Spring 4-23-2015

Theory and Testing of an Energy Signal Tool: An Application of Building Energy Performance Monitoring

Eric David Rodio Boxer
University of Colorado at Boulder, boxere@gmail.com

Theory and Testing of an Energy Signal Tool: *An Application of Building Energy Performance Monitoring*

By

Eric D. R. Boxer

B.S. University of Colorado at Boulder, 2007

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science in

Architectural Engineering

2015

University of Colorado at Boulder

Advisor: Dr Gregor Henze

This thesis, entitled:

"Empirical Testing of an Energy Signal Tool: *An Application of Building Energy Performance Monitoring* "

written by Eric D. R. Boxer

has been approved for the Department of

Civil, Environmental, and Architectural Engineering

_____

Professor Gregor Henze, PhD., P.E.

_____

Professor Moncef Krarti, PhD., P.E.

_____

Professor Balaji Rajagopalan, PhD.

_____

Adam Hirsch, PhD. (NREL)

Date:_____

A final copy of this thesis has been examined by the above signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Boxer, Eric D. R. (M.S., Architectural Engineering)

Empirical Testing of an Energy Signal Tool: An Application of Building Energy Performance Monitoring

Thesis directed by professor Gregor Henze

There is a growing amount of data being collected from commercial buildings that quantifies their energy use. Very few facilities managers have the means to transform this data to determine how the building is performing compared to expectations. Performance benchmarking of buildings can be used to draw this comparison, but first a benchmark must be established. Many tools exist for equipment *fault detection diagnostics* (FDD), but FDD alarms do not indicate fault severity. Other tools exist for rapid peer-benchmarking of energy performance, but statistically meaningful comparison groups are small. An Energy Signal Tool is proposed as a way to self-benchmark performance and quantify fault severity. Being able to quantify performance of buildings in a portfolio can help an energy manager prioritize operational changes or maintenance based on which facilities need attention most.

In this work, detailed building energy simulation modeling of a retail store is carried out with the OpenStudio / EnergyPlus software platform. Uncertainty analysis is used to enhance decision support with a probabilistic approach to energy consumption risk management. Expected model parameter distributions are characterized, and global sensitivity analysis is performed to quantify parameter significance. Latin Hypercube Sampling (LHS) batch simulation is used to sample from significant parameter distributions and generate expected ranges of energy consumption for four major building energy end-uses. The results of batch sampling form a probabilistic range that is used as the performance benchmark.

This work then goes on to demonstrate how sub-metered data can be put into the context of these expected ranges and transformed into plain output for self-benchmarking and energy management decision support with utility theory. The refined concepts of the Energy Signal Tool were tested synthetically in ten different fault scenarios across three climate zones. The results of the testing were processed to illustrate the fault sensitivity of the tool, and to demonstrate how such a tool could be applied to prioritizing actions across a portfolio of buildings. This work builds upon the concepts for an “Energy Signal Tool” originally proposed by Henze *et al.,* (2015), with the goal of making the tool suitable for industry application.

Acknowledgements

The process of conducting research and compiling it all into a masters' thesis was like walking a path, most of the time in the dark, having never been there before, but having some idea of what the destination would look like. Along the way, there were many individuals who helped by shining some light of knowledge, skills, direction, wisdom, encouragement and advice on the path.

Dr Henze, thank you for your patient guidance, enthusiasm, and optimism throughout this work. I also owe much gratitude to Anthony Florita at NREL for sharing his deep understanding of the material with me. Without taking class with Dr. Balaji, I would not have developed the skills necessary to satisfy my enjoyment of data analysis, nor would I have been able to complete this work. Thanks to Lincoln Harmer (M.S., 2013) for passing the torch of uncertainty.

A huge thanks to Adam Hirsch at NREL, whose leadership made this project possible, and who provided great feedback along the way. Thanks to Scott W., our retail partner chief contact, for providing tremendous input and enthusiasm for initiating this work. I wish to express much gratitude to Rois Langner at NREL for the hard work she put into gathering the information and assembling the initial EnergyPlus model of the case study building. This gave me a tremendous head-start on an ambitious project. Then equally immense gratitude goes to the entire OpenStudio development team at NREL, especially Andrew Parker, Dan Macumber, and David Goldwasser for their help in transitioning this model to OpenStudio at the leading edge of interface development. Thanks also to Nick Long and Brian Ball for all the work you have put into the Analysis Spreadsheet so that even a graduate student could run batch simulations in the cloud. Special thanks to Ron Judkoff at NREL, with whom I was privileged to have many meaningful conversations about this work.

To my parents, thank you for encouraging me to develop my own unique set of interests in life, and for all of your love and support. To my partner Krystal, thank you for your love, support, and understanding as we stayed close while living apart for two years during my graduate work.

Contents

List of Figures

List of Tables

List of Equations

Glossary of Abbreviations

.idf	EnergyPlus model file
AIC	Akaikes information criterion
AMY	Actual meteorological year
API	Application program interface
AWS	Amazon web services
BCL	Building Component Library (OpenStudio)
BMS	Building management system
CBECS	Commercial buildings energy consumption survey
CBP	Commercial Buildings Partnership (DOE, NREL)
CVRMSE	Coefficient of variation root mean squared error
E_c	Expected cost (vector)
EC2	Elastic cloud web server
ESTool	Energy Signal Tool
EUI	Energy use intensity (kBtu/ft^2-yr)
FDD	Fault detection and diagnostics
GOF	Goodness of fit
GUI	Graphical user interface
HVAC	Heating, ventilation, and air conditioning
LEED	Leadership in energy and environmental design
LHS	Latin Hypercube sampling
NGAS	Natural gas
NMBE	Normalized mean bias error
NREL	National Renewable Energy Laboratory (Golden, CO)
OAT	One-at-a-time (sampling algorithm)
PDF	Probability distribution function
REFR	Refrigeration
SA	Sensitivity analysis
SI	Significance Index (parameter sensitivity)
SR	Signal priority ratio
TMY	Typical meteorological year
UA	Uncertainty analysis
WBE	Whole building energy
X_{meas}	Measured energy use

1 Introduction

Buildings consumed 40% of total energy in the US in the year 2013 to provide comfortable indoor human environments and essential services (EIA, web). Green building standards such as LEED ® guide energy efficient design, but the energy savings predicted by the design models deviate significantly from actual building usage data (Frankel *et al.,* 2008). There are three possible reasons for this: inaccurate assumptions in the design energy model, improper execution of intended construction, and improper control of building systems. Building performance benchmarking can reveal "faulted" operational states; where actual energy use deviates from intended energy use. Establishing benchmark comparisons for a building portfolio can quantify the severity of these faults, and help an energy manager prioritize operational changes or maintenance based on which facilities need attention most (Hedrick *et al.,* 2011) There are two primary classes of building performance benchmarking: peer comparison and self-benchmarking. This paper is focused on a tool for performing building self-benchmarking.

Databases such as the DOE Commercial Building Energy Consumption Survey and EPA Energy Star Portfolio Manager are popularly used for peer benchmarking. They can provide a rapid assessment of building performance relative to somewhat similar buildings. However, they do not have adequate amounts of data for detailed performance assessments of individual buildings. When the multiple features that define a building (such as year built, size, hours of operation, type of HVAC system, *etc.*) are used to filter the "peer" group, the subset for comparison becomes quite small. With peer benchmarking, it is not possible to see efficiency gains when improvements in one end-use have been offset by another – possibly unrelated – end-use.

The advantage of building self-benchmarking is the ability to measure building performance based on unique and accurate expectations at levels of both the whole

building and detailed end-usages. Building commissioning is the most rigorous method for self-benchmarking in terms of identifying deviation from correct building operation. Mills (2009) showed how building commissioning leads to proven savings in new and existing buildings. However, full building commissioning is an expensive activity. Facilities managers (FM) need a more economical option for identifying and prioritizing energy efficiency measures in their building portfolios.

1.1 Motivation for an Energy Signal Tool (ESTool)

Installing sub-metering infrastructure that monitors and records building energy end-uses can provide rich data for building energy self-benchmarking, and is a prevailing trend for commercial buildings (Guo *et al., 2014*). Having more monitored end-uses and smaller time intervals of data recording can bridge the gap between intended operation and observed actual operation (NSTC, 2011). One recommendation of the NSTC report was for sub-metering information feedback that is *"tailored to intended users to effect operational and behavioral change"* (NSTC, 2011. Introduction, Pg. *x*). The key is that sub-metering itself does not reduce energy consumption – this data must be compared to some benchmark of expected performance. The challenge therein of applying this principle to energy management for a large set of buildings is that sub-metering data can be immense and inherently messy.

Until recently, the use of building energy simulation tools with parameter sensitivity and uncertainty analysis has been mainly for building design optimization purposes, rather than optimizing building operation (see the work of Struck & Hensen, 2007; De Wit & Augenbroe, 2002; Brohus *et al.,* 2009, Hopfe & Hensen, 2011). The rapid advancement in accessible computing power recently has opened the doors to using building energy simulation tools for operational verification.

1.1.1 Thesis Statement

There are vast amounts of building energy data available to facilities managers (FM) right now to aid in tracking performance. However, FM lack the resources to analyze large sets of monitored energy use data, and there is no standard for the transformation of this data to a useable form (O'Donnell *et al.,* 2013). Thus, facilities managers do not know if their buildings perform as they were intended to. Building simulation tools and computer processing power available now can allow for detailed energy modeling to be used as a means of benchmarking operational building performance in real-time.

There is a need for a tool that can transform monitored data into a usable output. This output would help to identify faulted operational states, energy savings opportunities, and also quantify their significance. By acknowledging uncertainty in the model, decision support can be made more credible and accurate. Developing this tool based on open-source energy modeling software commonly used for building design is the path to making this a reality for industry adoption. The first step in this work involved incorporating a meaningful range of uncertainty through parameter characterization and sensitivity analysis. Next, methods of calibrating an energy model to observed data for the purpose of establishing a realistic benchmark were investigated. Third, methods of distilling complex data streams into a simple action signal interface were explored. The action signal output was then tested and tuned for optimal functionality.

1.2 Research Objectives

The outline of this work was developed to meet the needs of a nationwide commercial big-box retailer (see section 1.3). The research was funded in part by a DOE Commercial Buildings Partnership grant through the National Renewable Energy Laboratory. The broad objective was to develop a way to implement the concepts for an Energy Signal Tool, proposed by Henze *et al.* (2015), using open-source detailed simulation modeling software. As such, it was hoped that the retail partner could replicate the effort

and use the re-designed ESTool to improve their current energy management strategies and get more out of their extensive energy sub-metering infrastructure. The tool would allow the facilities manager to assess the measured performance of each building relative to an expected range of its own performance on an ongoing basis. In this way, they would be able to prioritize maintenance and minor retrofit opportunities across the portfolio for energy efficiency and broader organizational goals. While the tool was designed around a big-box retail store case study, the intention is to show that it is broadly applicable to sets of commercial buildings managed under a single portfolio.

Specific objectives of this research were as follows:

- Document the process needed to create a self-benchmarking tool that uses OpenStudio, based on the conceptual design of Henze *et al.,* (2015).
- Make use of sub-metered energy use data for comparison of actual to expected performance.
- Investigate a parameter sensitivity analysis process to screen model variables and rank their significance to the model.
- Develop a structured logic to defining the risk tolerance thresholds for energy end-use monitoring and corrective action.
- Test the theory of the Energy Signal Tool in a realistic environment and adjust it as needed for optimum functionality.
- Demonstrate portfolio level energy fault prioritization.

1.3 Background of Project Development

This work was funded as an operational assessment follow-up to the DOE-sponsored Commercial Buildings Partnership (CBP) work that NREL has been involved with since 2008. The retail partner with whom the Energy Signal Tool concept was developed manages 1800 stores nationwide. In recent years, the retail business has seen degradations in profit margins, and investing in energy efficiency is one source of gains

they have been turning to. The retailerpartnered with NREL to design and construct a new high-efficiency store that opened in 2012.(Hirsch *et al.,* 2014). This has placed the retail partner on track, at least from the standpoint of design, to becoming a leader in energy efficiency. However, through observing deficiencies in actual operation compared to expected energy performance of the store, the retailer believes there are opportunities to reduce energy consumption by additional 2 to 5% through more intelligent active energy management of all their facilities. However, they did not have a benchmark established to identify potential sources of this marginal gain in other stores. Thus, the idea for an Energy Signal Tool to track energy performance was born.

1.3.1 Current Practices for Energy Diagnostics and Optimization at Retail Partner

Operations and maintenance management (O&M) is needed periodically to keep HVAC, refrigeration and other building systems operating as intended. Apart from scheduled equipment replacements, equipment failures will occur randomly due to flaws and other environmental factors. At other times, the operation of control systems needs to be adjusted due to human error or changes in building use. The retail partner manages these issues at stores across the country from a central headquarters. There are no trained facilities experts located permanently in any of the individual stores or regions. Thus, all information regarding facility performance is gathered remotely through BMS interfaces, energy use data logging, and utility bills.

A typical store maintenance call originates from a store manager who reports that the store or parts of the store have been too hot or too cold. According to conversations with the chief engineer, the retail partner currently prioritizes maintenance calls by letting them accumulate until the *estimated* value surpasses the cost of addressing them at any given store. Estimated value is usually based on equipment type and historical experience. When sufficient estimated value has accrued, they initiate corrective action at that location.

Any corrective action requires maintenance personnel or an engineering technician to visit the store. This incurs a cost whether or not the action results in a solution to the problem.

The retail partner does conduct some active energy management of its portfolio. All new stores (over 1000 so far) include enough sub-metering infrastructure to capture major electricity and natural gas energy end–uses. A typical store has 22 unique data points in total, which are listed in Table 1 below. These points are minimally sufficient to capture the energy end-use from major categories such as heating, HVAC, plug loads, and lighting. However, it is impossible with this configuration to have perfect end-use separation at each point; some points capture combinations such as unit heaters and plug loads. The retail partner is aware that better planning of sub-metering infrastructure would make sub-metering data more useful.

Table 1: Typical sub-metered points at a big-box retail store

POINT	End Use	UNITS	Phases	Metered Circuit(s)	CT Location
1	TOTAL LOAD	KW	3	MDS	MDS
2	Sales Flr Plug Loads	KW	1	HP-1	MDS
3	Misc Plug Loads	KW	1	LD-2	HP-1
4	Sales Flr Lights	KW	3	HL-1 & HL-2	MDS
5	Parking Lights	KW	3	HL-4	HP-1
6	Refrigeration Compressors	KW	1	RTCR-1	MDS
7	Refrigeration Cases	KW	1	LP-12A & B	HP-1
8	Refrigeration Walkin	KW	1	LPE-3	LPE-1
9	Refr Door Heaters	KW	3	LP-12A/cct 2,8	LP-12A
10	Dehumidification Unit 1	KW	1	DHU 1/cct 8	MDP
11	RTU 1	KW	1	cct1	HP-1
12	RTU 2	KW	1	cct2	HP-1
13	RTU 3	KW	1	cct7	HP-1
14	RTU 4	KW	1	cct8	HP-1
15	RTU 5	KW	1	cct13	HP-1
16	RTU 6	KW	1	cct14	HP-1
17	RTU 7	KW	1	cct1	MDP
18	RTU 8	KW	1	cct2	MDP
19	Dehumidification Unit 2	KW	1	cct19	HP-1
20	RTU 10	KW	1	cct20	HP-1
21	RTU 11	KW	1	cct7	MDP
22	Dehumidification Unit 3	KW	1	DHU 12/cct8	MDP
23	VAV Reheat	KW	1	HL-3/cct26	HP-1

The retail partner contracts a third-party energy management service to collect the data from the sub-metering equipment. This energy management data tracking service makes energy use data available for all stores in a single online portal. The data can be viewed as time series in graphical or tabular format. From conversations with the engineering team, it was gathered that they analyze this data by drawing comparisons of store energy use intensity (EUI), as well as other sub-metered end-uses, among stores of identical prototype within the same ASHRAE climate region. In this way, an attempt is made to identify which stores are performing the best, or which need attention the most. The challenge in this approach is determining where the performance advantages stem from so that actions to bring high-consumption stores to lower levels can be taken. Another challenge is accounting for unique qualities of each store that affect its energy consumption

Another strategy being implemented is the creation of spreadsheet meta-models that utilize the sub-metered data. These models predict "expected" energy consumption by observing past correlations between observed end-use consumption and weather data. Training a meta-model requires many points of hourly observed data correlated to environmental conditions. The idea is that if performance in one period deviates too much from the expectation generated by the model, then some problem may have arisen. The shortcoming they have encountered with this method is the need to re-tune the model for seasonal shifts, such as when dehumidification equipment is switched on.

A different third-party service is contracted to do some synthesis of the energy use data with building vital statistics data received the BMS – including controls logic, sensor readings and setpoints. They deploy automated logic to look for combinations data anomalies that would indicate equipment faults. This is an example of what shall be termed traditional FDD. For example, with an issue such as stuck OA dampers, they might correlate unusual mixed air temperatures from BMS data with higher than expected heating or cooling energy use over a time period of a week or more. Other issues such as refrigeration

system operation and indoor environmental comfort are more urgent to deal with. These are monitored by tracking sensor readings directly from the BMS. The company has a policy of resolving any issues related to refrigeration systems within four hours to ensure grocery preservation. The Energy Signal Tool is not capable of being so sensitive as needed to address the importance of correcting comfort and inventory preservation issues within days or hours. Rather, the objective is to quantify the severity of faults in a way that allows for non-immediate responses to be prioritized intelligently.

This understanding of operational context brought forth the role of the Energy Signal Tool as more of performance benchmarking tool than fault detection. Fault detection is better accomplished by looking at empirical evidence for specific pieces of equipment and comparing it to rule-based expectations. This way, root causes can be identified directly. Often though, FDD systems produce excessive alarms. This is where the ability of the Energy Signal Tool to quantify the severity of faults in terms of energy consumption deviation would be quite complementary to FDD. The benefits of performance benchmarking for the retail partner are discussed in Section 1.3.3 below.

1.3.2 Current Practices for Optimizing Building Operation at Retail Partner

The retail partner has been working to test and implement energy efficiency measures (EEMs) at stores for over twenty years. However, energy efficiency is a lower priority than the functions provided by the building systems. Therefore, EEM decisions are based on several criteria, in the following order of priority:

- The impact on corporate branding, the "feeling", or image within stores, as well as corporate sustainability.
- The value to improving operational stability. This includes considerations of maintenance costs, product inventory stability, and customer comfort.
- The internal rate of return of the EEM investment compared to other investments the company could be making (energy and maintenance cost savings).

- Energy efficiency resulting from improvements

Apart from the CBP work, new equipment and controls strategies are typically tested one at a time in stores. Their energy performance is typically gauged on the basis of the test period of performance (either utility data or sub-metered energy use data) compared to a past period in a season with similar environmental conditions. The effectiveness of operational stability for an EEM is gauged by waiting to see how they perform in near-design conditions (in the outer tails of environmental conditions distributions). These conditions occur relatively infrequently, and since this is a top priority, the retailer often waits through a long test period before mass implementation. When optimization practices are finally validated, quantifying the potential impact and prioritizing the rollout among stores is an energy management challenge that currently has no methodical approach. This is currently done with rough rule-of-thumb guidelines.

The retail partner has a corporate practice of making *portfolio-wide* rollouts of energy efficiency upgrades and control strategies to the extent that they are climate appropriate. This is in keeping with the philosophy of simplifying operations management across the portfolio as much as possible. In one example, they tried adjusting indoor thermostat setpoints. Lowering sales floor heating setpoints in all stores from 70^0F to 68^0F was successful. However, adjusting the cooling setpoint from 74^0F to 75^0F resulted in comfort complaints, and so the change was reversed.

1.3.3 Benefits of an Energy Performance Benchmarking Tool for Retail Partner

The Energy Signal Tool concept came about through discussions between NREL and energy management leadership at the retail partner. They have observed the need for a more proactive approach to energy management that would help energy managers to quantify the severity of energy faults across the portfolio of stores, rather than uncovering them slowly beneath mountains of data. The retailer is aiming to meet the operational targets for the CBP case study building, as well as to acheive a small marginal increase in

efficiency beyond the efforts from CBP (estimated at 2-5%, as noted above), so the tool must also be inexpensive to implement, requiring little additional infrastructure. One alternative considered was to install additional sensors and enter into an extended performance diagnostics contract with their third-party energy management service. This was deemed too costly for the marginal benefit available.

Incorporating all company priorities into energy-related decision support of the tool has promise to further enhance the basis of decision making beyond just available energy savings. The research team decided to investigate *detailed* energy modeling in conjunction with a statistical tool that leverages an existing sub-metering infrastructure. A detailed energy modeling approach gives a far superior benchmark compared to spreadsheet meta-modeling does. This comes from its level of detail and the ability to incorporate knowledge of the buildings. With a detailed energy model, it is possible to input seasonal fluctuations in schedules and changes in control strategy that help the model to match the patterns of actual operation. With a detailed simulation model, observed environmental factors act on all aspects of the building (envelope, equipment, controls, etc.), rather than being correlated to relatively few points of sub-metered equipment data available. As opposed to a spreadsheet based model, a detailed model can also be used to *predict* the effects of changes to the building, and does not have to be completely re-derived when changes are made (Coakley *et al.,* 2014).

