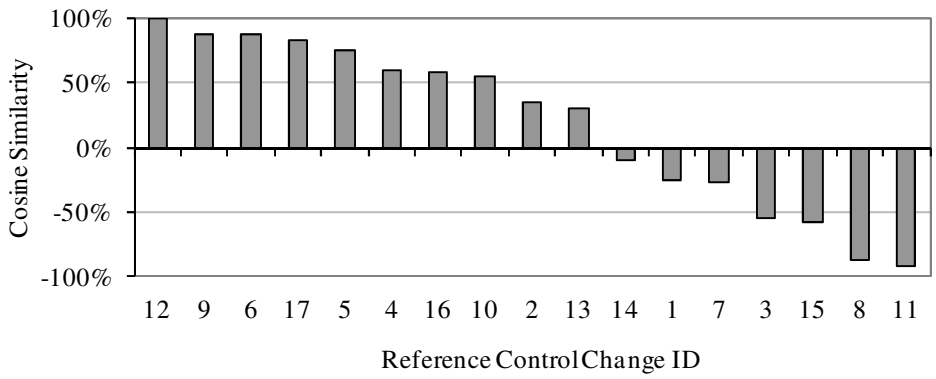

Figure C.11 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 7)

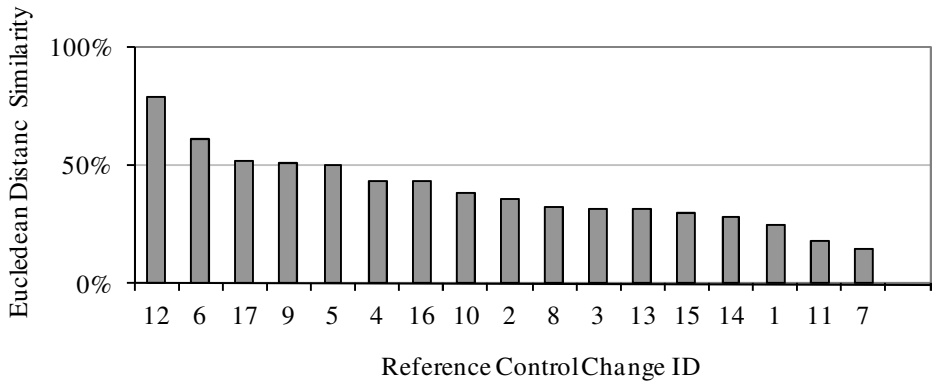

Figure C.12 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 7)

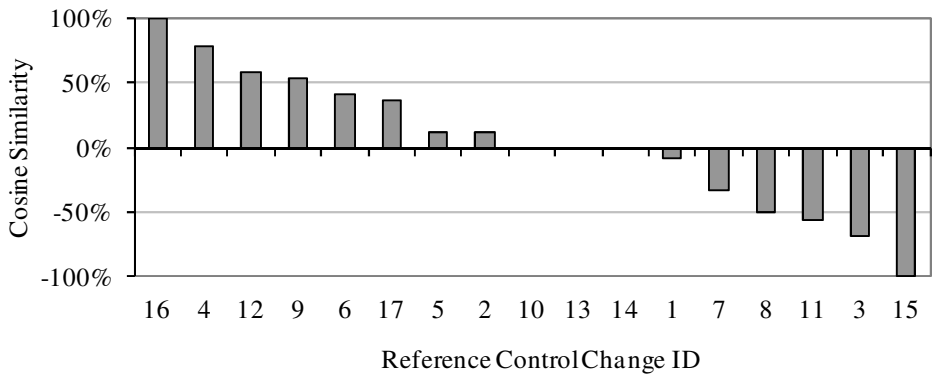

Figure C.13 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 8)