This tool could *supplement* the FDD information on operational faults by providing information that quantifies their severity. For energy management, an Energy Signal Tool promises to help the company to prioritize actions rather than operating in a mode of reacting to problems as they arise. Beyond performance assessment, a side benefit of having the tool based on detailed energy models of each store would be the ability to test energy efficiency strategies virtually. Essentially, this would entail running the energy signal tool in reverse; making changes to a model and using the new output as the

"observed" consumption. Then, EEMs could be ranked by the signal output. It would make it simple to quantify the potential savings impact of a measure on each store, and rank stores in order of implementation priority. This would also result in the ability to test EEMs in design conditions where the limits of comfort and functionality are most likely encountered. It would eliminate the traditional waiting time of performance validation, and reduce the risk that adverse performance would affect comfort or inventory.

1.3.4 Further Motivation: Results of CBP Comparison for Case Study Building

The retail partner has worked with NREL to design new stores that conform to the *50% Advanced Energy Design Guide for Big Box Retail* standards created with ASHRAE (Bonnema *et al.,* 2012). One such pilot store in Wisconsin, constructed in 2012, was expected to exceed ASHRAE 90.1-2004 energy performance standards by *52*%. This store is the case study building used in the remainder of this report. It was equipped with extensive sub-metering equipment. A design energy model was also available for comparison of performance to predicted levels. In the first eight months of monitored operation at the store (August 2012 to April 2013), 34% energy savings were observed. This was much less than the model expectation of 52% (NREL, 2013a). See Figure 1 below for a breakdown of energy expectation and consumption by end-use.

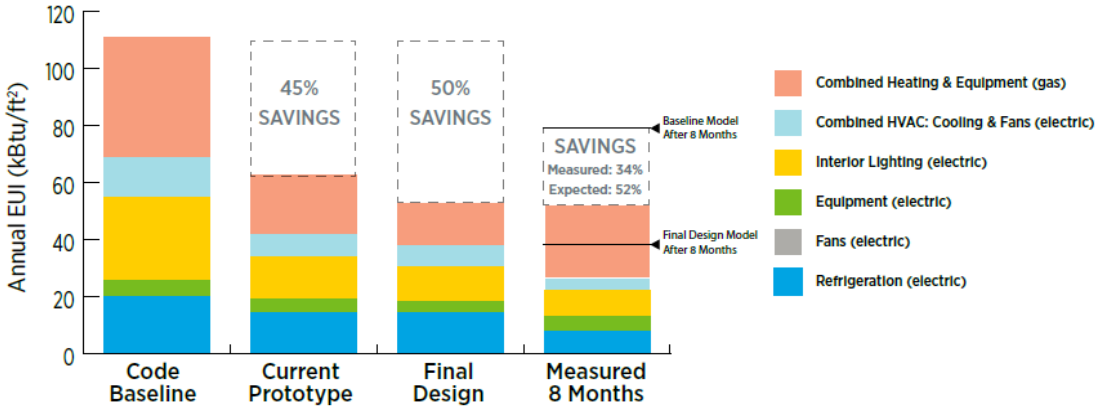


Figure 1: Comparison of actual savings to expected savings for the Wisconsin AEDG prototype

Image source: NREL (2013a)

Figure 1 shows that most of the deviation in consumption was due to excess natural gas use, which had been 180% of the predicted amount in the first eight months of store operation. The advanced pilot store was performing *worse* than standard prototypes. Following this report in September of 2013, NREL conducted a follow-up commissioning investigation and found that there had been faulty controls in the outdoor air unit operation schedule. The unit had been turning on at 3 a.m. on a nightly basis, introducing unnecessary amounts of outdoor air into the building. Correcting this issue eliminated some, but not all, of the unexpected degradation in energy performance. After the HVAC controls issue was corrected, the store achieved a 44% increase in performance over the ASHRAE 90.1-2004 baseline (Hirsch *et al.*, 2014). The expected baseline generated by the detailed energy model helped to demonstrate the existence of a problem. Acknowledging uncertainty in the baseline expectations, and then investigating some of the model uncertainties in operation, could have further validated the results. This evidence gave the retail partner insight into how beneficial uncertainty-based energy performance self-benchmarking can be.

2 Literature Review

2.1.1 Parameter Uncertainty Analysis and Sensitivity Analysis

Whole building energy models have thousands of input parameters that together characterize the usage patterns, operational strategy, and physical properties of a building and its systems. Even for existing buildings with extensive construction and operational documentation, some parameters such as outdoor air infiltration rate and system sizing factors cannot be specified with certainty. Many of these factors will also vary randomly over time. A single model is only the best guess representing a collection of guesses made for many unknown and un-measureable input parameters. Various combinations of parameter values in the model will result in a range of model outcomes. . UA is applicable in

the context of decision making based on simulation models, as it acknowledges the unknowns that go into establishing a baseline(de Wit, 2003).

Prior to utilizing UA, uncertain model inputs must be characterized with probability density functions that describe the range of probable values for each. Dominguez (2009), MacDonald (2002), Corrado (2009), and de Wit (2003) all provide information on which to base assumptions regarding the distribution of parameter uncertainty. Many assumptions though are still at the discretion of the user. Struck and Hensen (2007) provide examples of making assumptions for uncertainty in un-measurable physical parameters. Struck and Hensen (2007) demonstrate that UA is useful for supporting multi-criteria decision support in building design. They note that the early design phase of a building project is characterized by great uncertainty in the path leading from there to the final outcome. They use Monte Carlo analysis to generate the complete range of building performance uncertainty as it varies with the value of all uncertain parameters.

Studies of sensitivity analysis (SA) explore how model objective function outcome is affected by each model input parameter (Saltelli, 2000). Combining sensitivity analysis with UA can be used to quantify the impact that each uncertain parameter has on the model outcome. From here, model dimensionality (degrees of freedom) can be reduced by ignoring the effects of those uncertain parameters that have little impact on the model (Saltelli, 2000; Hopfe and Hensen, 2011). Burhenne (2013) notes that the uncertainty distributions of significant parameters must be examined carefully, and in this way, SA/UA can become an iterative process. Saltelli (2000) also notes that one of the primary reasons for performing SA is to determine which model variables merit additional research for improved estimation.

Burhenne (2013) used Monte Carlo analysis to explore a range of possible design methods for achieving Passivhaus standard residential construction. Monte Carlo filtering of the results made it possible to uncover which building parameters had the greatest

influence on design outcome, and thus which were the best investment. Along with building design parameter uncertainty, Burhenne explored economic and energy price uncertainty. He found that combining cost-benefit analysis with the SA/UA process created a decision support system that drew upon information about how present design investments would translate to future return.

Booth and Choudhary (2013) propose a framework of UA for assessing the potential impact of energy efficiency policy in UK housing. They use uncertainty assessment of parameters coupled with Bayesian inference to updates beliefs about parameter distributions. This is followed by an assessment of expected ranges of outcome for various retrofit measures that can be taken to maximize net present value. Their goal was to minimize the financial risk to the UK government by providing a range of expected outcomes that result from retrofit measures. With this information, the UK government can prioritize their spending based on the likely impact it is likely to have.

2.1.2 Building Energy Performance Benchmarking and Decision Support

Performance indicators (PI's) are quantifiable indicators that adequately represent a particular performance requirement. Verification methods to evaluate whether PI's have been met can include experiments, models, or tests; or a combination. Pati *et al*., (2006) argue that creating a common vocabulary of performance indicators can help to bring all of the disparate stakeholder needs together in the building design process for an optimal product. Similarly, defining operational performance indicators in a building can help bring diverse interests of energy management, maintenance, customer/employee satisfaction, and corporate strategies together.

Energy use is one performance indicator. Hedrick *et al.,* (2011) demonstrated an internal benchmarking system for energy use in restaurant portfolios. They were motivated by a study that found using the EPA ENERGY STAR portfolio manager was an insufficient way to compare energy use among restaurants. This is due to the multitude of

factors such as menu, kitchen appliances, and hours of operation that affect building energy use intensity. The study acted on the recommendation of creating a benchmarking system unique to each multi-use operator as the best way to evaluate performance. In the paper, the authors took a statistical approach to assessing performance. Regression models for expected performance were derived from best correlation of independent variables such as hours of operation, facility type, and observations in weather. This type of approach has the advantage of simplicity to the user, but may not be accurate if key independent variables such as local HDD and CDD are missing or inaccurate. It also had the disadvantage of only identifying poor performing stores on a qualitative basis; only revealing whether or not one particular store is an outlier in the data set.

In decision making with regards to optimizing building performance for energy, comfort, and other metrics the facilities manager has a tremendous amount of information to deal with. Modern BMS systems are capable of logging data being produced by hundreds of sensors in a single building. Energy sub-metering may track dozens of points, ranging from high-level end uses, such as total lighting energy, to very minute levels, such as individual refrigeration compressors. The facilities manager is largely unprepared to deal with this bombardment of information. O'Donnell *et al.,* (2013) propose that data transformation is crucial to helping facility managers do their jobs well without having to rely on experts. The authors note that, despite the availability of extensive monitoring data, there is typically the crucial step of *checking* missing from the "*plan, do, check, act*" cycle. They demonstrate an automated, rule based methodology of transforming data from a wide array of information sources into a holistic performance assessment tool. This tool takes into account performance indicators of function, efficiency, thermal loads, performance and legislation.

Doukas *et al.,* (2009) proposed an intelligent decision support model for improving building energy performance. These authors built upon the work of others in response to

several high-level European energy efficiency directives. They identified the need for an energy management decision support tool that utilizes of the data collected by the BMS. They use this data as "experience data" to help train a model to recognize abnormal operation in various components tracked at the BMS. Standard indices derived based on energy efficiency standards, prevailing weather conditions, and indoor environmental conditions are used to prescribe a "normal" performance level. They propose forming a priority list for intervention based on comparison of measured building operation to the standard indices of performance. This method of ongoing performance assessment requires that the building have very robust BMS infrastructure. It also requires the derivation of benchmark indices for each unique case. In the five years since this publication, advances in computing power have made the use of detailed simulation for performance benchmarks more realistic.

Costa *et al.,* (2013) investigated ways of using calibrated building models to help facility managers determine when buildings have faults without the need to rely on external expertise. They found that data availability allows for easy application of manual fault detection and diagnostics, and concluded by stating the significance of developing an automated system of FDD. They provide some examples of creative data visualization options that would help the facility manager to identify faults. The authors state explicitly that work to automate the FDD process, and "*integration of the FDD activity within an action management process*" (Costa *et al.,* 2013; *Pg.* 6), would greatly enhance their proposed methodology.

Henze *et al.,* (2015) have developed a conceptual prototype for an Energy Signal Tool for operational performance decision support. The authors developed the theory for a tool that alerts facilities managers of building consumption anomalies across a variety of monitored end-use categories. They show that sub-metered data and a range of uncertain operational expectations can merge with utility theory and be processed into the quick

visualization of a five-level traffic light. The resulting traffic light represents the action with the lowest *probable* cost thus acknowledging the fact that models will not perfectly reflect reality. The authors based their investigation on reduced order models developed in the Matlab environment to facilitate statistical processing of input and output data. They have also presented the idea that knowledge of parameter uncertainty can be updated over time using Bayesian inference, which may reduce the need for future model calibration (Pavlak *et al.,* 2013).

This thesis work builds upon the concepts for an Energy Signal Tool (ESTool) as proposed by Henze *et al.,* (2015). In reference to this recent work, additional development effort went into transitioning the ESTool model platform to open-source building energy simulation software common to engineering practitioners. Efforts were also made to give a solid basis for parameter sensitivity analysis, and incorporating these results into risk management. There is little previous research on using SA in the context of risk tolerance. A reasoned and methodological approach is proposed for setting risk tolerance thresholds for each end-use. The application of utility theory is enhanced by accepting cost input from an ordinary user with several organizational objectives in mind. This is a novel approach to customizing a cost matrix. This work also proposes that the Energy Signal Tool is most useful in the realm of portfolio energy benchmarking and retrofit prioritization, rather than FDD.

2.1.3 Energy Management and Performance Diagnosis Practices in Industry

Chung (2011) distinguishes between two types of building energy use benchmarking systems: those intended for public information, and those which generate results for internal use only. Public benchmarking systems, the author notes, are useful for promoting better energy performance in buildings with high interest in public image. Internal benchmarking, on the other hand, is most useful for internal decision making. One example of each type of benchmarking commonly found in industry is presented below.

Energy star portfolio manager

The US EPA offers vast sources of information on building performance assessment and benchmarking through their Energy Star program (www.energystar.gov, web) administered by the Environmental Protection Agency. The site reports a score of 1-100 (100 being the best) as a metric of a building's energy performance. The result is calculated based on a peer group of buildings, with similar location, construction, and operating characteristics. Rather than simply being a percentile rank within the peer group index, the Energy Star score is a product of "*a statistical regression model that correlates the energy data to the property use details" (www.energystar.gov)*. The EPA Energy star system is the most rigorous peer benchmarking system of its kind in the United States (Gao and Malkawi, 2014). It is commonly used for high-level performance assessments and portfolio managers to index building performance within a portfolio.

While the Energy Star score is certainly a good starting point for assessing overall building energy performance, there are several big problems with using this information for active energy management or detailed assessment. First, the CBECS data (http://www.eia.gov/) used to build the statistical regression model is maintained in a notoriously out of date fashion. The most recent CBECS data available to date comes from a 2003 survey; the results of which were not released until 2007 (eia.gov[1]). The CBECS peer comparison groups turn out to be rather small, if available at all. Thus, the aforementioned statistical regression model may be valid only for certain buildings types with high representation in the CBECS. For example, with the DOE Buildings Energy Data Book (http://buildingsdatabook.eren.doe.gov/), it is possible to select criteria matching a big-box retail super-center in Denver, Colorado built in 1998 and create a peer group from the

[1] At the time of writing an update to CBECS, with data from a 2012 survey, was expected in late 2015.

[2] Calibration of a model could only be completed to utility data because the sub-metered natural gas use was not available. Furthermore, the suite of software used permitted only monthly data calibration.

[3] A "measure" is a scripted method, as opposed to GUI or text-based, for altering elements of an OpenStudio

2003 CBECS dataset. With only these constraints (HDD>5500, CDD<2000, retail store, 50,000 to 200,000 ft^2, and built between 1990 – 2003) the sample size returned is four. In this case, it would not be possible to obtain an Energy Star score. To generate a more substantial peer group, one may search, as an example, for all buildings of type "Office", built between 1990 and 2003, with sizes of between 25,000 and 100,000 square feet, and obtain a sample size of 49. This however neglects many distinguishing features of the building that are large factors in energy use, such as climate zone, hours of operation, plug load intensity, envelope composition, *etc*.

Researchers presented a study in 2002 on using the EPA Energy Star Portfolio Manager benchmarking system to assess similar primary education facilities in the Northeast. (Hinge *et al.,* 2002). They found that scores depended heavily on the amenities of the building (some schools did not have ventilation systems up to code), and the experience level of the facilities management personnel. In one case, a school that updated its ventilation system so that it was providing code minimum requirement saw its Portfolio Manager score drop from 36 to 16. This major update should have changed the peer group for comparison, but CBECS data is not detailed enough to allow this. In this case, a fair assessment of performance would have placed the energy consumption of the school below the expected amount before the ventilation update. This would be possible with self-benchmarking. Gao and Malkawi (2014) state the need for a benchmarking method that is based on a multitude of features, and their work proposes a clustering method by which to classify buildings.

Energy charting and metrics tool (ECAM)

ECAM is a retro-commissioning tool, developed by PNNL, which seeks to leverage the large amounts of point measurement data collected by the BMS for a single building (Taasevigen and Koran, 2012). ECAM allows the user to input usage and occupancy schedules, as well as outdoor temperature. However, ECAM fails to include some factors

related to individual building performance, which can only be accounted for by a detailed simulation model. The ESTool is capable of taking information relating to occupancy, change in controls sequence, and equipment repairs/upgrades and generate forward-looking energy use benchmark projections that require minimal user time, and minimal analysis skill, to process. ECAM has excellent data processing tools that allow the user to observe and compare energy use profiles. It can be used to set up performance trend comparisons (*e.g.*, comparing cooling energy use across three different summers), but changes in external variables such as weather, occupancy and changing building use that influence cooling energy use make this of limited value under close scrutiny.

The other major difference between ECAM and the ESTool is UA. Where ECAM makes comparisons on a point-to-point basis, the ESTool considers the likely range of expected outcomes, based on uncertain parameters, and reports how actual consumption compares to this. ECAM can be quite useful to help visualize large amounts of BMS data and to identify areas in a specific building where systems are not operating as they should be, but it does not give the user and indication of how urgent it may be to address these specific operational deficiencies. The ESTool can prioritize issues to address based on a probable range of *expected cost* compared to the operation-as-intended baseline. Another summary difference is that ECAM is excellent for making use of all BMS data collected for a building for purposes of performance issue diagnosis, whereas the ESTool gives a high-level view of each building's performance within a portfolio of buildings, and the quality of output depends on the user knowledge of the building. In this way, ECAM and ESTool may be complementary, ECAM can be used to gather information about individual buildings, and the ESTool can be implemented to *compare and prioritize* efforts at addressing the gaps between expected and observed energy use. Then, when significant gaps are identified, the user can go back to ECAM for those specific buildings and analyze specific systems and equipment in greater depth.

3 Methodology Overview

This work took place over a 14-month period beginning in January of 2014. It was carried out under the supervision and guidance of thesis advisor Prof. Gregor Henze (CU-Boulder/NREL joint appointment) and project engineer Dr. Adam Hirsch (NREL).

As mentioned above, a big-box retail industry partner provided a case study building with which to carry out the investigation of the Energy Signal Tool. This building had been previously studied by the National Renewable Energy Lab (NREL, 2013a). A full energy audit was completed upon store opening in the summer of 2012 and a retro-commissioning study was done in late summer of 2013. A detailed building energy model had been created in EnergyPlus based on the as-built construction documentation, as well as information about store operation gained from the audit.

The retail partner provided the research team with access to an online database of utility information and hourly sub-metered data for the case study building. This had also been in place since the CBP work described above. At the initiation of this project, there was great enthusiasm and involvement from engineering management at the retail partner. There was even a vision to see this project through to a case study implementation at engineering headquarters.

3.1 Scope of the Work

Since the industry partner initiated the project and expressed early interest in working with the research team to test a working implementation, the author was fully committed to going through the process with data from an existing building. A realistic implementation of a self-benchmarking tool involves working to calibrate a detailed model to observed data. To this end, a significant amount of effort went into gathering data

describing the operation of the case study building and calibrating a model to match this data. This effort is described in Chapter 6. Unfortunately, it became evident in the eighth month of the project that industry partner was no longer available to commit resources to the project. Due to this unexpected change, there was insufficient information available to the author to properly investigate how the building operational intent corresponded to the operational data that was available. Additionally, it was discovered that the sub-metered energy data set was significantly flawed; this leaving still more uncertainty. Therefore, efforts were turned towards additional statistical procedures and robustness tests of the Energy Signal Tool theory. Results of this are found in Chapter 8.

3.2 Methods and Tools Used

3.2.1 Methods

In general, project workflow was as follows:

1. Gather input from the industry partner about building operational strategies and recent history of the case study building
2. Complete a detailed model of case study building with OpenStudio
3. Gather performance data for case study building, processing into model information
4. Characterize parameter uncertainty
5. Pre-screen parameters with local sensitivity analysis
6. LHMC sampling for calibration to monthly utility data[2]
7. Multivariate regression and global sensitivity analysis for parameter sensitivity quantification
8. Cluster analysis of the calibrated models solution set to filter parameter values and select a representative model

[2] Calibration of a model could only be completed to utility data because the sub-metered natural gas use was not available. Furthermore, the suite of software used permitted only monthly data calibration.

9. Decision tool development, including definition of risk tolerance thresholds, and incorporating user input into the cost matrix
10. Test the tool on synthetic fault scenarios generated by modifying the representative model
11. Classify signals displayed for skill in accurately identifying faults
12. Investigating ideas for further development of a user-interface

Figure 2 below gives a graphical overview of methods one through eight as listed above.