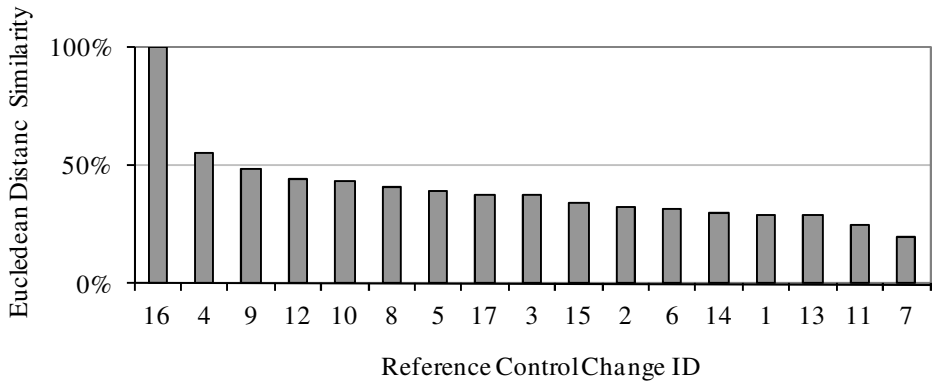

Figure C.14 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 8)

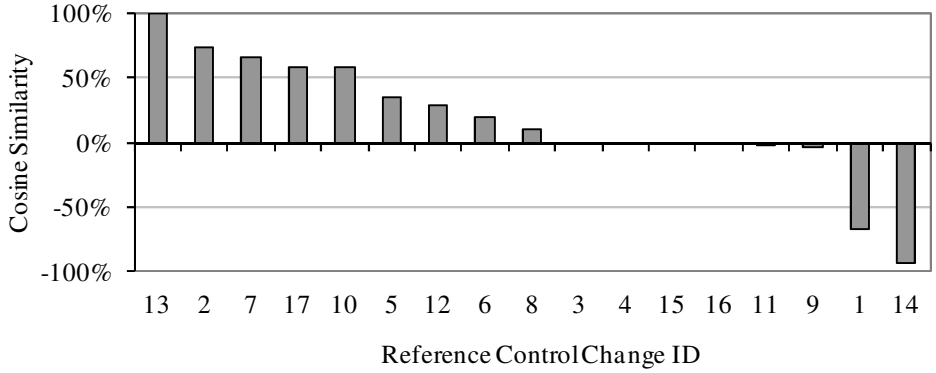

Figure C.15 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 9)

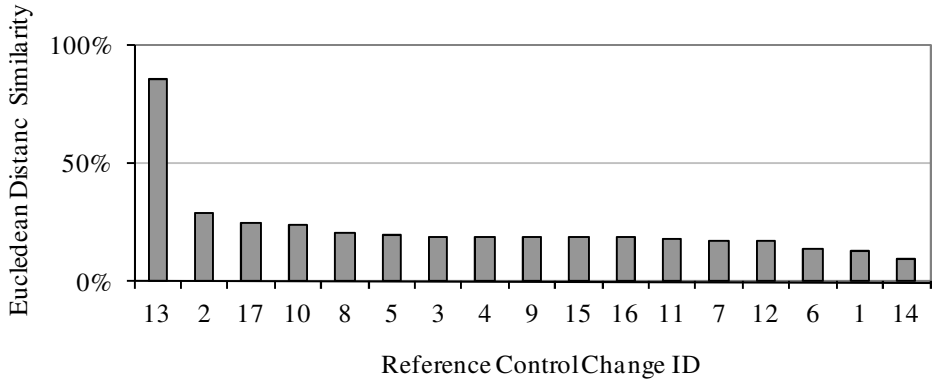

Figure C.16 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 9)

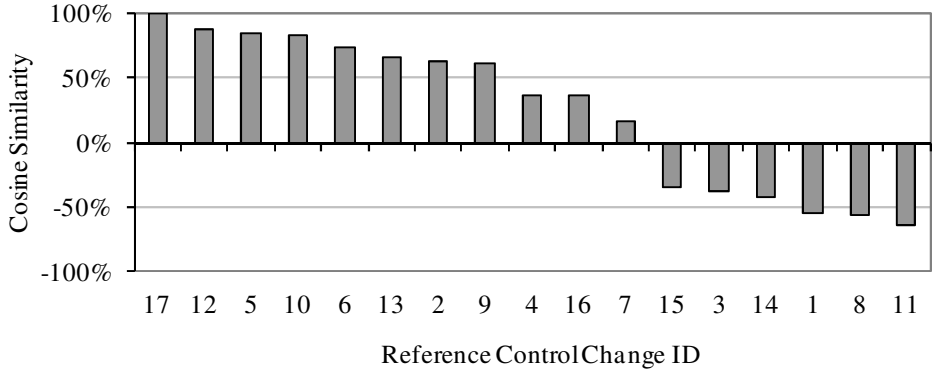

Figure C.17 Representative cosine similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 10)