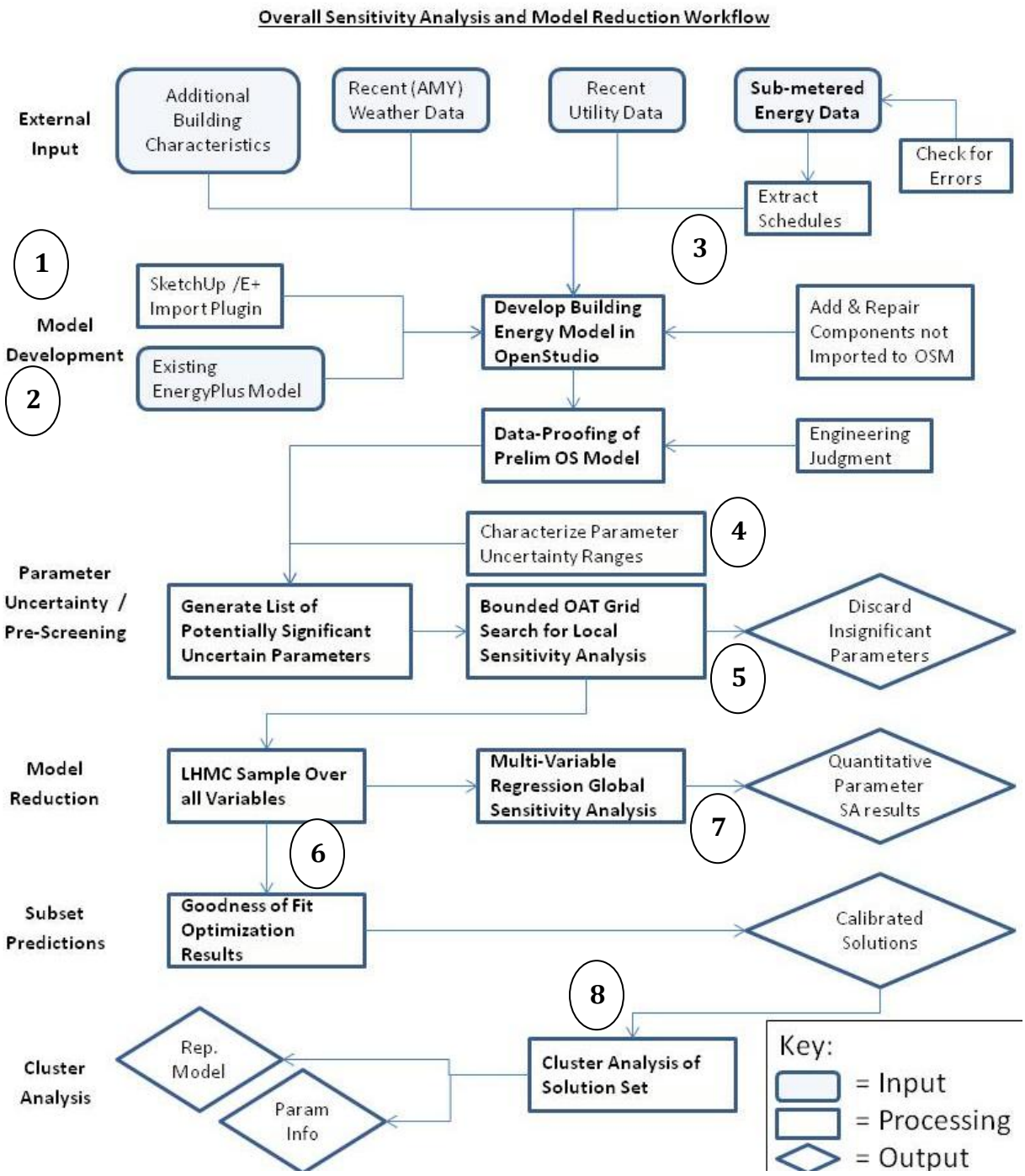


Figure 2: Workflow of model development, sensitivity analysis, and model reduction

3.2.2 Tools

It was decided to utilize OpenStudio, rather than EnergyPlus, for the detailed simulation component of the project. This would leverage the batch-simulation capabilities of the OpenStudio Analysis Spreadsheet and cloud computing power of OpenStudio server

(see Macumber *et al.,* 2014[i]); both under development at NREL. The author had the great fortune to work with the NREL OpenStudio development team to transition the model over to OpenStudio format. This required writing both OpenStudio and EnergyPlus measures[3] in Ruby script to create or modify objects that OpenStudio was not set up to handle (discussed in a later section). With the OpenStudio GUI, it was simple to further develop the model with additional data gathered such as sub-metered energy use, and building construction documentation.

Information regarding parameter uncertainty was managed with the OpenStudio Analysis Spreadsheet. The Analysis Spreadsheet communicates with the OpenStudio Server package to set up and execute large batches of model runs. These are executed on the Amazon Web Services (AWS) cloud computing platform. In this way, a typical batch of 1,000 model runs lasted about three hours at a cost of about $8.00. Results from batch simulations are conveniently output in R data-frame format. Because of this convenience, and for the extensive plotting and graphical display capabilities of the R environment, R was used to process data and generate results throughout this work. Figure 3 below gives an overview of the sequence of software tools used in the project.

[3] A "measure" is a scripted method, as opposed to GUI or text-based, for altering elements of an OpenStudio model.

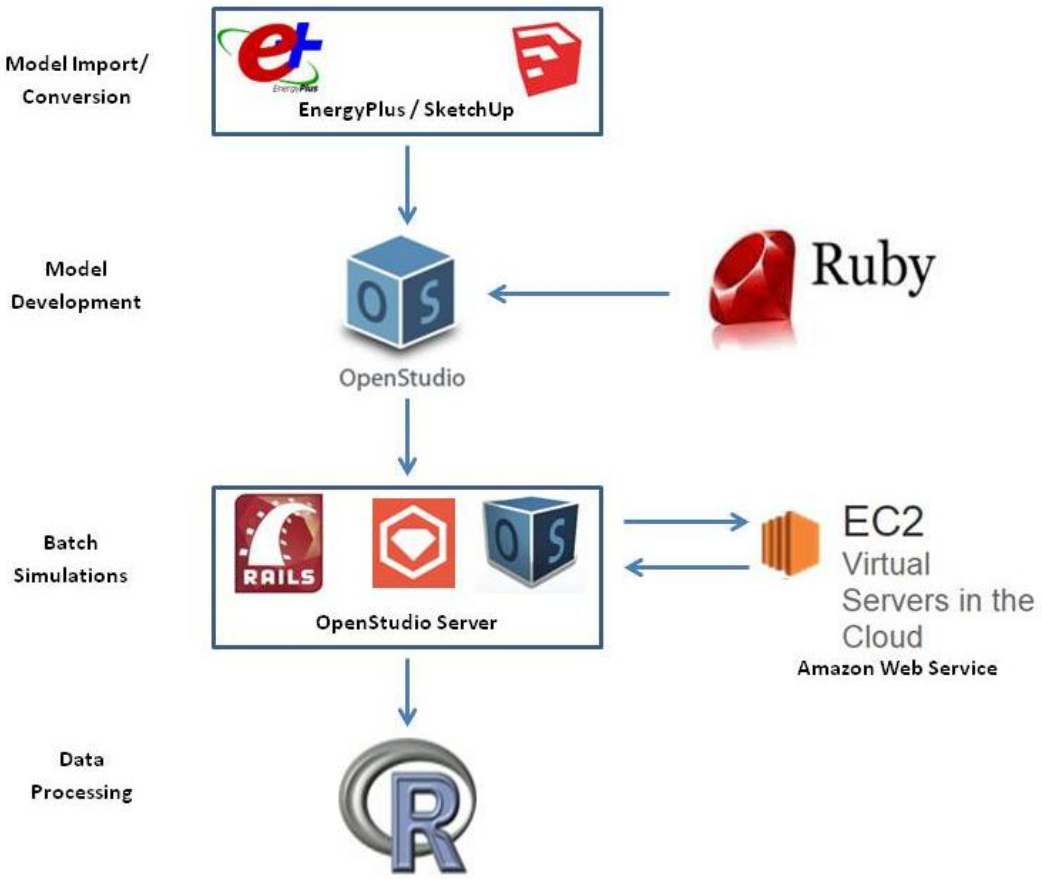


Figure 3: Workflow of software tools

3.3 Case Study Building Information

Figure 4: The case study building upon construction completion.

(Photo by Rois Langner, NREL.)

The case study building is a single story, 133,000 ft^2, big-box retail store located in suburban Wisconsin. It contains a large sales floor, a pharmacy area, a food vendor, a small kitchen, stocking areas, walk-in refrigerated storage, and a grocery area with refrigerated display cases. As mentioned above, the retail partner worked with NREL to incorporate multiple energy efficiency features across different building sub-systems in the design process in an attempt to achieve 50% energy savings compared to an ASHRAE 90.1-2004 baseline. At the envelope level, this included walls with R-12.3 insulation and R-25 roof insulation. Due to the sensitive nature of detailed information, exact specifications and floor plans for the case study building cannot be shared in this report.

Figure 5 below displays a high-level summary of energy use history, which spans over two years from when the store opened in July 2012, to the latest available data. Figure 6 and Figure 7 show how energy use corresponded to weather conditions in the first two years. Data was collected from monthly utility billing summaries and normalized to calendar months. For the period of data corresponding to the calibrated model used throughout as the case study, the store had an energy use intensity (EUI) of 68.4 kBtu/ft^2-yr. This compares to an average of 76.1 kBtu/ft^2-yr for retail buildings of similar size in the CBECS database (2003 survey) in the same climate zone[4]. The ASHRAE 90.1-2004 code minimum baseline model predicted the store would consume 111 kBtu/ft^2-yr (NREL, 2013a).

[4] CBECS data source: http://buildingsdatabook.eren.doe.gov/CBECS.aspx [accessed: 02/13/15] Categories selected for Climate Zone - CDD <2000, HDD>7000; Building size – 50,000 to 200,000 ft^2. The size of the CBECS survey sample set was seven buildings. The fact that this is a relatively small improvement in energy efficiency, despite being designed for superior energy performance, is most likely due to the extended service nature of the case study store; which contains small restaurant, extensive refrigeration equipment, and has relatively long operating hours.

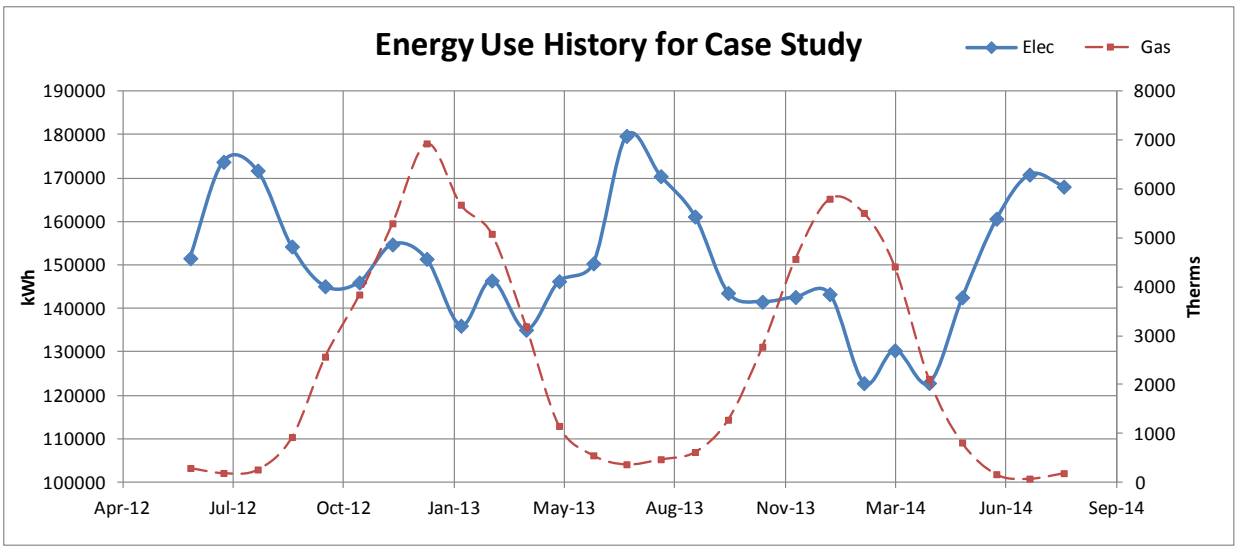


Figure 5: Energy use history for case study building

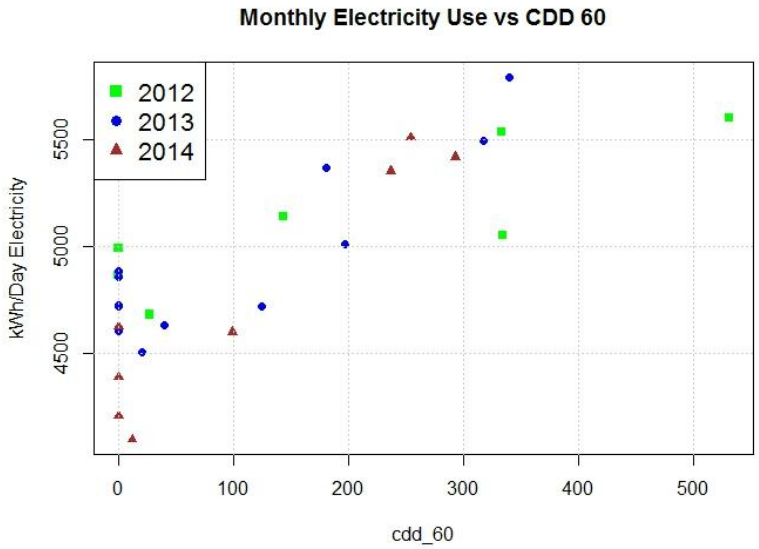


Figure 6: Plot of monthly electricity use vs. cooling degree-days observed

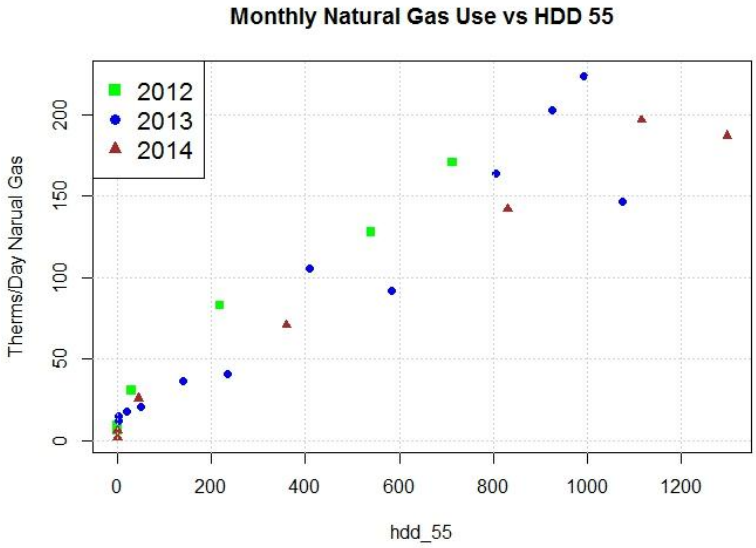


Figure 7: Plot of monthly natural gas use vs. heating degree-days observed

As Figure 6 and Figure 7 above illustrate, the inconsistent operational strategies at the case study building have not generated energy use data with regular correlation to outdoor conditions. This is also partially due to energy use being dominated by internal loads, being approximately 58% of total energy consumption. Figure 8, below, gives a summary of energy consumption by end use at the case study store.

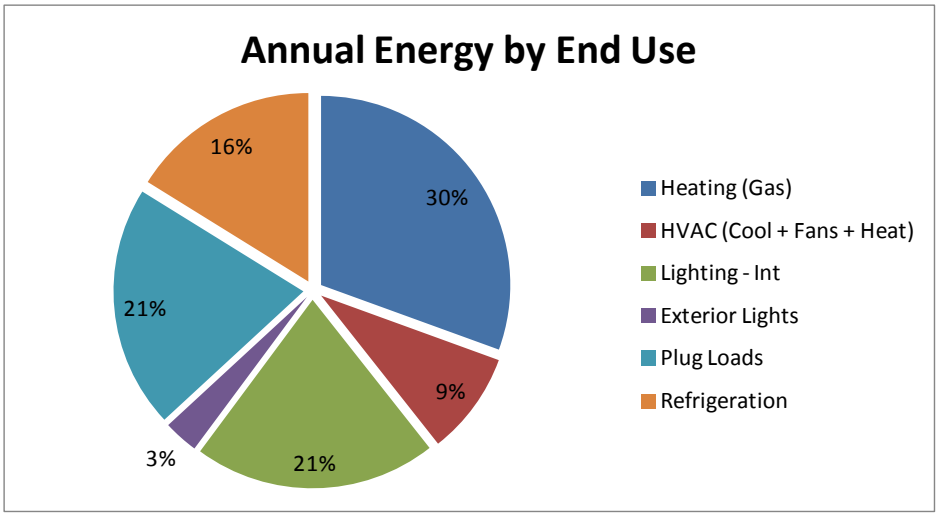


Figure 8: Annual measured energy by end-use for the period 4/1/2013 to 3/31/2014

3.3.1 HVAC and Refrigeration Equipment

Ten, single-zone, rooftop air-handling units (RTUs) with variable speed fans condition interior spaces of the store. One, multi-zone, system conditions the office and pharmacy area with VAV terminal boxes, electric reheat, and economizer mode. Each RTU contains packaged DX cooling and gas furnace heating. Three dedicated outdoor air units provide conditioned ventilation air to the store. These are equipped with energy recovery heat exchangers. The ventilation strategy is to supply a constant outdoor air quantity of 0.08 cfm/ft^2 to the entire store. The ventilation supply air temperature is reset seasonally; at 16°C in Winter, and 13°C in Summer. A dehumidification unit equipped with a desiccant wheel for evaporative condensing pre-cooling is used to keep indoor dew-point temperature below the desired control threshold of 12°C in the grocery area.

The store has 31 refrigeration display cases split between two racks of rooftop compressors. Cases operate between -26°C (frozen meats) and +2°C (dairy). Medium temperature cases do not have doors installed, while low temperature cases do. All freezer cases are equipped with automatic defrost controls which stage on in ten-minute intervals to minimize demand. All display cases with doors contain anti-sweat heaters, which modulate based on the indoor ambient dew-point temperature. There are seven walk-in refrigeration and freezer storage units in the store. They supply inventory to the food service kitchen, and stock general inventory for the grocery area.

Assumptions about infiltration had to be made for each zone of the store. This was one parameter of considerable uncertainty, as it is very difficult to measure, and can vary with a combination of sales volume, weather, and HVAC system modes. As described later on, parametric uncertainty was introduced into this variable with a multiplier sampled from a triangular distribution of +/- 40% variability. Figure 9 below shows the nominal modeled infiltration rates modeled for each zone.

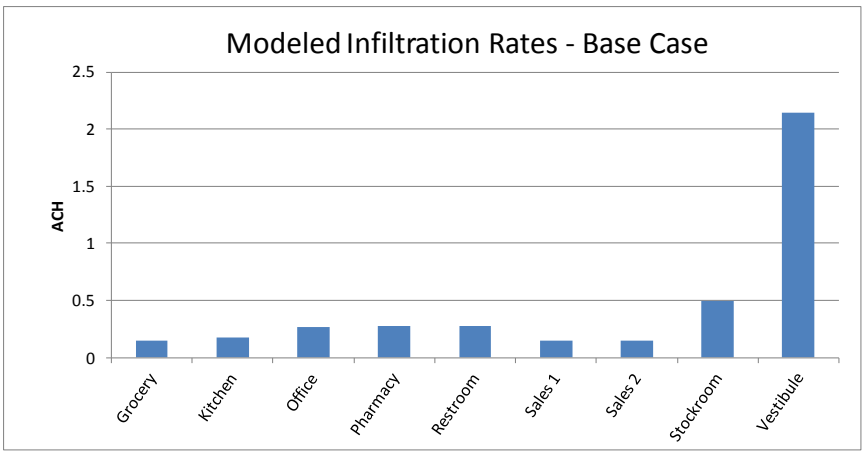


Figure 9: Modeled infiltration rates by zone

3.3.2 Schedules of Operation

As mentioned above, the retail partner has remote control of all store BMS systems from a single central engineering headquarters. All system operation schedules are set from headquarters. The store is open to the public between the hours of 8 a.m. and 10 p.m. Monday through Saturday, and 8 a.m. to 9 p.m. on Sundays. Employees are present between the hours of 6 a.m. and 11 p.m., seven days per week. Indoor temperatures are set down/up during hours when the store is closed to the public by 4°C for cooling, and 5°C for heating. There is a morning warm-up/cool down period one hour before store opening built into the thermostat schedules, shown in Figure 10 below.

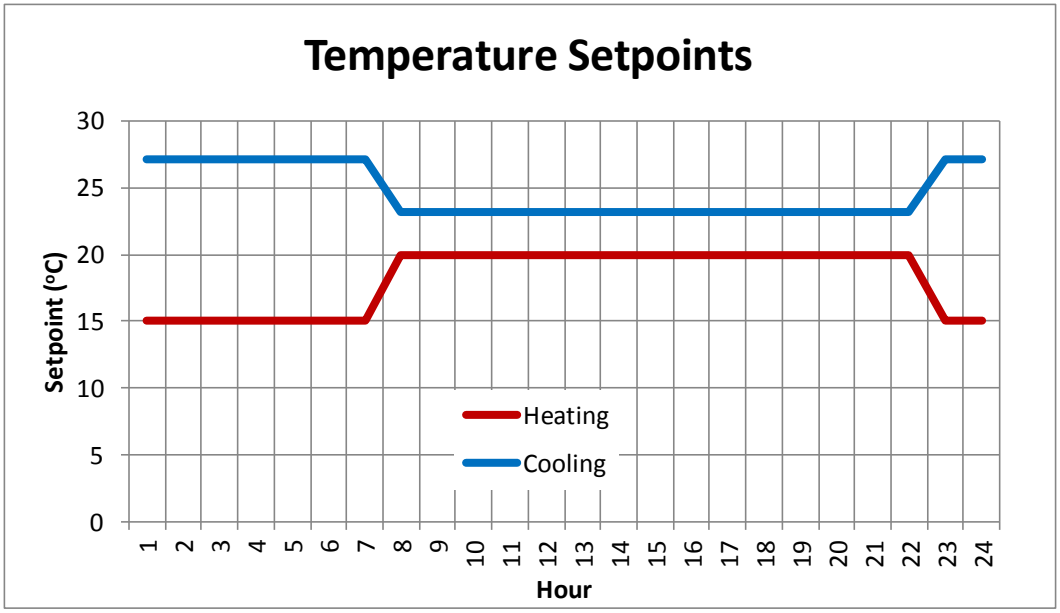


Figure 10: Indoor temperature setpoints (main sales area). Thick lines indicate a float of 0.3°C.

General interior lights are scheduled to be off during hours when no one is present, and on during all occupied hours. Refrigeration case lights are also shut off at night. Parking lot and other exterior lighting operates on an astronomical clock to be on during all hours when employees are present.

3.3.3 Lighting and Other Equipment Loads

In designing the new store, emphasis was placed on minimizing internal loads. Low-energy T-8 lighting fixtures were installed on the sales floor. Interior lighting power density is quite low, at 0.71 W/ft^2 on the sales floor. All low and medium-temperature refrigerated cases and walk-ins have LED lighting installed. Checkout stations are equipped with a stand-by mode that shut off during hours when not in use. Timers control the sales floor plug loads to be off during hours when the store is closed to the public. This includes display case lighting. Occupancy sensors shut lights off automatically in walk-in refrigerators and restrooms. In Figure 11 below, profiles of general internal occupancy, lighting, and plug-load equipment operation in the store are given. There are also cooking loads from the food service, which consume a relatively small amount of natural gas.

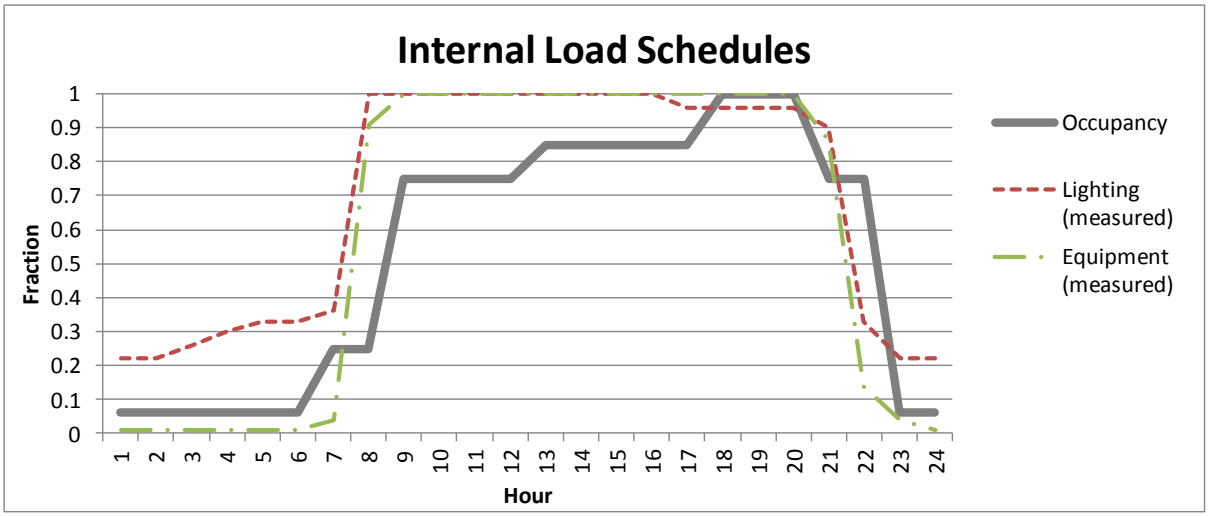


Figure 11: Schedules of internal loads (assumed and measured)

Even with the energy efficiency measures listed above, interior equipment loads are significant. Table 2 below summarizes interior equipment loads across all zones as measured in the store and modeled, including small stand-alone refrigeration cases.

Table 2: Lighting and equipment electric power density

Zone Name	Total Equipment Power (kW)	EPD (W/ft^2)	LPD (W/ft^2)
Grocery	4,050	0.26	0.68
Kitchen	23,000	12.84	1.48
Office	4,500	1.00	0.72
Pharmacy	2,629	1.35	0.91
Restroom	0	0.00	0.86
Sales 1	6,747	0.15	0.71
Sales 2	6,747	0.15	0.71
Stockroom	8,136	0.50	1.37
Vestibule	0	0.00	1.00
Walk-in Refrigeration	0	0.00	0.82

Figure 12: Interior view of case study building.

(Photo by Rois Langner, NREL.)

3.3.4 Electrical Service and Submetering Equipment

Four sub-system transformers step voltage down from a nominal 480V primary to a secondary of 208V for some sub-panels. Measured voltage at the main meter was reading between 480 and 490V before late-March, 2014, but then changed to between 460 and 470V. This was due to corrective action by the utility.

There are 88 unique energy sub-metering points at the building (not counting multiple phases of a single point). These points are metered with current transducers installed at panel, sub-panel, and equipment locations. Although there is a great quantity of energy metering, the layout was optimized for FDD and not energy management. Because of this, the process of extracting end-use consumption values involved a great deal of data wrangling. With this effort, the sub-metered energy use data did prove to be quite useful for extracting information about building operation as discussed in the next section. A full summary of sub-metered points can be found in the appendix.

4 Methods of Detailed Simulation Modeling with OpenStudio

4.1 Modeling Approach

At the outset of this project, detailed building information had already been investigated and entered into an EnergyPlus simulation model. Efforts in this work went into transitioning this model to OpenStudio (OS). The OpenStudio platform provides a sophisticated graphical user interface (GUI) to the well-developed EnergyPlus simulation engine. As discussed in the next chapter, uncertainty analysis requires large amounts of model samples. OpenStudio offers the basic user access to cloud computing batch modeling capabilities needed to execute thousands of model runs. This is through the OpenStudio Analysis Spreadsheet, which pairs a sample generation interface with a cloud-computing API (Macumber *et al.*). The open-source software platform upon which OpenStudio is written also allows the user to share and create scripts for uncertainty analysis (called "measures"; discussed later on). Compared to the reduced-order grey-box modeling approach taken by Henze *et al.* (2015), a detailed simulation model platform such as OpenStudio is more realistic to implement in the energy management industry, and is easier to modify as changes occur in the building. These reasons, and the availability of colleagues who developed the software at NREL, solidified the decision to use OpenStudio for this study.

When a simulation is run with OpenStudio, the model is processed and converted to an .idf file, which is then simulated with the compatible version of EnergyPlus. While EnergyPlus requires substantial training to use, the OpenStudio development team aims at making make OS a tool accessible to a less expert user group. It is conceivable that, within a few years, most facilities managers, energy managers, or third-party energy management consultants would be able to make small changes in an OpenStudio model that reflect

minor retrofits or operational control changes that have occurred in the building or are being tested for viability. OpenStudio was under development throughout the course of this project, with software updates released on a monthly basis.

A SketchUp plugin was used to import building geometry and numerous other features form the original EnergyPlus model to a new OS model. Figure 13 below displays building geometry as modeled. Features that were copied in included building geometry, construction materials, schedules, thermal zones, occupancy, and internal loads information. This imported .idf content was visible in the OS GUI. For things that did not get copied in completely or accurately, there were three subsequent options for inclusion:

1. Use the OpenStudio GUI to re-build components.
2. Write scripted Ruby measures that act upon the OpenStudio model directly to modify existing components when a simulation is run. These are known as "OpenStudio Measures".
3. Write Ruby measures that inject additional components, not compatible with the OSM platform, in .idf language into the .idf file that results from the initial OSM processing stage. These are known as "EnergyPlus Measures".

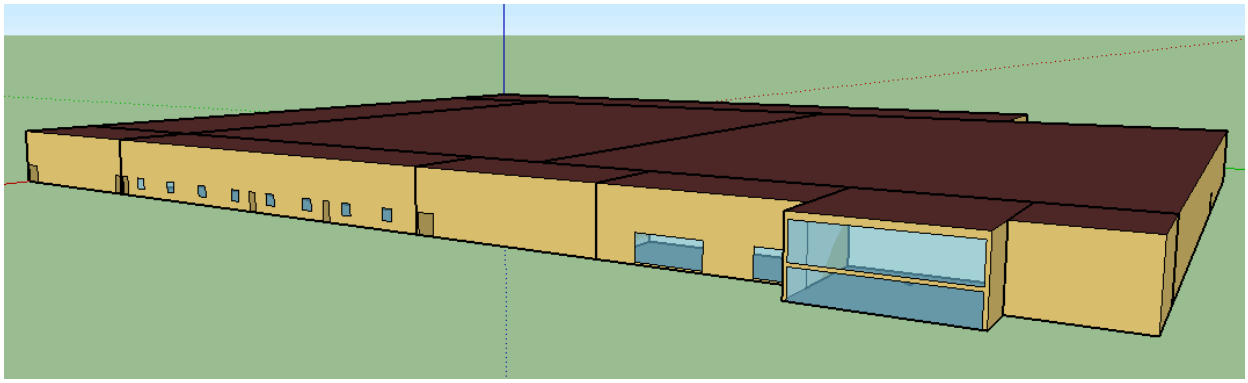


Figure 13: View of modeled case study retail building rendered with OpenStudio plugin for SketchUp

Measures are programmatic instructions (written in Ruby language) read by the OS application to make custom changes to a model. The online Building Component Library (BCL) for OpenStudio is a repository of hundreds of measures useful for creating a building

model (see Fleming *et al.*, 2012[ii]). The BCL is sometimes useful for patching missing building elements. In this work, numerous unique elements required custom measures, not available in the online BCL, to re-build. During the conversion from .idf, HVAC systems did not get copied in. In this case, most systems were packaged DX – so they were fairly simple to rebuild in the GUI. One packaged unitary system had to be imported as an EnergyPlus object. OpenStudio automatically generates individual curves that describe operation and performance of cooling equipment. A measure was applied to map the old curve data from the EnergyPlus model to the new curve names generated by OpenStudio.

Refrigeration systems did not get copied during the transition from EnergyPlus, butover either. OpenStudio was capable of modeling them with few exceptions. Several minor OS bugs in refrigeration system descriptions were encountered, which required custom measures to fix and model properly. Once again, OS automatically generates individual curves that describe operation and performance of refrigeration equipment. It is not possible to select other curves in the GUI, so these had to be assigned with another Ruby measure. Since OS generates a unique curve for each parameter of each piece of equipment, it was simple to write a measure that switched the names and handles of the .idf curves with the default curves generated by OS. In this way, the default curves were replaced with the original .idf curves.

Other model objects, created with EnergyPlus, were not supported by the OpenStudio user interface at the time of model development. These required custom EnergyPlus Ruby measures to add them to the .idf file produced in the OS output. They included:

- Outside Air zone mixing objects
- Kitchen exhaust fan
- Humidity controls
- Packaged unitary system

- Dehumidification controls using desuperheater waste heat
- Walk-in refrigeration systems
- Various Energy Management System controls

In the future, it is not expected that importing an EnergyPlus model will be so laborious, as OpenStudio content catches up.

4.2 Sampling and Batch Simulation Methods

Sampling is a process whereby a variety of model parameter combinations are tested for their impact on model outcome. Setting up the uncertainty analysis portion of the Energy Signal Tool was done through the Analysis Tool spreadsheet interface, which is under development at NREL (Ball and Long, 2014[iii]). The Analysis tool was used to set up parameter uncertainty distributions and simulate large batches of OpenStudio models. To use this tool, it is required to have measures (or arguments within a measure) written that are capable of perturbing parameter values over the range of uncertainty. A screenshot in Figure 14 below illustrates the user interface for parametric input. When combined with batch sampling, this will explore the effects of various parameter values on the building model and its end uses. Although many measures are available in the BCL, they had to be customized or new measures had to be created for this work. The Analysis Tool uses a Ruby-On-Rails automated program interface (API) to tap into AWS cloud computing servers and execute the batches of runs. The API is initiated through several command line arguments.

				Continuous Variable Description			
Measure Display Name			Nominal	Min	Max	Mean	Distribution
HVAC CURVE mapping	*RubyMeasure*						
SetDefrostEnergyRefCase	*RubyMeasure*						
SetNamesofZoneAirNodes	*RubyMeasure*						
AdjustThermostatSetpointsByDegrees	**RubyMeasure**						
variable	cooling_adjustment	Double	0	-2	2	0	triangle_uncertain
variable	heating_adjustment	Double	0	-2	2	0	triangle_uncertain
argument	alter_design_days	Bool	FALSE				
ReduceSpaceInfiltrationByPercentage	**RubyMeasure**						
argument	space_type	Choice	"Entire Building"				
variable	space_infiltration_reductior	Double	0	-40	40	0	triangle_uncertain
argument	material_and_installation_c	Double	0	0	0	0	uniform_uncertain
argument	om_cost	Double	0	0	0	0	uniform_uncertain
argument	om_frequency	Integer	1	1	1	1	uniform_uncertain
SetHeatingandCoolingSizingFactors	**RubyMeasure**						
variable	htg_sz_factor	Double	1.2	1.1	1.4	1.2	uniform_uncertain
argument	clg_sz_factor	Double	1.2	1.1	1.4	1.2	uniform_uncertain
AddPckgdUnitarySystem	**EnergyPlusMeasure**						
ModifyEnergyPlusEMSObjects	**EnergyPlusMeasure**						
argument	zone_mix	Double	2	0.5	3.5	2	uniform_uncertain
argument	cycle_time	Double	1800	1200	2400	1800	triangle_uncertain
AddEnergyPlusObjects	**EnergyPlusMeasure**						
argument	econ_low	Double	10	8	12	10	uniform_uncertain
argument	econ_high	Double	18.33	16.33	20.33	18.33	uniform_uncertain
variable	float	Double	0.278	0.2	0.8	0.278	triangle_uncertain
argument	fan_low_spd	Double	0.4	0.3	0.6	0.4	uniform_uncertain
argument	fan_med_spd	Double	0.8	0.7	0.9	0.8	uniform_uncertain
FanLoopsEnergyPlusmodifyFanEffFan	**EnergyPlusMeasure**						
variable	pressureRise_mult	Double	1	0.7	1.3	1	uniform_uncertain
variable	fanEfficiency	Double	0.45	0.35	0.48	0.45	triangle_uncertain
argument	fanMotorEfficiency	Double	0.9	0.75	0.93	0.9	uniform_uncertain

Figure 14: Screenshot of Analysis Spreadsheet parametric input

The Analysis Spreadsheet has an MS Excel interface for inputs regarding cloud computing configuration, sampling algorithm, parameter uncertainty distribution and output reporting. Figure 15 below provides an example of spreadsheet configuration input. It has numerous sampling and model optimization algorithms built in. Those used for this work included "pre-flight" (for OAT parameter pre-screening) and "LHS" (for Latin hypercube sampling, which was used extensively throughout). On the "Setup" tab of the spreadsheet, the user may select the option of sampling "all variables" or "individual variables". For the pre-flight algorithm, the option of sampling individual variables was utilized to isolate the effect of each variable on model outcome. For the LHS algorithm, the option of sampling all variables simultaneously was selected. This means that each sample generated was globally unique in that it is an assemblage of unique random values within the uncertainty range of each parameter. It is expected that a more advanced version of the Analysis Tool will be incorporated into the Parametric Analysis Tool in the near future as part of the standard OpenStudio GUI package.

Settings				
Spreadsheet Version	0.3.6			
User Id	eric_target			
OpenStudio Server Version	1.8.0			
Cluster Name	Default Cluster			
Server Instance Type	c3.4xlarge	Worker Only - Recommended for Worker	16 Cores with 160 GB	$0.84/hour
Worker Instance Type	c3.8xlarge	Worker Only - Recommended for Worker	32 Cores with 320 GB	$1.68/hour
Worker Nodes	1		Total Cost	$2.52/hour
Running Setup				
Analysis Name	target_lhs_month4_1000			
Measure Directory	../measures			
Export Directory	../analysis			
Allow Multiple Jobs	TRUE			
Use Server As Worker	TRUE		22	
Simulate Data Point Filename	simulate_data_point.rb			
Run Data Point Filename	run_openstudio_workflow_monthly.rb			#NAME?
Problem Definition	**Problem Type**			
Analysis Type	lhs			
Algorithm Setup	**Values**	**Allowed Values and Description**	**(Enter values here to**	
Sample Method	all_variables	individual_variables / all_variables	all_variables	
Number of Samples	1000	simulations is this times each variable)		

Figure 15: Screenshot of Analysis Spreadsheet "setup" tab

When a batch of simulations is initiated, the spreadsheet sequences them to run on the cloud-based web server with OpenStudio server code. The Ruby "gem" to install and run the package is publically available online[5].

The model samples generated by the spreadsheet are sent out for execution to the Amazon Web Service Elastic Computing Cloud (AWS-EC2) service. The EC2 offers a virtual computing environment for large batches of EnergyPlus simulations to be processed in a much shorter amount of time than would be required at a desktop computer. Currently, normal users can call up to 64 cores of server power, which collectively decreases the amount of time required to run simulation batches by a factor of 50. The simulation is set up by selecting the server and worker instance types to be run with EC2. There are eight classes of servers and four classes of workers to choose from, each with a different level of processing power, memory storage, and EC2 pricing. Cloud server memory storage is just as important as cloud server count for effective batch simulation. During this work, it was

[5] https://rubygems.org/gems/OpenStudio-analysis

found that server memory limits and large SQL results output files are limiting factors to total batch size.

There are several visualizations of data generated with batch simulation. These compare data points of parameter input to objective function output help the user to see the strength of correlation and possibly identify parameter multi-colinearities. Some examples of OS Server visualization are included in the appendix.

4.3 Weather Data

Independent of other factors, year-to-year fluctuations in weather can have a significant impact on energy consumption. Using an Actual Meteorological Year (AMY) weather file, rather than a Typical Meteorological Year (TMY) file, can reduce the amount of uncertainty in benchmark expectations. Wang *et al*. (2012[iv]) found that by taking historical data from 10-15 year timeframes in four climates, there was up to an 8% spread in building energy use. This amount of uncertainty is greater than risk tolerance thresholds used to generate energy signals. As the figures below illustrate, there were significant differences between a typical year and the actual year at the case study project site. In a real application of the Energy Signal Tool (as was planned at the outset of this project), the user would need updated AMY weather files for each application site. There are tools and methods available for creating such files from freely available data, but the process is quite complicated. Hourly observed solar radiation data is rarely available. Installing a weather station that can log hourly data for each Energy Signal Tool implementation site in a large portfolio would provide the best access to weather data.

For this project, weather data for the nearest official weather station to the case study site, which is logged at the Timmermann Airport (station KMWC). An (AMY) of data was obtained from Weather Analytics for a fee. This came in directly as an EnergyPlus (.epw) formatted file, spanning the 12-month period from April 1, 2013 to March 31, 2014.

This is the period corresponding to the remainder of the work presented with the case study building.

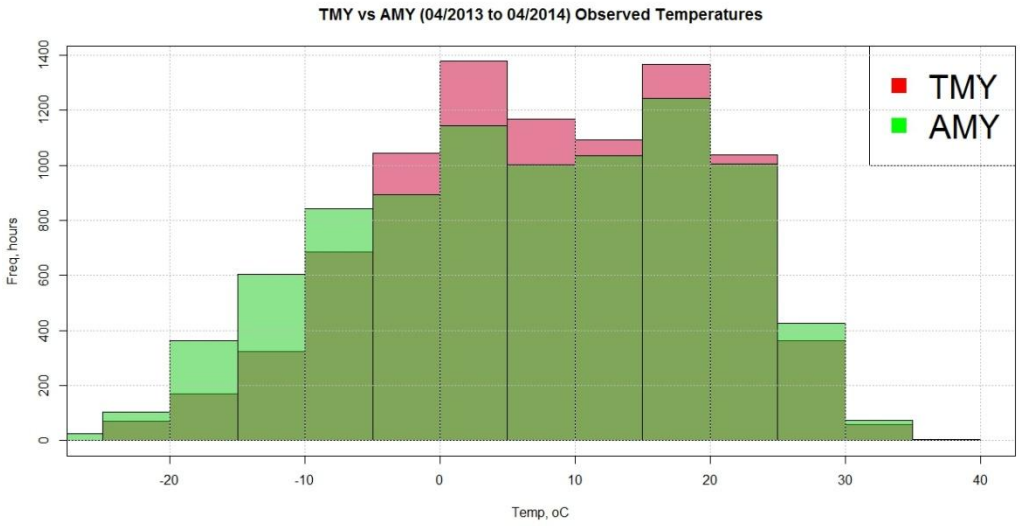


Figure 16: Comparison of AMY vs. TMY hourly outdoor dry-bulb temperatures

As seen in Figure 17 below, the AMY period contained a significantly greater number of heating degree days than a TMY. The winter of 2013-2014 was a cold one in the upper Midwest region.