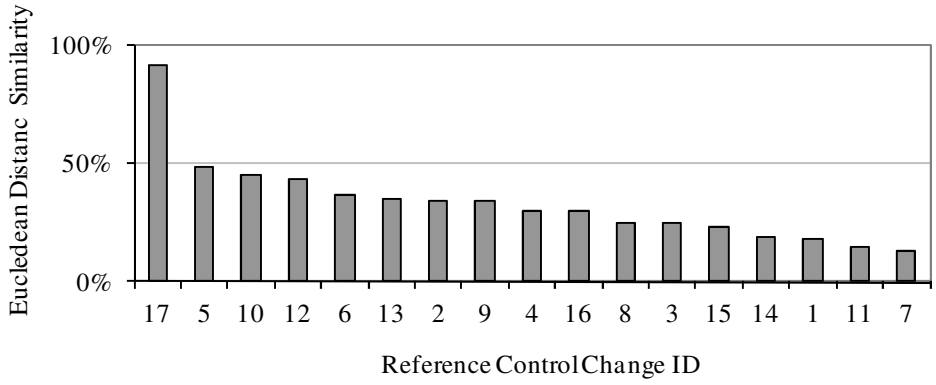

Figure C.18 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Bush Academic Building (Case 10)

Table C.19 Cosine similarity results in Case 1 “Outside airflow ratio increase of 7.5%” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
2	Xoa increase	1.00	1.00	1.00	1.00	1.00	1.00
3	Tpreh decrease	0.11	0.10	0.10	0.09	0.09	0.11
4	Tpreh increase	-0.14	-0.14	-0.15	-0.16	-0.17	-0.14
5	Tcl decrease	0.21	0.23	0.25	0.30	0.34	0.34
6	Tcl increase	-0.12	0.11	0.30	0.44	0.53	0.53
7	Xmin decrease	-0.62	-0.63	-0.60	-0.57	-0.54	-0.54
8	Xmin increase	0.61	0.61	0.61	0.61	0.61	0.61
9	Trc decrease	0.80	0.86	0.90	0.91	0.92	0.92
10	Trc increase	-0.71	-0.67	-0.62	-0.57	-0.50	-0.50
11	Trh decrease	0.03	0.03	0.03	0.04	0.04	0.04
12	Trh increase	-0.03	0.07	0.11	0.12	0.13	0.13

Table C.20 Euclidean distance similarity results in Case 1 “Outside airflow ratio increase of 7.5%” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.16	0.10	0.07	0.04	0.03	0.16
2	Xoa increase	0.38	0.58	0.90	0.72	0.47	0.90
3	Tpreh decrease	0.20	0.16	0.13	0.12	0.12	0.20
4	Tpreh increase	0.20	0.13	0.08	0.05	0.02	0.20
5	Tcl decrease	0.08	0.01	0.00	0.00	0.00	0.08
6	Tcl increase	0.07	0.03	0.02	0.01	0.00	0.07
7	Xmin decrease	0.14	0.08	0.05	0.03	0.03	0.14
8	Xmin increase	0.27	0.23	0.16	0.09	0.04	0.27
9	Trc decrease	0.33	0.25	0.05	0.01	0.00	0.33
10	Trc increase	0.10	0.04	0.01	0.01	0.00	0.10
11	Trh decrease	0.19	0.14	0.11	0.09	0.08	0.19
12	Trh increase	0.20	0.15	0.11	0.08	0.05	0.20

Table C.21 Cosine similarity results in Case 2 “Outside air preheat temperature increase of 3.3°F” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.14	0.14	0.14	0.14	0.14	0.14
2	Xoa increase	-0.14	-0.14	-0.14	-0.14	-0.14	-0.14
3	Tpreh decrease	-0.90	-0.87	-0.84	-0.82	-0.82	-0.82
4	Tpreh increase	1.00	1.00	0.98	0.96	0.94	1.00
5	Tcl decrease	0.42	0.41	0.41	0.40	0.39	0.42
6	Tcl increase	-0.45	-0.49	-0.37	-0.20	-0.04	-0.04
7	Xmin decrease	-0.30	-0.31	-0.32	-0.33	-0.35	-0.30
8	Xmin increase	0.30	0.29	0.29	0.28	0.28	0.30
9	Trc decrease	0.01	0.00	0.00	0.00	0.00	0.01
10	Trc increase	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
11	Trh decrease	-0.78	-0.77	-0.77	-0.78	-0.79	-0.77
12	Trh increase	0.78	0.77	0.77	0.77	0.76	0.78