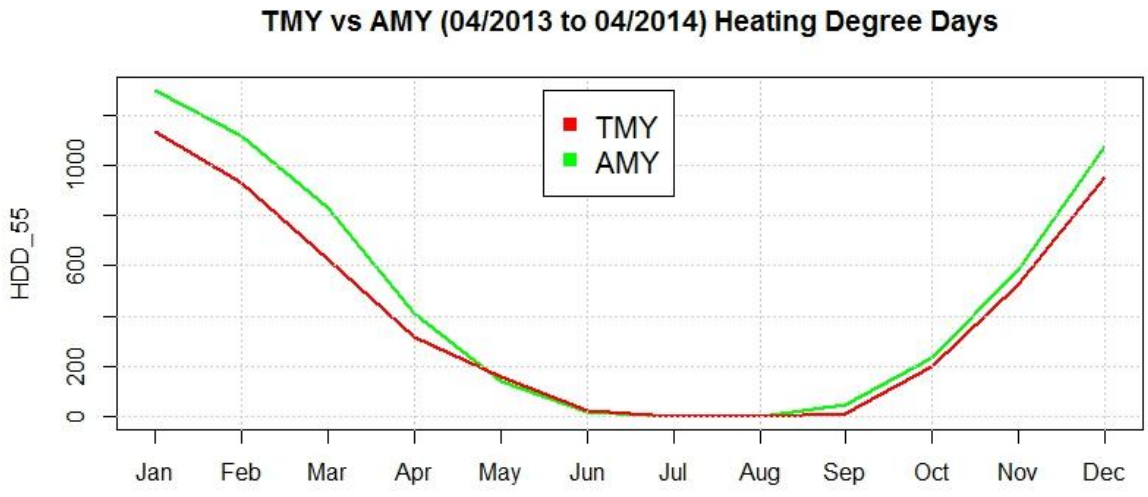


Figure 17: Comparison of AMY vs. TMY heating degree days per month

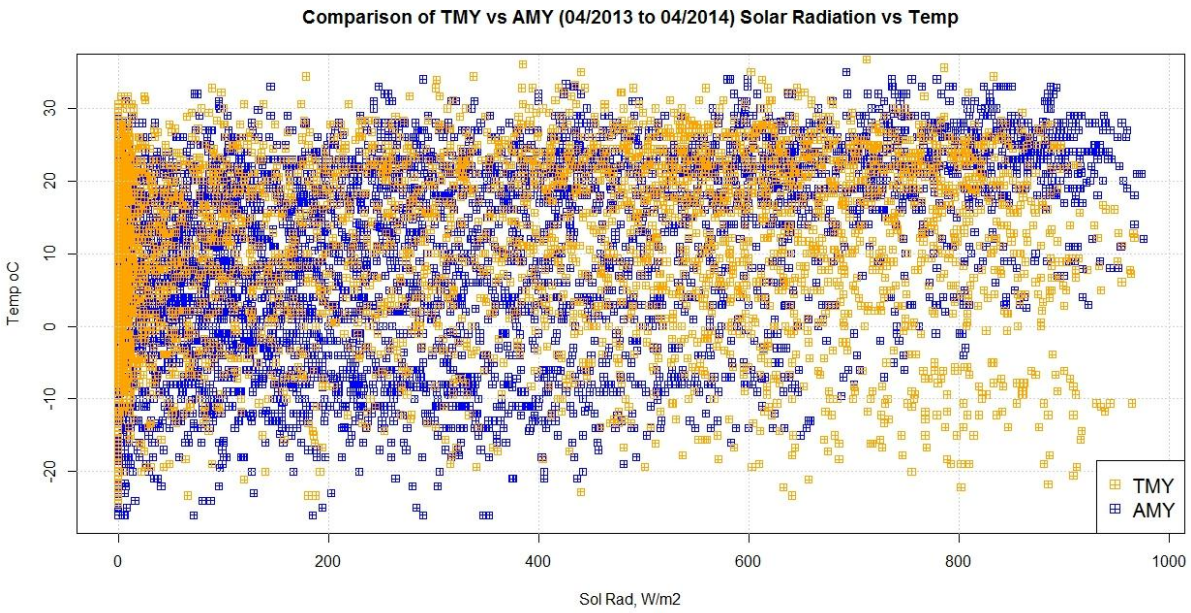


Figure 18: TMY vs. AMY solar radiation and temperature

The figure above illustrates how TMY expectations are for sunnier conditions during cold weather than were observed during the AMY period. From the figure below, it seems that times of low solar radiation are recorded in a slightly different manner for AMY data sets.

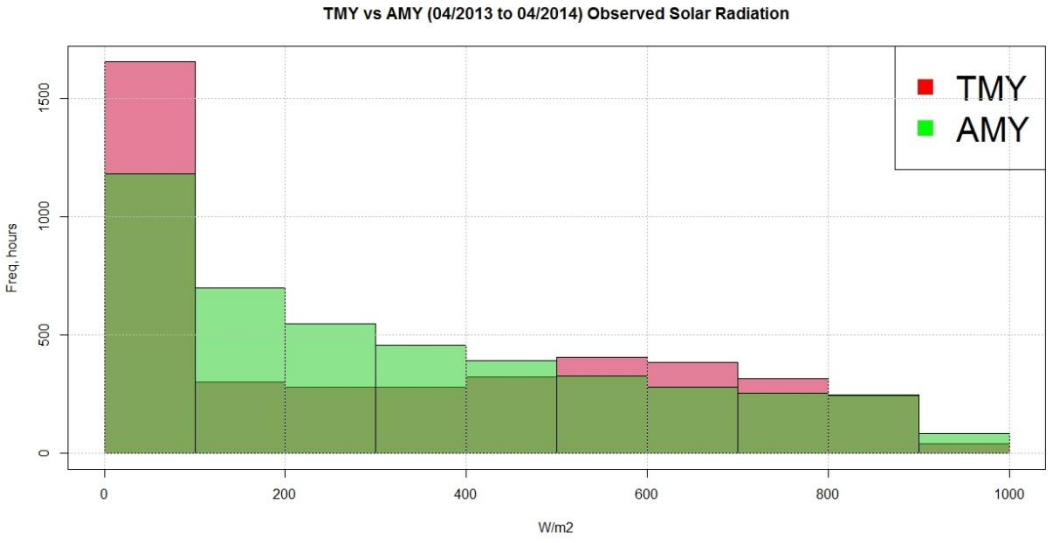


Figure 19: Comparison of AMY vs. TMY hourly solar radiation data

4.4 Use of Sub-Metered Energy Data

The benefits of having sub-metering equipment installed are in giving facilities managers a quantitative picture of how the building is performing, encouraging performance accountability, and providing a framework for tracking behavioral change (NIST, 2011). The case-study retail store is equipped with 88 points of sub-metering consisting of hardware measuring data from main panels distribution panels, sub-panels and down to the level of specific equipment, and even equipment components in some instances (VAV boxes, RTU compressors, RTU fans, Etc.). By installing a larger number of monitoring points in the case study store, the retail partner hoped to learn from their investment in the high-performance design. The greater than normal extent of this sub-metering infrastructure made it possible to sum specific sub-metered points into category by end-use without much difficulty. Refer to the appendix for a full list of sub-metered data points and groupings.

4.4.1 Sub-Metered Data Analysis and Preliminary Heuristic Calibration

Sub-metered data from the case-study building were available through an online energy management application from as early as October of 2013 up to the present. There was a major retro commissioning effort of the data collection system between September and October of 2013. During this period, no data was collected, and several changes were made that effectively rendered previous data of poor comparison. Due to the scope of this work, the heuristic model calibration needed to take place before the entire years worth of sub-metered data was available. Therefore, data for the period October 2013 through April 2014 was used to extract information pertaining to schedules of operation. The figure below shows that there is relatively little variance in the load profile for each hour. For the most part, operation is predictable. The extreme outliers represent days when the store was open for 24 hours, or was closed for a holiday. In this and other “box-plot” style visualizations, the blue shaded range represents the 25th to 75th percentile (interquartile)

range of observed data. The median of the data is represented by a black line within the shaded region. The "whiskers" bound the range encompassing 1.5x the interquartile range, and the outlying points are all those beyond. Boxplots with narrow blue regions represent distributions with less variation. Those with median values close to the geometric center of the shaded region have more normal distribution characteristic.

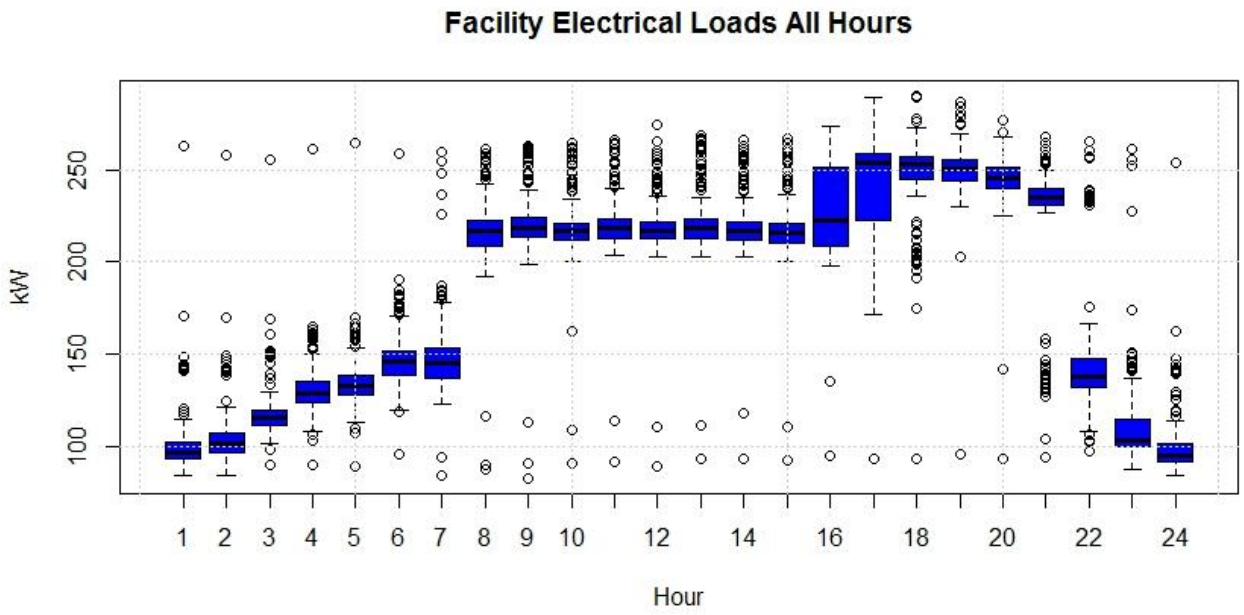


Figure 20: Boxplot of electrical load profile as observed from 10/2013 to 4/2014 (170 days)

The hours of 4 p.m. and 5 p.m. exhibit significantly more variation than most. It seems that this is partially because occupancy increases more in these hours during the holiday shopping period than others. Also, the shift from more occupancy at the 4 p.m. hour to the 5 p.m. hour in the early spring is indicative of parking lot lights that operate on an astronomical clock. In Figure 21 below, building electricity load data for the hours of 4 p.m.

and 5 p.m. is given from the period 10/1/2013 to 4/6/2014.

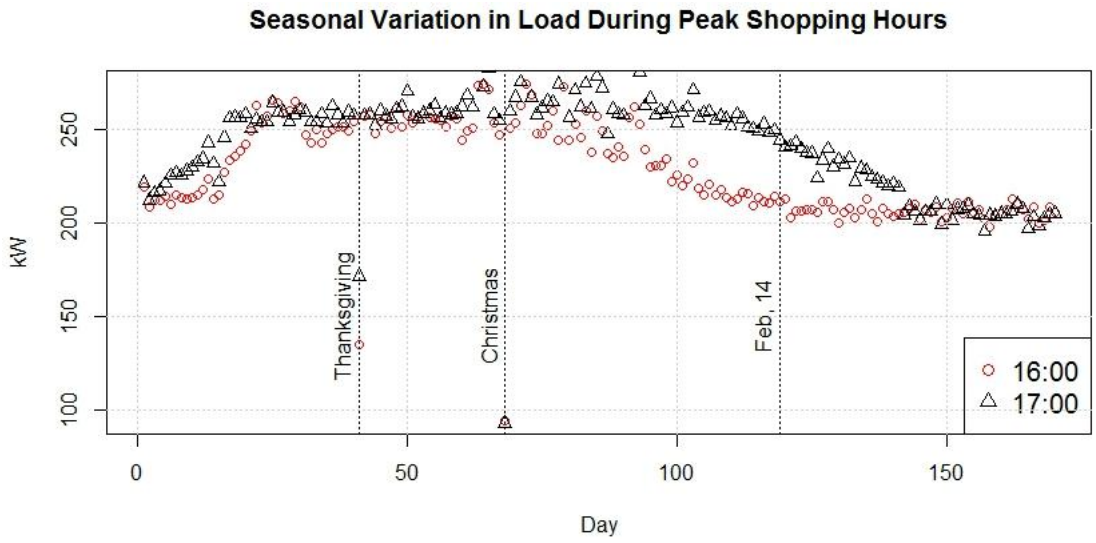


Figure 21: Variation in total building electricity load

Even after all physical information about the building is gathered and entered into the energy model, there is additional work in the heuristic calibration process. Processing sub-metered data into energy model load schedules can enhance the heuristic model calibration. The EnergyPlus feature '*Schedule:File'* can be used to process sub-metered data into very accurate schedules for end-use implementation[v]. However, to reduce model complexity, the useful sub-metered data was analyzed and converted to typical day schedules. This was deemed appropriate for a facility with identical operation hours seven days per week. Of the sub-metered end uses, the following were identified as candidates for this method of schedule building. They are all cyclical on a diurnal basis, and have no variation due to interactions with environmental factors, or other systems in the building.

- Stock/Receiving Lights
- Office Area Lights
- Sales Floor Lights
- Parking Lights

- Sales Floor Plug Loads*[6]
- Food Service Equipment Loads*[7]
- Electric Water Heating Load*[8]

For those items listed without an asterisk, the data presented itself as following a very predictable daily pattern, where the daily maximum was within close range of the maximum taken from the entire 170-day (4080 hour) data set. These end-uses exhibited little hourly variation. For these end-uses, the following formula was used to generate schedules:

Equation 1: Hourly schedule derivation from sub-metered data

$$Hourly\ Values\ (\%) = \frac{\sum_{d=1}^{170} \sum_{h=1+(d-1)*24}^{24+(d-1)*24} (Hourly\ Value)\ /170}{Max(\sum_{d=1}^{170} \sum_{h=1+(d-1)*24}^{24+(d-1)*24} (Hourly\ Value)/170)}$$

Schedules for an energy model are built by normalizing the load at each hour to some maximum load, which is assigned a value of 100%. In the equation above, the mean load value from the observed data set is taken for each hour (170 samples of each hour) and divided by the maximum of those means.

The sales floor plug loads exhibited significant seasonal variation, as shown in the following figure:

[6] Plug loads did exhibit some seasonal variation due to seasonal displays.

[7] Kitchen equipment loads exhibited significant, random, daily variation of hourly load profile.

[8] Water heating loads exhibited significant, random, daily variation of hourly load profile.

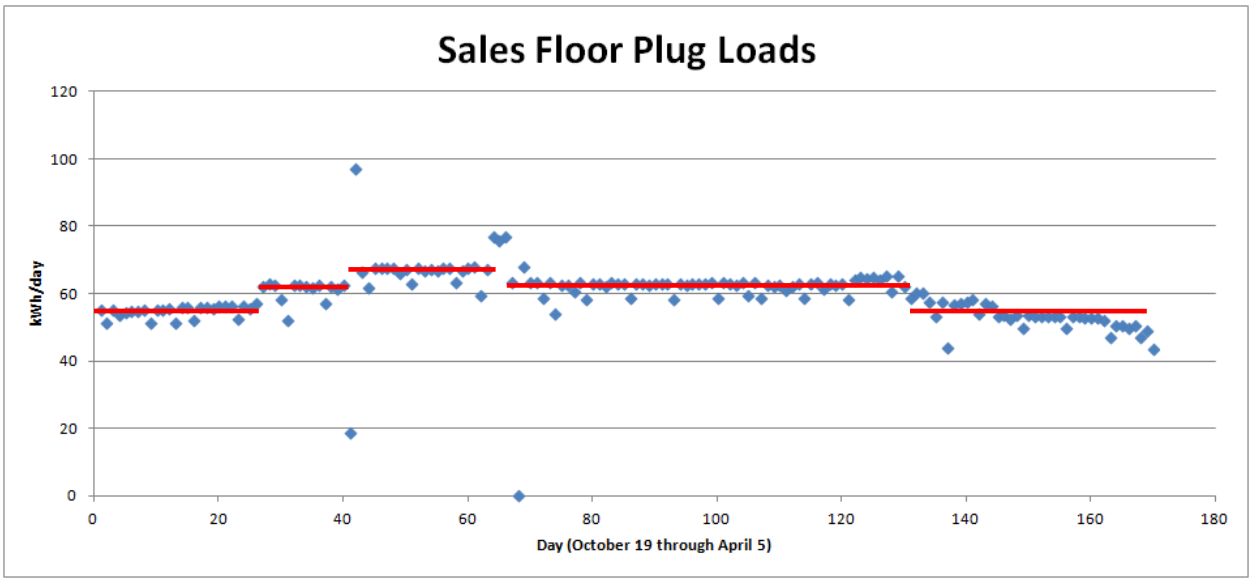


Figure 22: Seasonal variation in retail store plug loads

Sales floor plug loads are elevated for the winter holiday sales season, and are then reduced to a base value between March 1st and November 15th. The differences between holiday and non-holiday periods in the daily load frequency distribution are seen below in Figure 23. A difference of 5% is observed in the elevated median of the holiday period. Due to this noticeable variation, the sales floor plug load schedule was created by normalizing to the peak hour in the holiday season, adjusting all hours up or down accordingly.