Table C.22 Euclidean similarity results in Case 2 “Outside air preheat temperature increase of 3.3°F” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.36	0.27	0.18	0.12	0.08	0.36
2	Xoa increase	0.33	0.24	0.17	0.12	0.08	0.33
3	Tpreh decrease	0.25	0.17	0.13	0.11	0.11	0.25
4	Tpreh increase	0.69	0.81	0.43	0.22	0.11	0.81
5	Tcl decrease	0.14	0.01	0.00	0.00	0.00	0.14
6	Tcl increase	0.09	0.02	0.01	0.01	0.00	0.09
7	Xmin decrease	0.27	0.16	0.10	0.07	0.05	0.27
8	Xmin increase	0.38	0.23	0.12	0.06	0.03	0.38
9	Trc decrease	0.24	0.07	0.01	0.00	0.00	0.24
10	Trc increase	0.24	0.10	0.04	0.02	0.01	0.24
11	Trh decrease	0.27	0.19	0.13	0.10	0.08	0.27
12	Trh increase	0.55	0.52	0.39	0.27	0.18	0.55

Table C.23 Cosine similarity results in Case 4 “Minimum airflow ratio increase of 4.9%” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.61	-0.61	-0.61	-0.61	-0.61	-0.61
2	Xoa increase	0.61	0.61	0.61	0.61	0.61	0.61
3	Tpreh decrease	-0.26	-0.26	-0.25	-0.25	-0.25	-0.25
4	Tpreh increase	0.29	0.29	0.31	0.33	0.34	0.34
5	Tcl decrease	0.88	0.88	0.89	0.90	0.92	0.92
6	Tcl increase	-0.79	-0.62	-0.44	-0.26	-0.08	-0.08
7	Xmin decrease	-0.98	-0.97	-0.95	-0.94	-0.92	-0.92
8	Xmin increase	1.00	1.00	1.00	1.00	1.00	1.00
9	Trc decrease	0.27	0.41	0.51	0.58	0.62	0.62
10	Trc increase	-0.09	-0.03	0.03	0.08	0.16	0.16
11	Trh decrease	-0.36	-0.36	-0.35	-0.35	-0.35	-0.35
12	Trh increase	0.36	0.39	0.40	0.41	0.41	0.41

Table C.24 Euclidean distance similarity results in Case 4 “Minimum airflow ratio increase of 4.9%” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.11	0.08	0.05	0.04	0.02	0.11
2	Xoa increase	0.21	0.23	0.21	0.16	0.12	0.23
3	Tpreh decrease	0.13	0.11	0.09	0.08	0.08	0.13
4	Tpreh increase	0.18	0.16	0.13	0.09	0.05	0.18
5	Tcl decrease	0.41	0.05	0.00	0.00	0.00	0.41
6	Tcl increase	0.03	0.01	0.01	0.00	0.00	0.03
7	Xmin decrease	0.08	0.04	0.03	0.02	0.01	0.08
8	Xmin increase	0.33	0.71	0.65	0.29	0.13	0.71
9	Trc decrease	0.16	0.08	0.02	0.00	0.00	0.16
10	Trc increase	0.13	0.08	0.04	0.02	0.01	0.13
11	Trh decrease	0.13	0.10	0.08	0.06	0.05	0.13
12	Trh increase	0.18	0.18	0.15	0.12	0.09	0.18

Table C.25 Cosine similarity results in Case 5 “Room heating set-point temperature increase of 2.3°F” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09
2	Xoa increase	0.09	0.09	0.09	0.09	0.09	0.09
3	Tpreh decrease	-0.67	-0.65	-0.63	-0.61	-0.61	-0.61
4	Tpreh increase	0.76	0.78	0.83	0.84	0.84	0.84
5	Tcl decrease	0.40	0.40	0.39	0.39	0.39	0.40
6	Tcl increase	-0.41	-0.40	-0.27	-0.08	0.12	0.12
7	Xmin decrease	-0.41	-0.42	-0.43	-0.45	-0.47	-0.41
8	Xmin increase	0.40	0.40	0.39	0.39	0.39	0.40
9	Trc decrease	0.17	0.16	0.15	0.14	0.14	0.17
10	Trc increase	-0.17	-0.17	-0.16	-0.16	-0.16	-0.16
11	Trh decrease	-0.98	-0.98	-0.97	-0.97	-0.97	-0.97
12	Trh increase	0.99	1.00	1.00	1.00	1.00	1.00

Table C.26 Euclidean distance similarity results in Case 5 “Room heating set-point temperature increase of 2.3°F” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.27	0.18	0.12	0.08	0.05	0.27
2	Xoa increase	0.29	0.21	0.15	0.10	0.07	0.29
3	Tpreh decrease	0.23	0.16	0.12	0.10	0.10	0.23
4	Tpreh increase	0.48	0.47	0.39	0.24	0.12	0.48
5	Tcl decrease	0.13	0.02	0.00	0.00	0.00	0.13
6	Tcl increase	0.07	0.02	0.01	0.01	0.00	0.07
7	Xmin decrease	0.21	0.12	0.08	0.05	0.04	0.21
8	Xmin increase	0.34	0.22	0.12	0.06	0.03	0.34
9	Trc decrease	0.23	0.07	0.02	0.00	0.00	0.23
10	Trc increase	0.18	0.08	0.03	0.01	0.01	0.18
11	Trh decrease	0.23	0.16	0.11	0.09	0.07	0.23
12	Trh increase	0.54	0.87	0.72	0.45	0.28	0.87

Table C.27 Cosine similarity results in Case 7 “Room cooling set-point temperature decrease of 1.7°F” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	-0.85	-0.85	-0.85	-0.85	-0.85	-0.85
2	Xoa increase	0.85	0.85	0.85	0.85	0.85	0.85
3	Tpreh decrease	0.00	0.00	0.00	0.00	0.00	0.00
4	Tpreh increase	0.00	0.00	0.00	-0.01	-0.03	0.00
5	Tcl decrease	0.07	0.09	0.12	0.16	0.20	0.20
6	Tcl increase	0.02	0.20	0.35	0.46	0.53	0.53
7	Xmin decrease	-0.37	-0.36	-0.33	-0.29	-0.26	-0.26
8	Xmin increase	0.37	0.37	0.37	0.38	0.38	0.38
9	Trc decrease	0.99	1.00	0.98	0.95	0.94	1.00
10	Trc increase	-0.92	-0.88	-0.85	-0.81	-0.76	-0.76
11	Trh decrease	-0.02	-0.02	-0.02	-0.02	-0.03	-0.02
12	Trh increase	0.02	0.15	0.19	0.21	0.22	0.22

Table C.28 Euclidean distance similarity results in Case 7 “Room cooling set-point temperature decrease of 1.7°F” for the Veterinary Research Building

ID	Reference Control Change	Magnitude					Max
		I	II	III	IV	V	
1	Xoa decrease	0.09	0.06	0.04	0.03	0.02	0.09
2	Xoa increase	0.20	0.27	0.33	0.30	0.22	0.33
3	Tpreh decrease	0.13	0.11	0.09	0.08	0.08	0.13
4	Tpreh increase	0.13	0.10	0.07	0.04	0.02	0.13
5	Tcl decrease	0.05	0.01	0.00	0.00	0.00	0.05
6	Tcl increase	0.09	0.04	0.03	0.02	0.01	0.09
7	Xmin decrease	0.12	0.09	0.06	0.05	0.04	0.12
8	Xmin increase	0.14	0.12	0.09	0.05	0.03	0.14
9	Trc decrease	0.39	0.63	0.12	0.02	0.00	0.63
10	Trc increase	0.05	0.02	0.01	0.00	0.00	0.05
11	Trh decrease	0.12	0.10	0.08	0.07	0.06	0.12
12	Trh increase	0.13	0.11	0.09	0.07	0.06	0.13