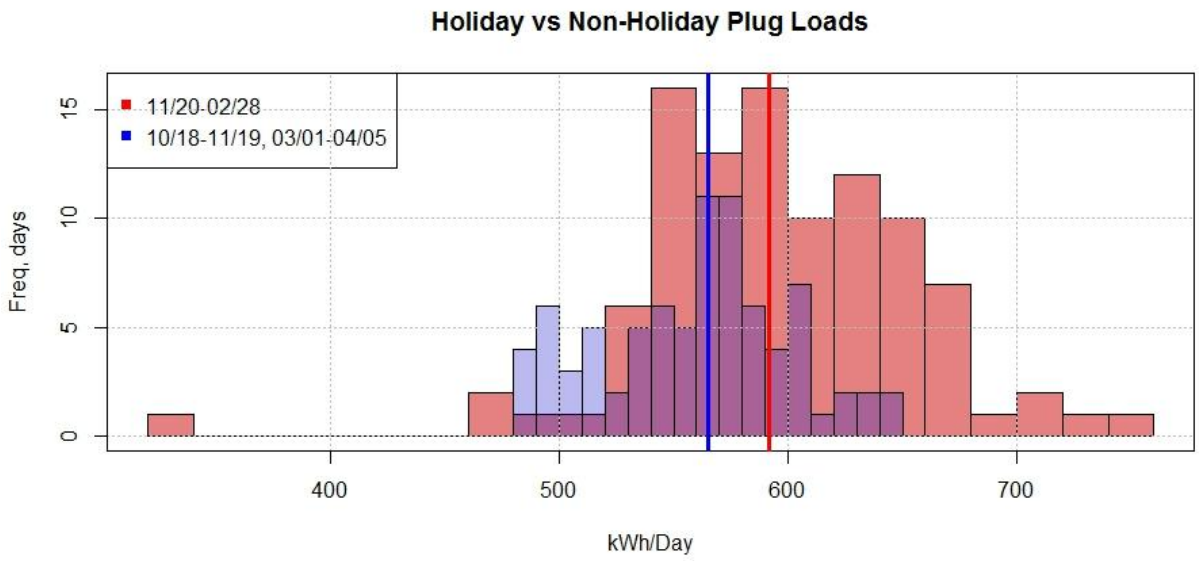


Figure 23: Holiday vs. non-holiday plug load measurements taken between October 18, 2013 and April 5, 2014

The food service plug loads and the water heating load exhibited enough *hourly* variation across the 170-day sampling period that simply taking the mean value at each hour would have been accurate from a total energy use standpoint, but would have neglected a significant portion of the peak load contribution. The water heater loads are distributed as follows, with relatively wide hourly variations across the 170-day period:

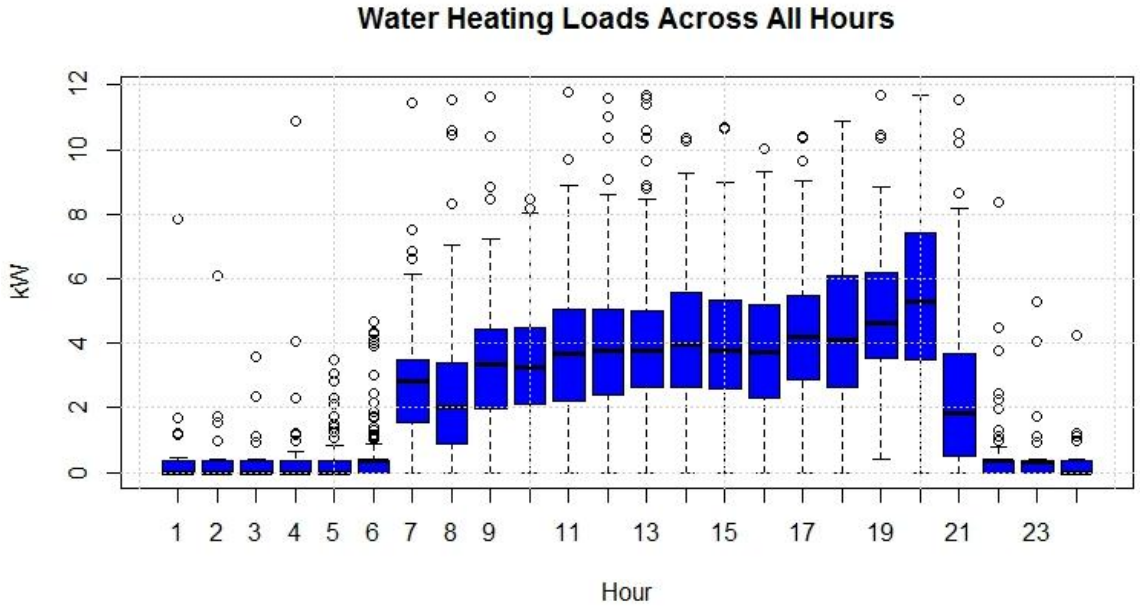


Figure 24: Domestic hot water loads, recorded hourly over the 170-day monitoring period

Whereby the most likely hour for the daily peak domestic hot water load to occur is 20:00 to 21:00, as verified by the figure above. The 0.995 quantile load for this hour was 10.6 kW. Thus, this hour was assigned the peak load of 10.6 kW, and the other hours' mean values were adjusted down by an equal share of the peak offset to create the typical daily profile schedule. The figure below shows the "average" load (average of all hours) compared to the "typical" load as modeled, including the typical daily peak.

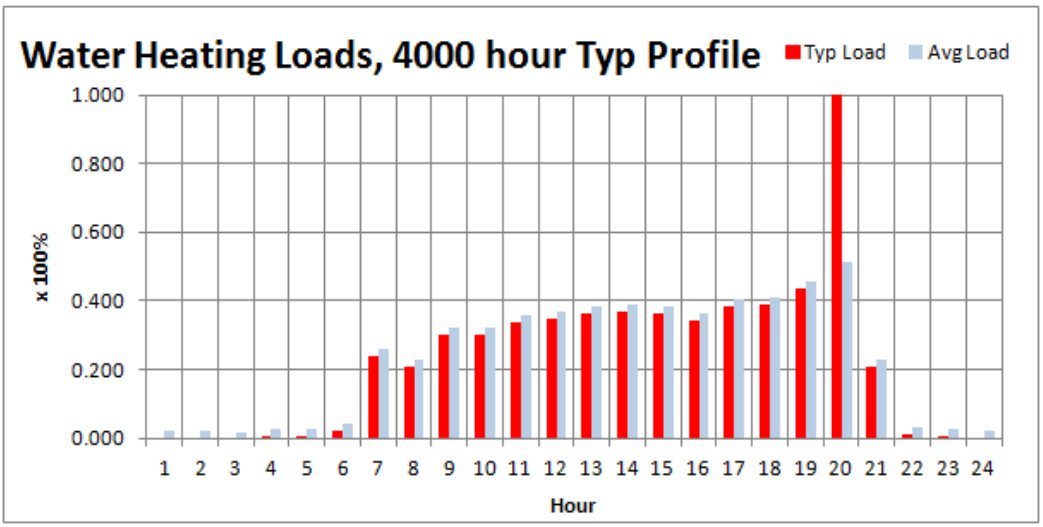


Figure 25: Final scheduled water heating loads compared to averages

The kitchen equipment loads, which are comprised of food service equipment and various other plug loads, are distributed as follows in Figure 26:

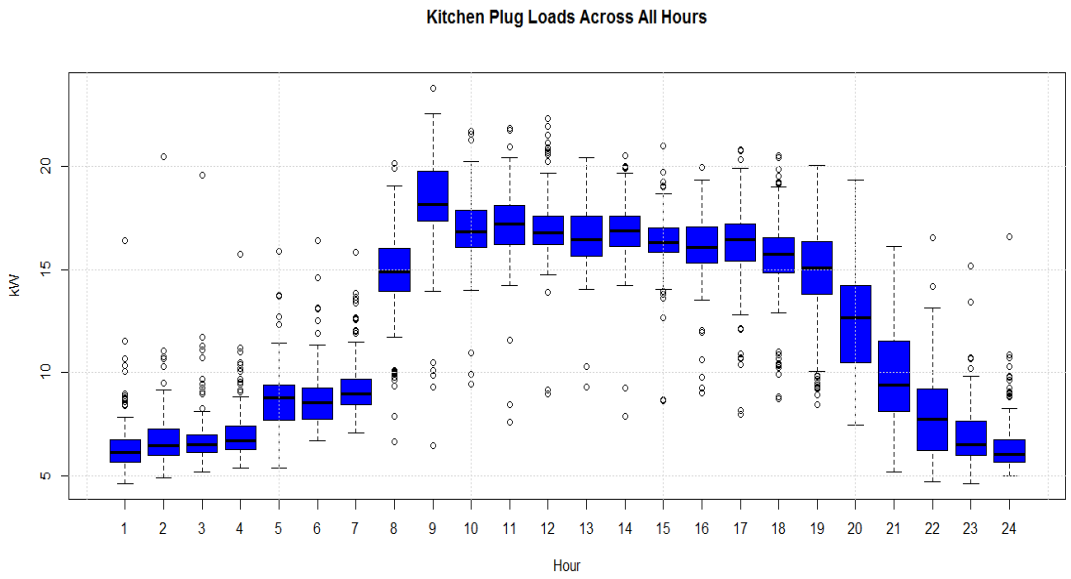


Figure 26: Boxplots of kitchen plug loads from 170 days of hourly data

From Figure 26, it can be concluded that the most likely peak occurs in the 09:00 to 10:00 hour; likely when food preparation is underway. The 0.995 quantile of the distribution for this hour was 22.4 kW. Thus, this hour was assigned the peak load of 22.4 kW, and the other hours' mean values were adjusted down by an equal share of the peak offset to create the typical daily profile schedule.

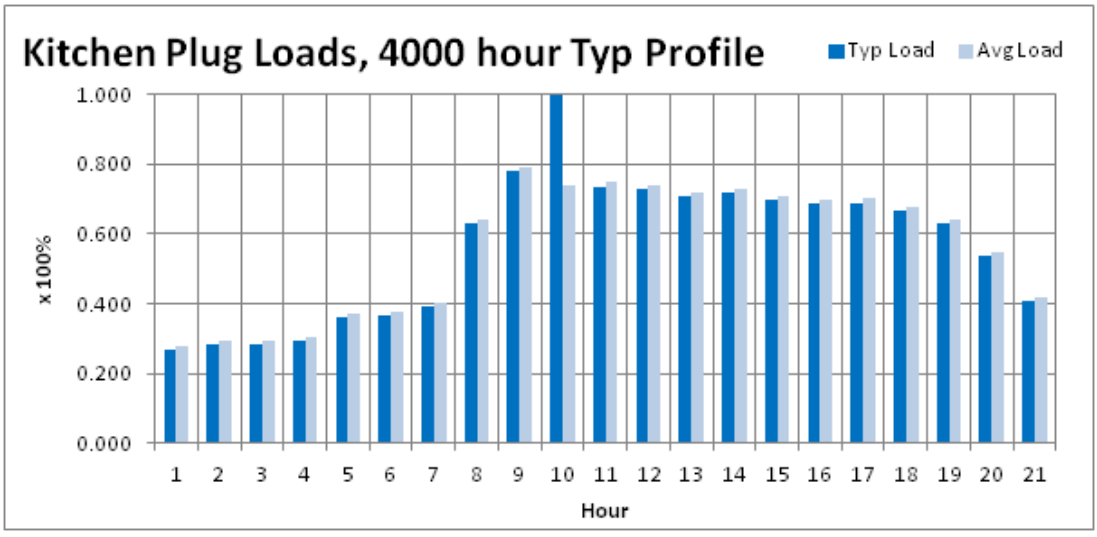


Figure 27: Final scheduled kitchen equipment loads compared to averages

The figure below shows a summary of daily electricity use distributions for the major loads for which information was extracted from the sub-metering data from the 170-day data collection period. With the exception of the sales floor plug loads, all other loads examined show a very normal distribution of daily load; with the box plots being centered about the mean. Thus, it was valid to match the sum of daily load schedules equal to the daily mean in the model; except for the sales floor plug loads as described above.

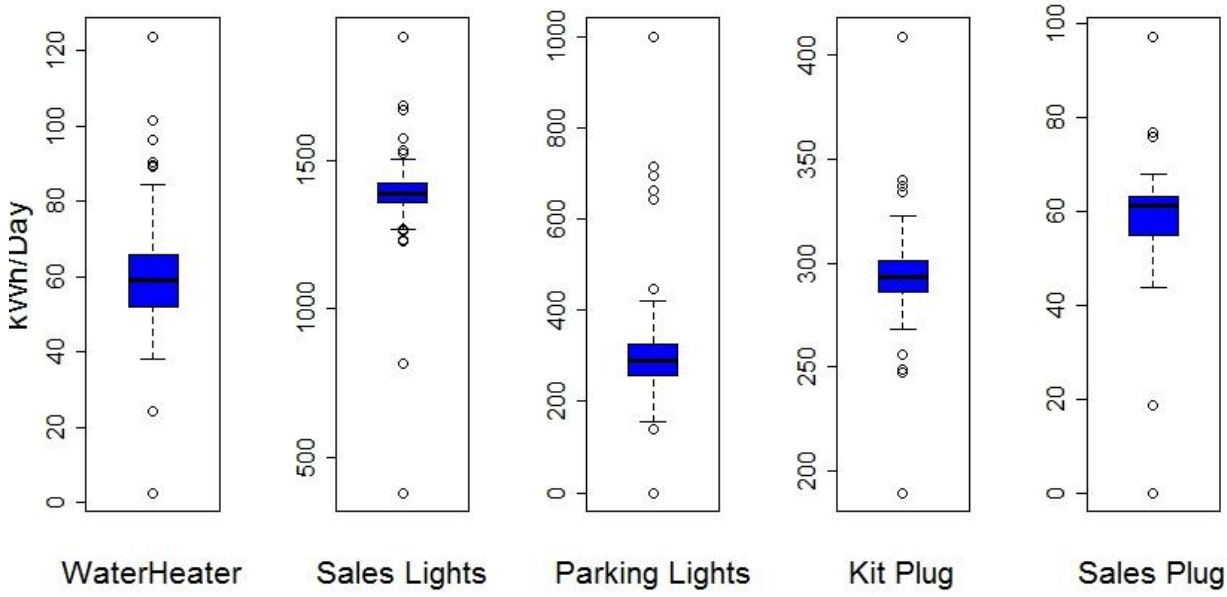


Figure 28: Distributions of daily loads from sub-metered end uses

As discussed later on, it was important to note the normalcy of the load distributions for lighting and plug loads. This is illustrated in several figures in the appendix. Due to modeling limitations, it was not possible to include schedule uncertainty in the uncertainty analysis.

4.4.2 Uncertainty in Sub-Metered Data

Sub-metered data observations form the basis on which the Energy Signal Tool assists with decision making. As such, validation of data quality is paramount. Error in sub-metered energy data can come from a variety of sources. This error can be grouped into four categories, and the extent of uncertainty will vary with the level of quality in the sub-metering infrastructure.

Physical implementation uncertainty:	• *Is every piece of equipment included in the sub-metering infrastructure?* • *Are some sub-metering devices installed improperly?* • *Is there phase in-balance in the electrical distribution system that is not accounted for with sub-metering of only one of three phases?*
Front-end calculation uncertainty	• *Are conversions from current transducer to units of power being done properly?* • *Is the building power factor being accounted for in real time?* • *Was the right voltage accounted for, or is it being accounted for in real time?*
Back-end calculation uncertainty	• *Are sub-meter aggregation assumptions accurate?* • *Is there duplication in sub-meter aggregations?*
Building electrical system interference uncertainty	• *Has a change in building voltage or power factor impacted the front-end calculations?* • *Has a change in wiring or transformer infrastructure impacted the losses between the main meter and the sub-meters?* • *Has a change in wiring impacted phase balance?* • *Was power factor correction equipment installed?*

For the case-study building, it was possible to deduce end-use data for the four major categories of HVAC, Lighting, Plug Loads and Refrigeration. This was done by taking combinations of sums from among the 88 sub-metered points available on the web-based energy management system. Also available were separate measurements of energy use taken at the building electrical service entrance. Comparing the sum of end-uses to the readings gathered from the main meter was done to check the assumption of accuracy in sub-metering equipment. Initially, it was expected that the sum of end-uses would be slightly less than the main metered amount due to secondary transformer losses and line losses in the building electrical distribution system. However, as evidenced in Figure 29 below, the sum of end-uses was found to be predominantly greater than the main meter reading.

$$Sub - meter\ Error = \sum End\ Uses - (Main\ Meter\ Reading)$$

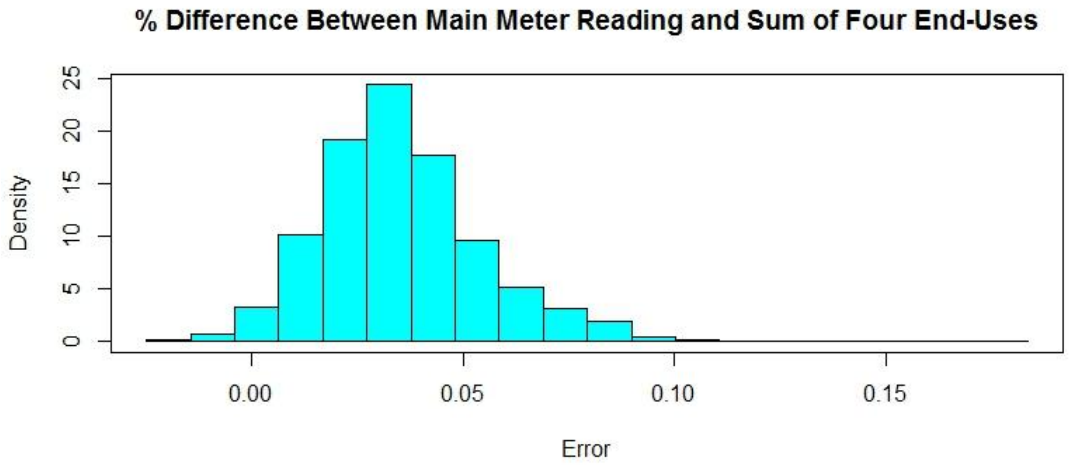


Figure 29: The normal biased error distribution of sum of end use meters vs. main meter

Error is assumed to be in reference to the main meter, which is a metered point located at the building electrical service entry. As seen in Figure 29 above, percent error is normally distributed around a mean value of +3.9%. This means that readings taken from

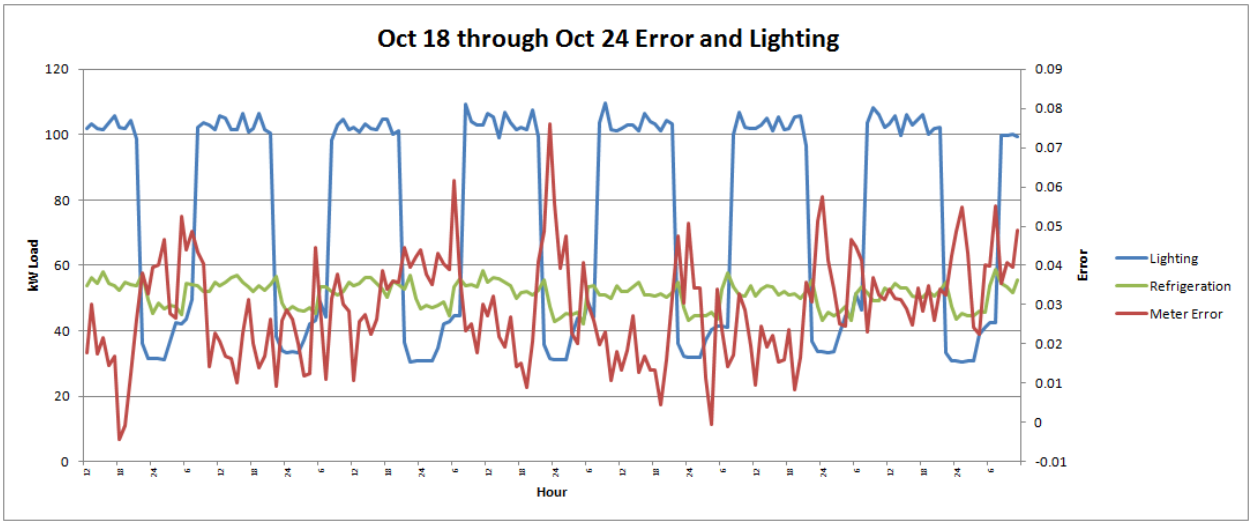
summing end-use meters are higher than the reading taken at the main meter by this amount.

A small data correlation study was done in an attempt to identify the root cause of this error in one of the end-uses. To do this, the time series data set of hourly measurement error was compared to the concurrent magnitude of load in each end use. An assessment was made for the Pearson correlation coefficient of each pair. It was expected that as the load carrying the root cause of the error would have the highest correlation to the measurement error.

Table 3: Correlation between end-use time series data and measurement error

End-Use	ρ
HVAC	0.424
Plug Load	0.342
Refrigeration Load	0.301
Lighting Load	0.443

Correlation to measurement error was found to be strongest for the lighting end use. The discussion below provides further evidence found to validate this.



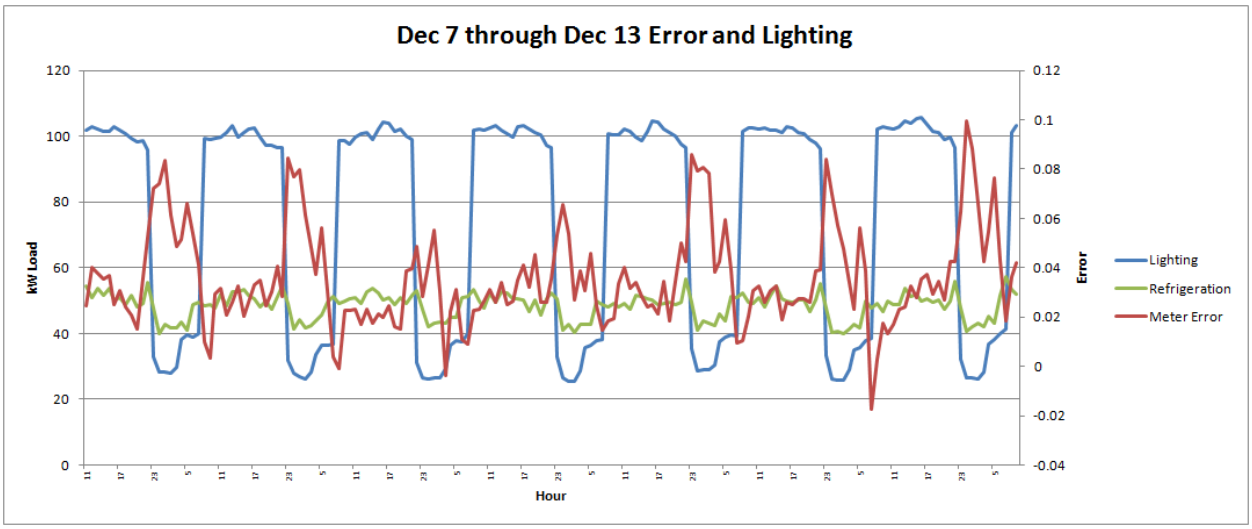


Figure 30 a,b: Visual correlation between lighting energy load and sub-meter error

As illustrated in Figure 30 above, measurement error is always highest at night when most lighting and plug loads are off. Dominant loads at night are refrigeration, which increases refrigeration as a percentage of total end uses.

As mentioned earlier, there had been a 4% service voltage drop at the building. An assumption of service voltage is important to the front-end calculations of sub-metered energy if it is not measured continuously. This service voltage step change did produce any noticeable change in the error observed between the main meter and sum of end-uses, as illustrated in Figure 31 below. Thus, the difference in voltage reading versus assumed voltage is not a factor contributing to error in the measurement.

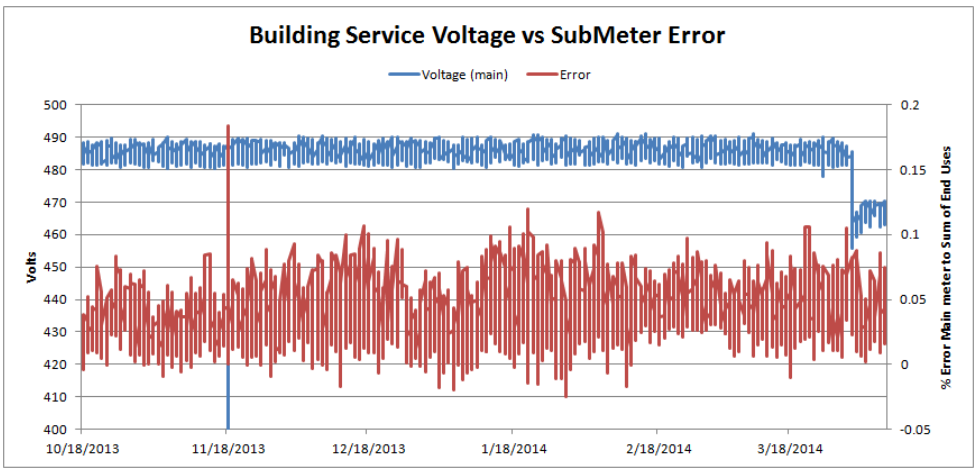


Figure 31: Building service voltage and sub-meter measurement error

After these investigations, a conversation with a facilities engineer working for the retail partner revealed the fact that all current transducer devices measuring end-use energy consumption are installed only on one out of three phases (Phase "A"). The exception to this is the interior lighting system, which has current transducers on all three phases. Thus, the positive error of sub-meter readings is a reflection of the phase imbalance on the electrical sub-panels being biased toward more loading on Phase A, on average, across all panels. The interior lighting system has a load of 80 kW, which is approximately 33% of the total daytime running load of the building. When the interior lighting system is turned on, it dilutes this bias. When the interior lighting system is off, all three-phase sub-metered loads are calculated by assuming phase *balance* – which is not accurate. Thus, this measurement error was found to be due to phase imbalance of the electrical systems being measured; placing it in the category of physical installation.

Finally, according to facilities engineers, the measurement error tolerance for the type of current transducers installed in the case study building is +/- 5%. Thus, the mean measured error, although biased for the possible reasons listed above, is within this tolerance. There are higher quality sub-metering systems that are rated for measurement error between 0.05 and 3%, depending on cost and application (NIST, 2011).

In addition to measurement error caused by problems with the physical installation and phase imbalance, it is common to have gaps or unexplained shifts in sub-metering data. Often, it is possible to fill minor gaps with simple regression techniques. The root cause of pattern shifts can be ascertained most often by speaking with a facilities manager. As explained later on, some quite large gaps were identified in the data set. In addition to this, the sub-metered data seemed to suggest drastic shifts in operation of dehumidification equipment. After having utilized the "good" data as much as possible, and then losing contact with the case study sponsor, the author decided it was outside the scope of this work to attempt to reconcile sub-metered data for model calibration purposes. This also led to the decision to make primary testing of the Energy Signal Tool based on synthetic end-use data. Refer to subsequent sections for additional explanation of data validation tests.

5 Parameter Uncertainty and Sensitivity Analysis

Assessing the entire range of possible model outcomes due to the full range of parameter uncertainty is a computationally intensive task. Uncertainty analysis quantifies the uncertainty in model outcome due to the uncertainty that exists in the set of model input parameters (Saltelli *et al.*, 2000). A sensitivity analysis can be used to reduce the subset of input parameters to that which have the biggest impact on energy use without impacting the overall magnitude of uncertainty propagation.

5.1.1 Parameter Uncertainty Analysis

Uncertainty is important to incorporate in the model because without it, the model would be simply a static representation of many point estimates, producing another single point estimate of the objective functions. By considering uncertainty in input parameters, the model is capable of providing a probabilistic range of the objective function. This makes

the model more useful to a decision maker concerned with building performance assessment.

There are five main sources of uncertainty in energy simulation (de Wit, 2003[vi])

1. Specification uncertainty (whereby the building may be constructed differently from design specifications)
2. Model uncertainty (the model is just a simplification of a real process; minimize with calibration)
3. Model input uncertainty (some specifications are unknown; address with sensitivity analysis)
4. Numerical uncertainty (model convergences on things such as system sizing may vary with iterations)
5. Scenario uncertainty (user/occupant behavior, external factors such as weather)

Sources (1), (3), and (5) above are addressed in this work with sensitivity analysis and will continue to propagate into the model results.

Wang *et al.* (2012) studied the possible ranges of eight significant model/scenario parameters for their contribution to model uncertainty for a typical office building. They observed that the combined uncertainties produced an 85% spread in annual energy consumption in climate zone 5a. When taken individually, the lighting and plug load schedule parameters contributed the most uncertainty to the models. In combination, HVAC parameters contributed far greater uncertainty. This is due to the highly interactive nature of HVAC uncertainties in the systems comprising a building; which is revealed in Monte Carlo sample-based whole building energy analysis. There is thus a strong case for using detailed building energy analysis in the Energy Signal Tool. This can quantify operational uncertainty while accounting for the interactive effects of uncertain parameters. Other available tools, such as ECAM, are quite capable of tracking end-uses that do not change with interactive effects. To track lighting and plug load energy

consumption, for example, a facilities manager needs only to create an energy data verification tool that tracks when a certain set of rules is violated. For this reason, it is suggested that the Energy Signal Tool not be set up to track these independent end-uses.

5.1.2 Parameter Uncertainty Characterization

The characterization of parameter distributions is what ultimately determines the degree of expected variability in energy consumption, and thus impacts the sensitivity of the Energy Signal Tool. While best practice would involve conducting audits of multiple similar facilities and noting the observed range of operational parameter values, the purpose of this research was to demonstrate the extent to which sub-metering is useful on its own and this type of audit survey was outside of the scope. Burhenne (2013) describes four potential sources for information that can be referenced to characterize parameter uncertainty distributions:

1. Physical bounds (whereby a uniform distribution is assumed lacking further information)
2. Literature references
3. Expert knowledge
4. Measurements

A combination of sources (1), (2), (3), and (4) were used to describe the probable range of distribution of each parameter in the work presented here. Many building parameters were deduced with high confidence during the initial building audit and through subsequent conversations with facility engineers. Others were not possible to deduce from any available knowledge of the building or its physical properties. To this end, a literature review provided some insight into parameter characterization with regard to best guess values and probable ranges. Notably missing from the literature are sources on the bounds and distributional characteristics of operational equipment, such as DX cooling equipment of a certain age, refrigeration capacity factors, temperature and flow

sensor/setpoint error, etc. Experience and judgment of the research team was applied to characterize these distributions and their bounds. The distributions of uncertain parameters will of course also depend heavily on idiosyncrasies of the particular building type. The following parameters' uncertainty ranges were characterized with methods from the sources identified:

- Infiltration (2013 ASHRAE Handbook of Fundamentals[vii], and Stein, 2012[viii])
- Wall and Ceiling conductivity (Dominguez, 2009[ix])
- Wall mass specific heat/thermal mass (MacDonald, 2002[x])
- Indoor temperature variations (Corrado, 2009)

Initially, a list of 37 parameters was selected, for which certainty was deemed to be less than 100%, describing the building, its occupancy, its internal loads, setpoints, HVAC and refrigeration equipment and controls. Refer to the list of these in Figure 32. These distributions were given probability density curves that would best describe them (uniform or triangular, as limited by sampling interface). Following the lead of many sources (including: de Wit, 2003; Hopfe & Hensen, 2011; Burhenne, 2013) sampling was done from the PDFs' central 95% confidence intervals. Refer to the appendix for a detailed description of each initial uncertain parameter and its distribution range.

5.2 Sensitivity Analyses and Model Reduction

5.2.1 The Importance of Model Reduction

Coakley *et al.* (2014[xi]) point out the major flaw in using detailed building energy simulation models is that they are so detailed as to be over-parameterized and under-determined. This is what makes the calibration process so challenging; there are thousands of knobs to turn which can potentially impact the outcome of the model. Testing all possible combinations requires a great deal of computational power (see Table 4 below). Eisenhower *et al.* (2012a[xii]) demonstrated an analysis capable of assessing over 1000 uncertain building parameters for optimization of design. This approach involved deriving

a meta-model of a building used a quasi-Monte Carlo parameter space search. Of the 1009 parameters tested in sensitivity analysis, only 2% of these were found to be significant – those with total significance index of greater than 0.05. A meta-model approach, which is essentially a statistical fit of the parameters to a curve, can be useful for a static model in design optimization. However, it would need to be re-derived multiple times for a dynamic model of an operational building viewed in different time scales under different operating conditions.

Table 4: Reducing degrees of freedom increases sampling effectiveness

# Of Parameters	# Combinations; three sampling levels
50	7.18E+23
20	3.49E+09
15	1.43E+07
10	59,000
7	2,190

A more appropriate approach given the current computing power available for this work (which is far greater than was available10 years ago, and will have greatly improved again in another 10 years, see Burhenne, 2013[xiii]) was to selectively diminish the number of uncertain model parameters. Saltelli *et al*. (2008) recommend that for computationally expensive models with many uncertain input parameters, a combination of local pre-screening and global sensitivity analysis be used to reduce the parameter space. De Wit and Augenbroe (2002) used the Morris method of sensitivity analysis to examine 89 model parameters under uncertainty. They isolated 13 that were significant to design decisions related to thermal comfort performance of an office building. Another study used stepwise regression and the standardized regression rank coefficient to find that eight most influential physical parameters appropriately characterize an office building model for the purposes of design decisions (Hopfe and Hensen, 2011). Corrado and Mechri (2009[xiv])

found, using the Morris method, that only eight of the initial 104 input parameters were significant to characterizing the energy rating of a single family home. Bucking *et al.* (2014) found that of 26 variables examined for importance factor, only eight were influential for achieving a Net Zero Energy house design.

In this work, the amount that the parameter space can be diminished is limited by the need for sufficient parameters to fully describe the expected distributions of energy end-uses. While ten or fewer parameters are sufficient to characterize *whole* building energy use, they would likely not be sufficient to describe the potential uncertainty in *each* energy end use. The goal of parameter reduction is to achieve a balance between computational efficiency and sufficient propagation of uncertainty in the distributions of expected consumption for each end use. If there is insufficient uncertainty in the model, the Energy Signal Tool will be over-sensitive to anomalies.

The first step in model reduction, as noted by Saltelli *et al.* (2000), is the elimination of redundant parameters. This was done as part of the initial selection of the 37 potentially influential parameters. Examples of redundancy in uncertain model parameters includes:

- Minimum outdoor air ratio, and infiltration rate
- Wall/roof insulation thickness, and insulation R-value
- Number of occupants, and heat gain per occupant (lacking CO_2 controlled ventilation)

5.2.2 Local sensitivity analysis parameter pre-screening

From this list of potentially significant variables, a local sensitivity analysis was performed to identify those parameters that exhibited the most significant impact on the energy end use categories. The term local implies that each parameter is tested only at one location of the sample space. A "one at a time" (OAT) local sensitivity analysis will qualitatively rank the effects of individual parameters on the model. This is an acceptable method of parameter pre-screening which minimizes computational power (Saltelli *et al.*,

2000). Saltelli notes that basic OAT methods are effective for the initial reduction of parameter dimensionality, but they are not sufficient for full reduction or quantitative parameter influence ranking. See section 5.3 for a discussion of secondary (global) sensitivity analysis. In other examples of OAT pre-screening, Eisenhower *et al.* (2012b[xv]) used the "significance index" metric (SI), similar to the "importance factor" defined by Bucking *et al.* (2014[xvi]). To carry out an OAT analysis, model outcome is tested by varying one parameter at a time, while other parameters in the model remain constant (Saltelli *et al.*, 2000[xvii]). In this work, the results of parameter pre-screening shall be characterized by the nomenclature SI (significance index). Results come from the following formula:
For Parameter *i* relative to each end use, *j*;

Equation 2: Pre-screening parameter significance index for local SA

$$\boldsymbol{SI_{i,j} = \frac{(E_{j,high} - E_{j,low})}{E_{j,high}}}$$

Where, E_{low} and E_{high} represent the energy use due to the extreme high and low values from the parameter distribution 95% confidence range for all parameters, *p*. The objective functions, *j*, which each parameter is tested on include the annual sums end use categories of the following model objective outcomes:

j_1: Whole building energy use (WBE)

j_2: Natural Gas use (NGAS)

j_3: Fan, Heating, and Cooling Electricity Use (HVAC)

j_4: Refrigeration Energy Use (REFR)

The results of parameter pre-screening are presented below in Figure 32 and Figure 34. More detailed results can be found in the Appendix. A significance threshold is drawn at the 50th percentile.

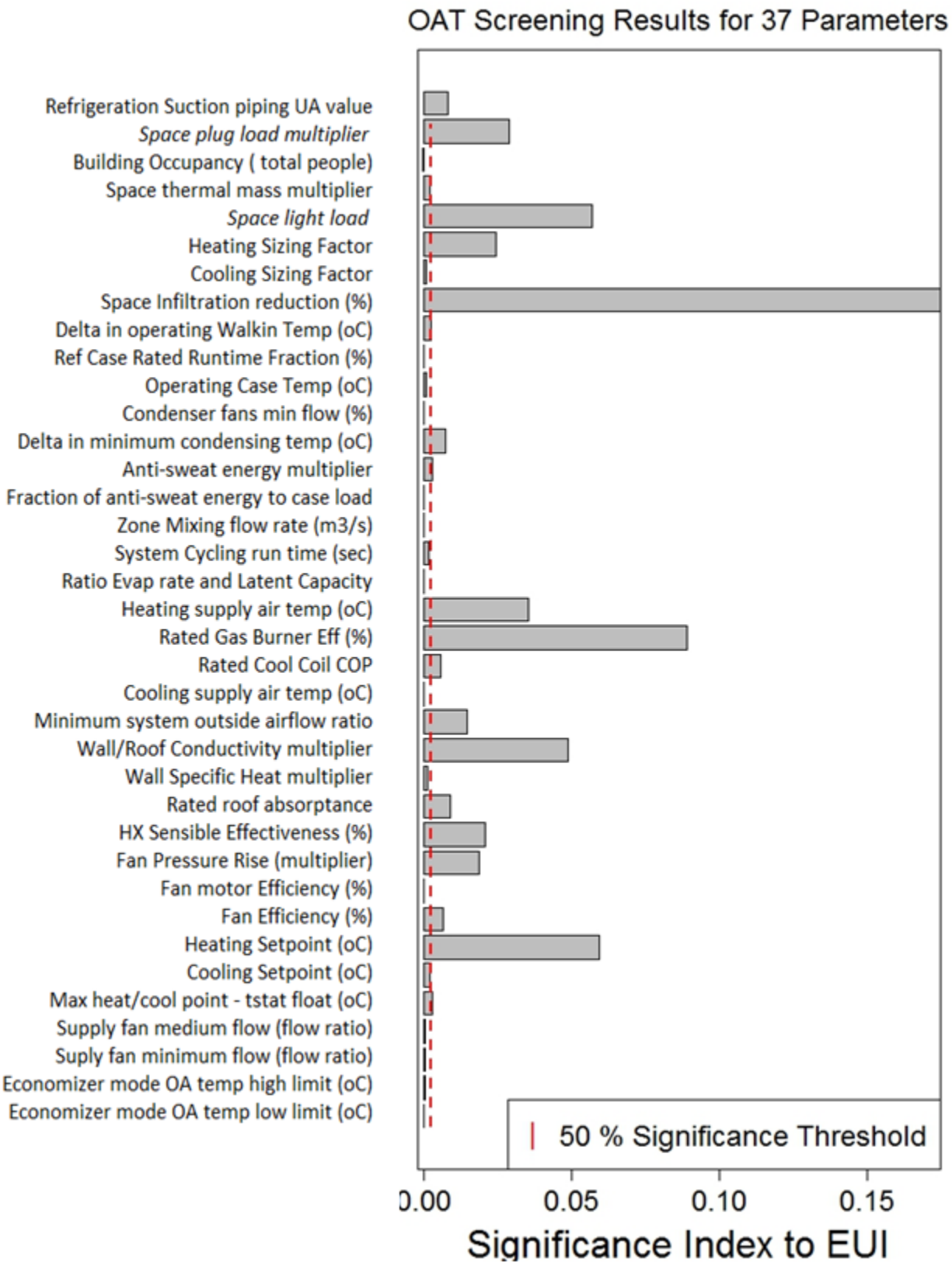


Figure 32: Results of OAT parameter pre-screening for effect on total energy use (EUI)

The results of parameter pre-screening identified 20 model parameters with uncertainty ranges that propagated to greater than 0.2% significance in overall building energy use. It was possible to model uncertainty distributions in 18 of these[9]. Only nine of these contributed greater than 2% significance to overall building energy. Figure 34, below, gives results on the basis of uncertainty propagation to individual end-uses – showing how many parameters have greater than 1% significance to each end-use. There are between only four and as many as 13 parameters that contribute significantly to uncertainty (greater than 1%) in each end use (refer to Figure 33 below). Those four that were significant to refrigeration were ranked in the bottom ten in relation to any other end use. Providing enough uncertainty range for refrigeration energy was important justification for retaining all 18 parameters in the model for the Energy Signal Tool.

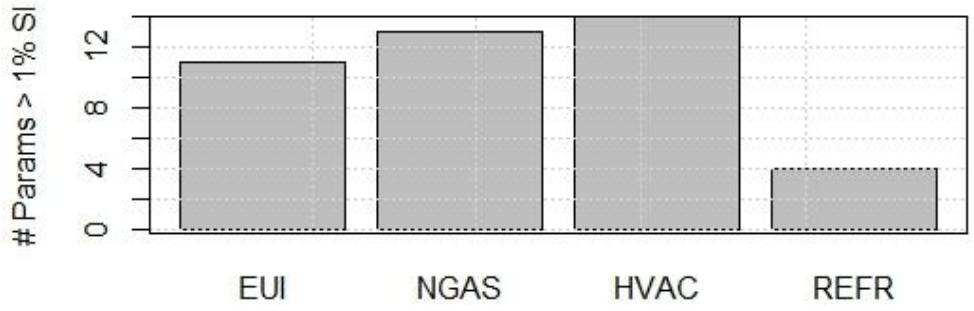


Figure 33: Summary of number of parameters significant to each end-use

[9] Due to technical limitations of the OpenStudio measure writing language, it was not possible to model uncertainty in lighting or plug load schedules. Furthermore, as determined through analysis of sub-metered data, lighting and plug load schedules exhibited normal distributions in their ranges of variation within any season; meaning that there will be little long-term aggregate effect on observed consumption. Schedules with seasonal variation were incorporated into the model.