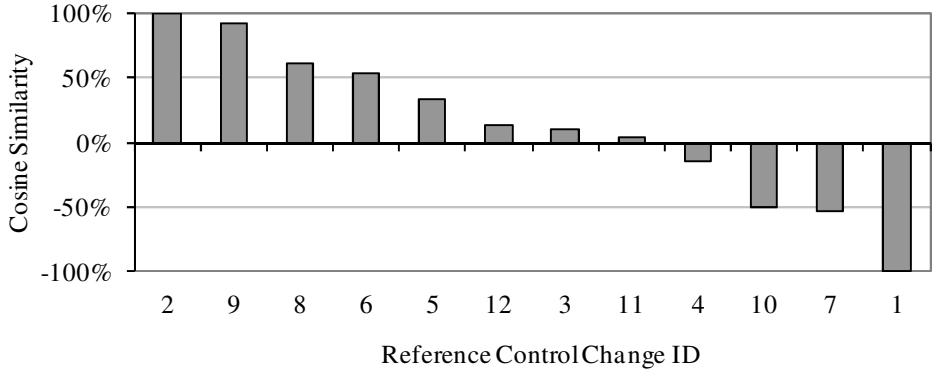

Figure C.19 Representative cosine similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 1)

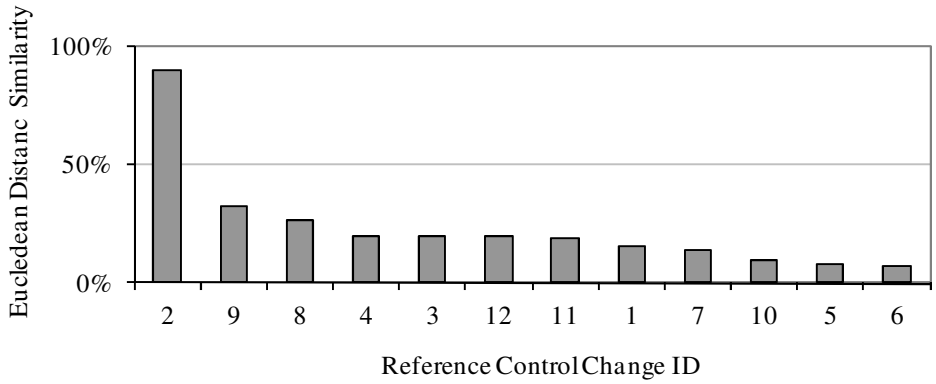

Figure C.20 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 1)

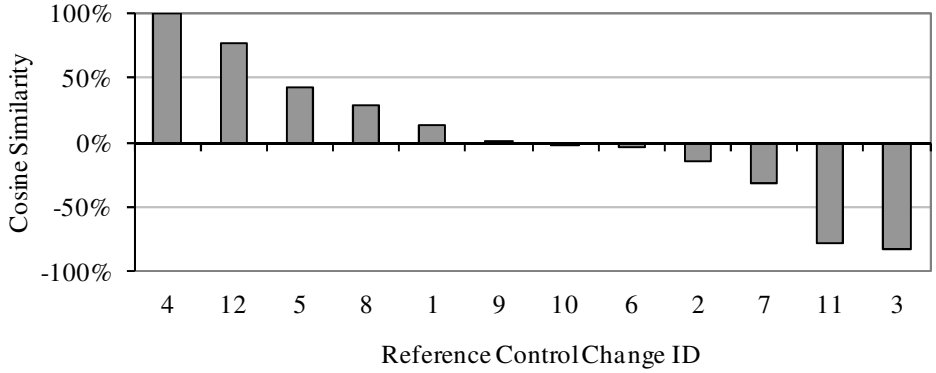

Figure C.21 Representative cosine similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 2)

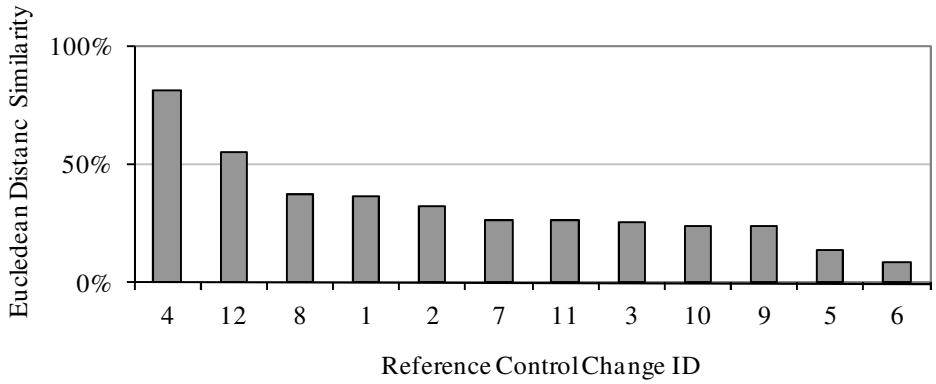

Figure C.22 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 2)

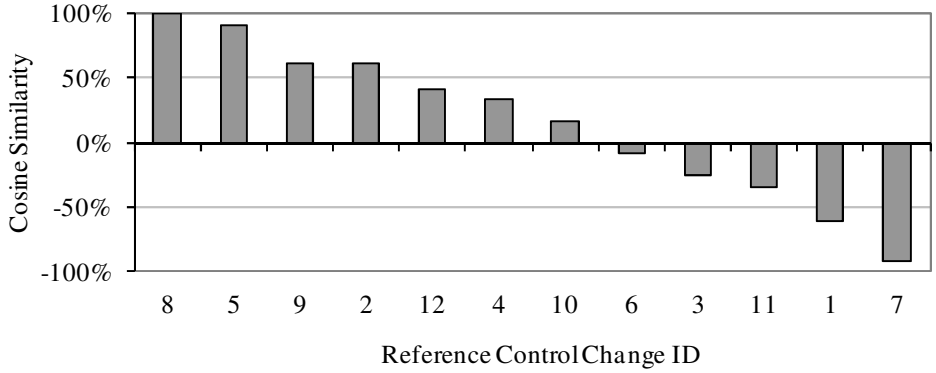

Figure C.23 Representative cosine similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 4)

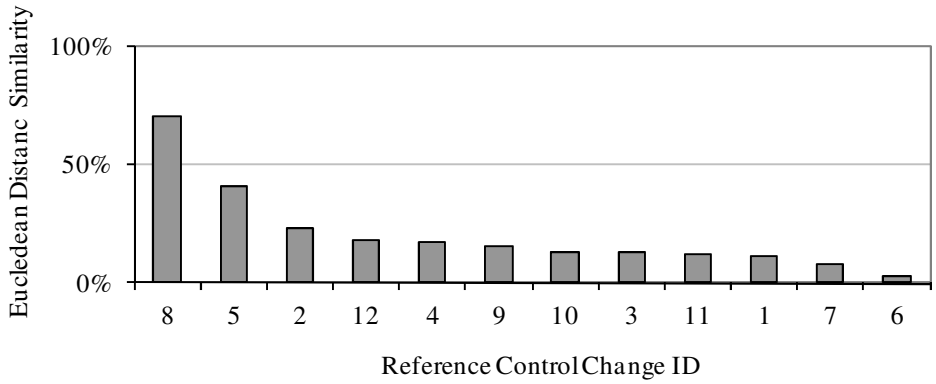

Figure C.24 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 4)

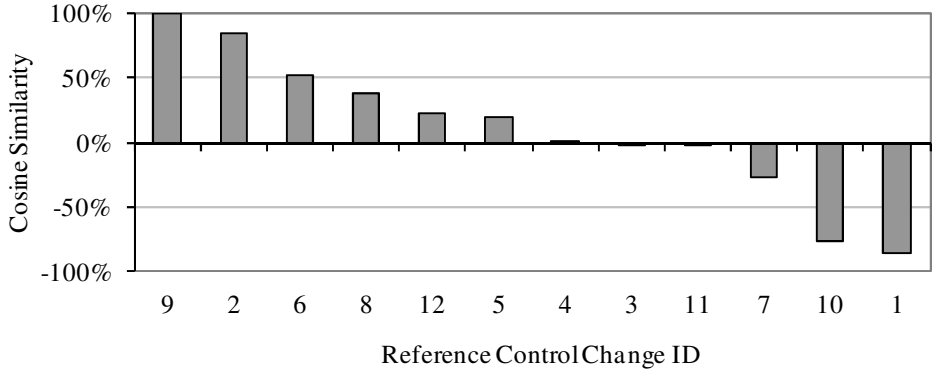

Figure C.25 Representative cosine similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 5)

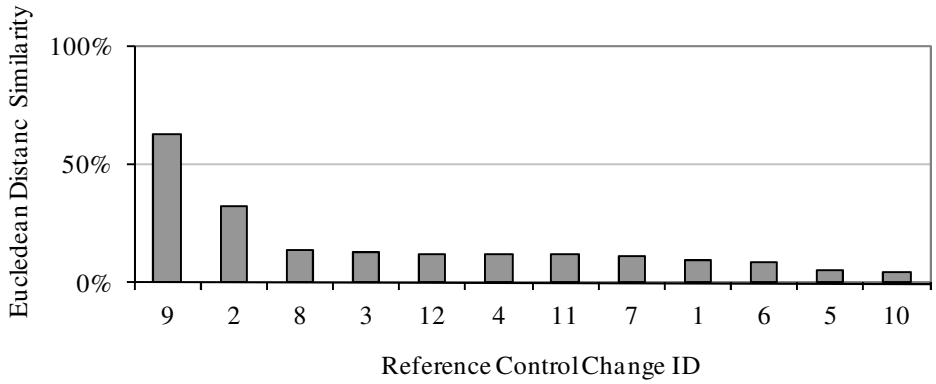

Figure C.26 Representative Euclidean distance similarity for different reference control changes sorted in descending order for the Veterinary Research Building (Case 5)

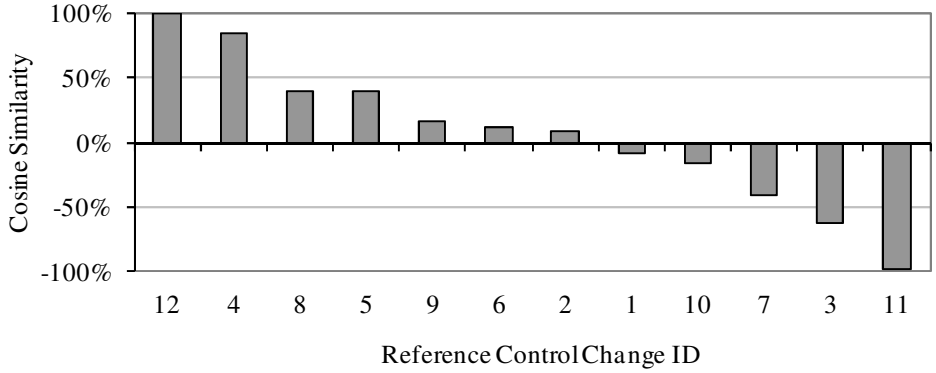

Figure C.27 Representative cosine similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 6)

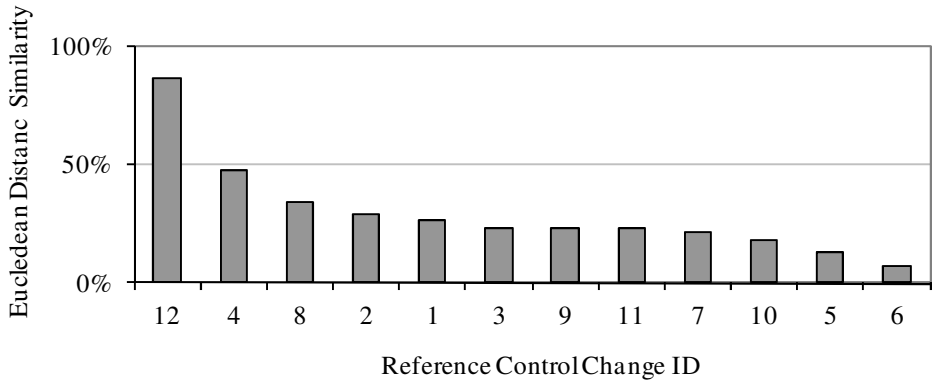

Figure C.28 Representative Euclidean distance similarity values for different reference control changes sorted in descending order for the Veterinary Research Building (Case 6)