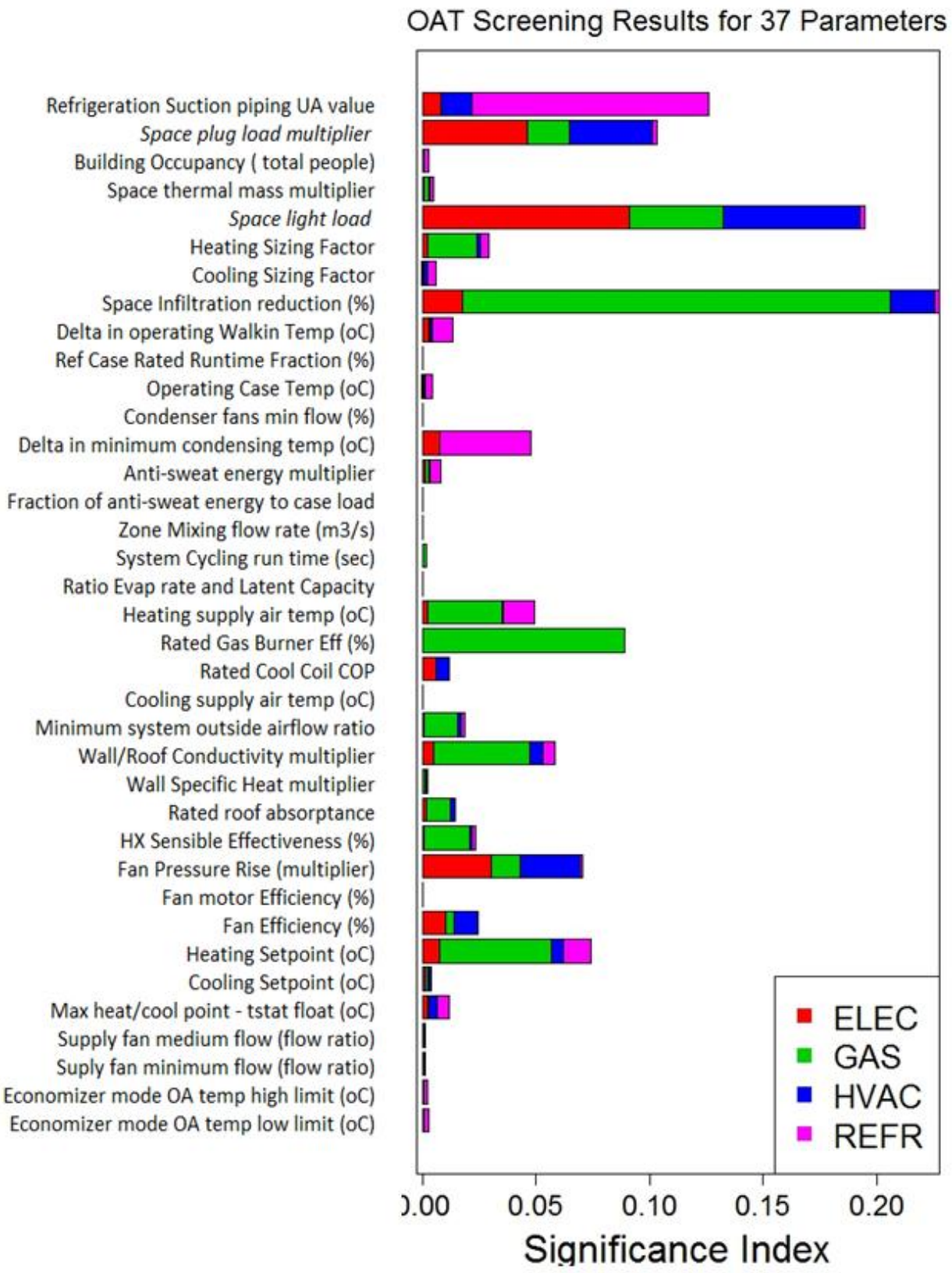


Figure 34: Results of OAT parameter pre-screening with SI results for each energy end use

5.3 Global Sensitivity Analysis

Like local sensitivity analysis, global sensitivity analysis seeks to quantify the relative significance that uncertainty within each parameter has on the model. Global SA is superior to local SA in that it acknowledges that the effects of a differential change in one set of parameter under certain model conditions could produce a different result under different global model conditions. Global SA is important for any non-linear model, and especially so for multivariate models. The result of the global sensitivity analysis is intended to be a quantitative ranking of parameters' influence; where a parameter is influential if its affect on the model objective outcome is proportionally larger than that of others (Bucking *et al*., 2014). Although no further reduction in the model parameter space was desired (for the reasons mentioned above related to end-use variance), global sensitivity analysis was used to assign the quantitative influence of each parameter on each end use. This is for later application to the Energy Signal Tool described in Section 7.3.1.

Two methods of global sensitivity analysis were tested in this work, and these are described in the following sections.

5.3.1 Global Sensitivity Analysis using the Chi Squared Statistic

Reddy (2007) suggests combining the LHS process of sampling for model calibration fit with identifying significant model parameters. This can be done using the chi-squared (noted as χ^2) metric of sensitivity. An initial batch of (as Reddy suggests) 1000 LHS sampled simulations are run, and total goodness of fit (GOF) criteria are computed for each model. From the sampled results that match this criteria, an analysis of the distributions of parameter values is performed. Generally, those parameters that exhibit more random distributions across the set of models which conform to GOF criteria will be weaker parameters. This non-random pattern test can be quantified using the chi-squared metric as follows:

Equation 3: Chi-Squared metric for calibrated solutions

$$For\ all\ runs\ which\ the\ following\ is\ true:$$

$$GOF_{NBME} < 5\%, and\ GOF_{CVRMSE} < 15\%;\ \chi^2_{param} = \sum_{s=1}^{b} \frac{(P_{obs,b} - P_{ex})\text{^}2}{P_{ex}}$$

Where $P_{obs,s}$ and P_{ex} are the observed and expected count of parameters found in the GOF solutions from each bin range of the input. The term "*b*" is the number of bins into which the sampled parameter values are discretized. This equation has been adapted and generalized slightly with the addition of the *b* term to incorporate parameter values coming from a continuous distribution, rather than discrete ranges as suggested by Reddy. For example, in a simple example where a parameter is assumed to be distributed in a uniform fashion with five bins, and there are 25 sample models which fit GOF criteria, a completely random observation would result in five instances of each parameter value within each bin. Where distributions of sampled parameters are non-uniform, determining the bin divisions is slightly more complicated.

For any number of bins, the threshold value of each bin division is determined by setting the integral over the range equal to the inverse of the total number of bins, as follows:

Equation 4: Parameter range bin boundaries for Chi-Squared metric

$$\int_{x=min}^{x=B1} \boldsymbol{f(x) = \frac{1}{\#Bins}}, \quad \int_{x=B1}^{x=B2} \boldsymbol{f(x) = \frac{1}{\#Bins}}, \ldots \int_{x=\#bins-1}^{max} \boldsymbol{f(x) = \frac{1}{\#Bins}}$$

Where f(x) describes the uncertain range of each parameter, and {B1, B2, …B-1} are values being solved for that make the above equations true, and form the bounds of the bin ranges. $P_{obs,b}$ is found by counting the occurrences of variable values that fall in each bin:

Equation 5: Probability observed in the calibration solution set for the Chi-Squared metric

$$\boldsymbol{P_{obs}} = \sum_{P=min}^{P=max} (\boldsymbol{Bin\ lower\ bound}) < value(\boldsymbol{P}) < (Bin\ upper\ bound)$$

Polly *et al.* (2013) also used the Chi-Squared metric to produce a ranked sensitivity analysis and isolate "strong" parameters from weak ones. They executed 2,400 model runs of a residential building and computed the Chi-Squared statistic from a subset of 73 models meeting monthly calibration requirements for GOF.

The advantage of determining the expected number of observations using the modified integral bin approach when calibrating in parallel with sensitivity analysis is that the uncertainty *distribution characteristics* of each variable are preserved, and calibration sampling is more effective. Reddy suggests that three discrete values be used for each variable. This is quite limiting, and in this work, each variable will be classified based on five bins (using five divisions is consistent with a method used later on).

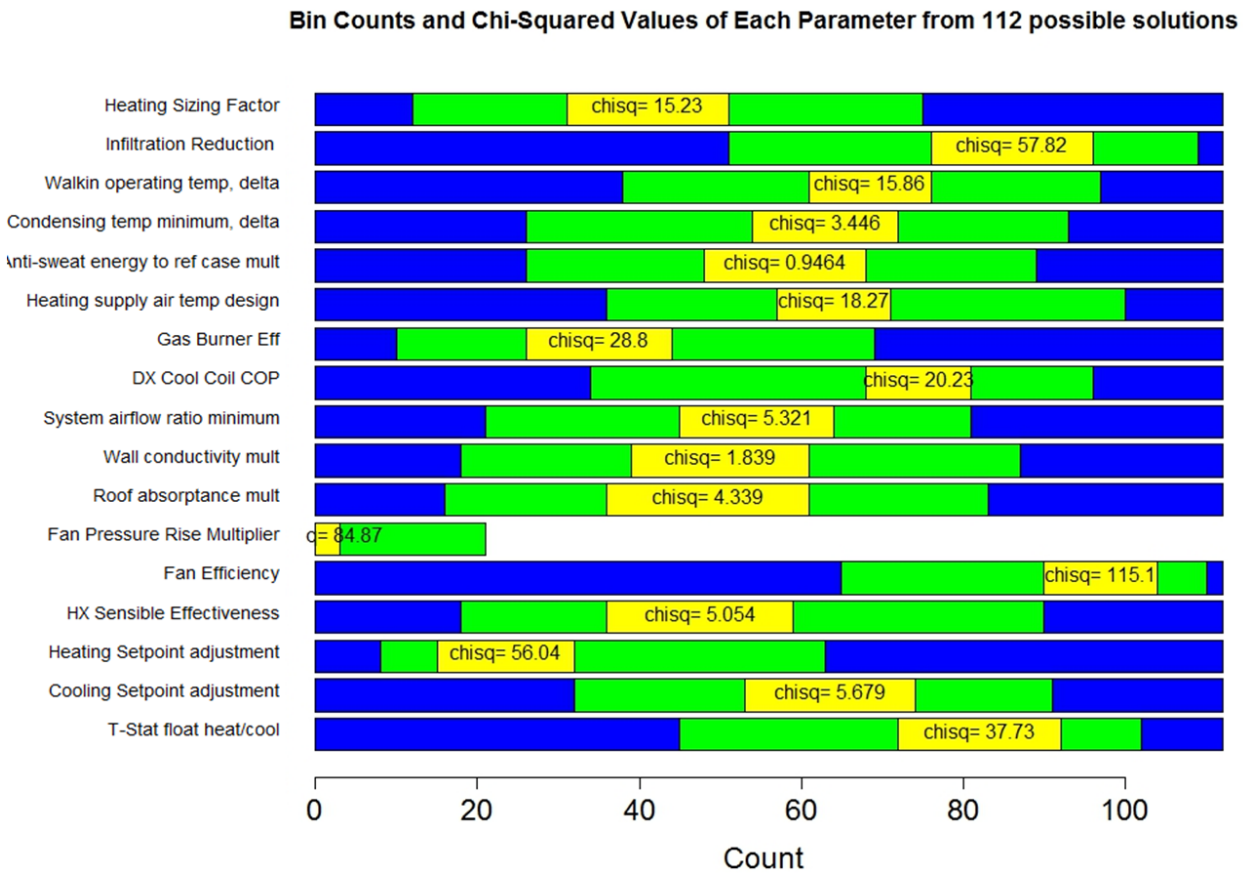


Figure 35: Chi-Squared test results. Each color band is a bin, shown in order from Bin #1 to Bin#5.

Table 5: Chi-Squared test results

Parameter	χ^2 value
Fan Pressure Rise Multiplier	281.35
Fan Efficiency	121.43
Infiltration Reduction	100.37
Heating Setpoint adjustment	90.46
T-Stat float heat/cool	50.62
DX Cool Coil COP	50.13
Heating Sizing Factor	40.78
Walkin operating temp, delta	39.40
Heating supply air temp design	22.24
Gas Burner Eff	15.09
Wall conductivity mult	15.09
Cooling Setpoint adjustment	13.71
HX Sensible Effectiveness	12.65
System airflow ratio minimum	10.70
Roof absorptance mult	9.97
Condensing temp minimum, delta	7.61
Anti-sweat energy to ref case mult	2.81

The chi-squared metric is used in the following hypothesis formulation to assess the distribution of samples within the vector space of the potentially calibrated solutions:

H_0: samples are randomly distributed and the parameter is weak

H_1: samples are not randomly distributed and the parameter is strong

Test Statistic: **T_o** ; Is abs(t_o) > $t_{(\alpha/2, n-2)}$? For:

- $\alpha = 0.01$ (parameter is definitely strong)
- $\alpha = 0.5$ (parameter is stronger than half of parameters)

If this is true, we reject the null hypothesis and declare the parameter to have strong influence on the calibrated solution.

Reddy (2007) states that any parameter with a chi-squared value of greater than 9.21 (α = 0.01) is likely, with a 99% chance, to be a strong variable. The threshold division found in this work between strong and weak parameters (α = 0.5, 50th percentile of strength) is a chi-squared value of 16.

The results presented in Table 5 above are not at all consistent with the results from earlier OAT pre-screening analysis. In fact, they are so strongly different as to de-merit their use. If the reason for non-randomness is convergence of a parameter value to a solution, then there are several reasons that a parameter distribution may exhibit this behavior. While the Chi-squared statistic may be valid as Reddy demonstrated with uniform parameter uncertainty distributions, it may not be valid in this case of non-uniform continuous parameter distribution ranges. If, for example, the initial guess of the shape of a parameter distribution is quite close to the calibration results, then the Chi-Squared test will declare the results to be random, when actually they simply reflect a good initial guess. Chi-Squared results may also be attributed to parameter correlation in the solution set. Figure 47 above shows there is some significant correlation between fan efficiency and actual cooling setpoint. If the results show a bias towards some value in the range of cooling setpoint, then the fan efficiency may follow (or vice versa).

5.3.2 Multivariate Regression Sensitivity Analysis

Saltelli *et al.* (2000) recommend utilizing a stepwise regression analysis to sort out the most significant parameters from the least significant. Similar to the chi-squared metric, this can also be achieved in parallel to LHS sampling for calibration. When there are multiple variables affecting the outcome of some model, the multiple regression method can be used to determine a linear model that is a function of all variables (Montgomery and Runger, 2010). This linear model attempts to show the relationship between model outcome and the value of each input variable. Per Montgomery and Runger (2010), the general form of a multivariate regression model is shown below:

Equation 6: General multivariate regression model form

$$\boldsymbol{Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \ \varepsilon}$$

Where;

Y = objective function outcome (energy end-use)

x_j = the jth input variable

β_j = the expected range in response of Y for x_j when all other parameters are held constant

ε = error between linear approximation and true model

This equation holds true for cases where model variables are independent; that is, when there are not interactive effects amongst variables on the model outcome. Typically, the method of least squares (LSE) or ordinary least squares (OLS) regression is used to estimate the regression coefficients (β_j) of the model. With the goal of minimizing the sum of squared error term (ε), least squares regression coefficients are estimated based on multiple samples (n) taken over all model variables (k). LSE is computed as follows according to Montgomery and Runger:

Equation 7: Least squares error computation for regression fit

$$\boldsymbol{LSE = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(y_i - \beta_o - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2}$$

The stepwise regression method can be used to select regression coefficients based on a large sample of data containing variable values and true model outcome – like that produced with LHS sampling. Stepwise regression builds a linear model by successively adding variables in order of their F-statistic (significance) on the overall model to minimize LSE and optimize fit (Montgomery and Runger, 2010). A good fit is important to achieve in this method, but model fit is not the end goal of this analysis. Rather, the goal of multivariate regression can be stated as quantifying the value that each variable contributes to the overall model. The incremental change in the coefficient of

determination shows how closely related the changes in the key output indicators are related to changes in the assumed parameter value.

In this work, the AIC (Akaike's information criterion) value is used to measure of overall goodness of fit of a model of variables to observed data. The AIC value essentially informs the user of how much information is lost from the model when any given parameter is removed (Akaike, 1974). During an AIC stepwise regression algorithm, model variables and coefficients are combined in succession of a decreasing AIC value assigned to each as it is added to the model. A lower model AIC value indicates better model quality. The AIC value, rather than the F-statistic, was selected based on the availability of a statistical processing package (described below).

While a linear model will not replicate exactly a real process such as building operation, many researchers have shown this method to form adequate approximations. Hopfe and Hensen (2011) take a regression model construction approach to parameter sensitivity analysis. In their study, a stepwise regression is performed (following rank transformations), whereby each parameter added to the model gives a progressively higher R^2 coefficient of determination value. The stepwise regression continues, adding parameters individually until the model R^2 value cannot be further improved. Other authors have used multivariate stepwise regression techniques to identify significant building features used in developing statistical comparisons for building performance data to buildings in benchmarking databases. Sharp (1998) demonstrated the usefulness of OLS regression for narrowing from 75 possible categorical variables to six which are significant determining building EUI. Gao and Malkawi (2014) also utilized stepwise OLS regression for extracting significant features impacting building energy performance.

In this work, a multivariate regression model was created for each end use based on the data generated by the 8422 LHS sample runs used for calibration. The building energy model has inherent non-linear relations between some inputs and objectives. Saltelli

(2000) recommends performing a rank transformation; whereby the values of each input and output variable are simply converted to ranked values (i.e., 1:8422) within their set. Following this substitution, the '*stepAIC*' function in R (see 'MASS' package documentation; Venables and Ripley, 2002) was utilized to assess AIC value of the linear best fit. The stepAIC function finds an optimal linear model by minimizing the AIC value of the model containing all of the parameters with any significance to the objective function. Once this is done, the function provides the AIC value that the model *would* have if each parameter were to be removed from the linear regression model. The magnitude of the increase in the AIC value of the model *without* each parameter will be less for relatively weak parameters and greater for strong parameters.

The relative difference in the AIC value (rather than R^2 value) was used to quantify the significance of each parameter relative to the model output of each end use. As mentioned, both parameter value and objective function data existed for 8422 LHS sample runs. For each objective function (EUI, NGAS, HVAC, and REFR), a linear fit was calculated, and differential AIC values were calculated for all parameters in the linear model. For all objective functions, there were some parameters that had such weak influence on the model they were not included in the linear approximation. The relative significance of each parameter in the context of each end use was then calculated as follows:

For all parameters, *k* and objective functions, *j the regression significance is calculated as*:

Equation 8: Parameter significance index for global SA

$$SI_{k,j} = \frac{AIC_{k,j} - AIC_{all}}{AIC_{max,j} - AIC_{all}}$$

Where;

$AIC_{k,j}$ = The AIC value of the model for objective function j with parameter k removed

AIC_{all} = The AIC value of the best fit model with all parameters (minimum AIC)

$AIC_{max,j}$ = The maximum AIC value of the model with any one parameter removed

The rank transformation was found to improve results of the fit (lower coefficient of determination, R^2) only for the refrigeration energy model. The EUI, HVAC, and Natural Gas models all produced better results without a rank transformation. Thus, the rank transformation method is used only for the refrigeration end use in this work. The following table gives a summary of this comparison.

Table 6: Comparison of regression methods; with and without rank transformation of variables

Objective:	*With rank transformation*		*Without rank transformation*	
	R^2	**AIC**	**R^2**	**AIC**
Total Energy Use	0.9016	118,824	**0.9615**	48,021
Natural Gas	0.9012	119,149	**0.9579**	47,328
HVAC (Cooling and fans)	0.9418	114,278	**0.9866**	12,975
Refrigeration	**0.8215**	123,590	0.7932	8,849

5.4 Final Parameter Significance Results from Global Sensitivity Analysis

Parameter significance values were calculated for those parameters that were found to be significant to each end-use. For each end use, the parameter with the most significance had its *SI* value normalized to unity, such that all those with lesser significance are expressed as having a fractional *SI* value. Results are shown below in

Table 7, with all rankings being in reference to ranking for parameter significance on whole building energy.

Table 7: Significance results of multivariate regression analysis, RS values shown

Parameter	WBE	NGAS	HVAC	REFR
Space infiltration reduction (%)	1.000	1.000	0.376	0.027
Gas Burner Eff (%)	0.652	0.711	0.000	0.000
Heating Setpoint (oC)	0.408	0.353	0.021	0.269
Wall/Roof Conductivity multiplier	0.302	0.282	0.079	0.033
Heating supply air temp (oC)	0.087	0.052	0.001	0.153
Heating sizing factor	0.095	0.076	0.223	0.002
HX Sensible Effectiveness (%)	0.157	0.169	0.001	0.003
Fan Pressure Rise (multiplier)	0.056	0.027	1.000	0.015
Minimum system outside air ratio	0.062	0.061	0.006	0.008
Rated roof absorptance	0.037	0.053	0.094	0.000
Refrig suction piping UA value	0.007	0.002	0.036	1.000
Condensing temp minimum, delta	0.003	0.000	0.001	0.467
Fan Efficiency (%)	0.005	0.006	0.555	0.001
DX Cool Coil COP	0.011	0.000	0.375	0.000
Anti-sweat energy multiplier	0.000	0.000	0.000	0.000
T-Stat float heat/cool (oC)	0.034	0.004	0.243	0.057
Operating walkin ref temp (oC)	0.002	0.000	0.022	0.139
Cooling Setpoint (oC)	0.010	0.000	0.053	0.000

Compared to the OAT pre-screening results, the results from global SA were expected to be slightly different, but not entirely so. This expectation held true. Figure 36, below, summarizes the comparison between results of OAT parameter screening and multivariate regression significance index calculations for each end-use[10]. For the end-uses of whole building energy, natural gas, and refrigeration results from the two methods were quite similar. The HVAC end-use exhibited the greatest differences, which is explained by the fact that the global analysis revealed more of the interactive effects between parameters affecting HVAC energy.

[10] The global sensitivity results show that the model could have probably been reduced further than 18 parameters. However, a significant amount of computational time and expense had already been invested in sampling with 18 parameters, and another round of sampling was impractical.

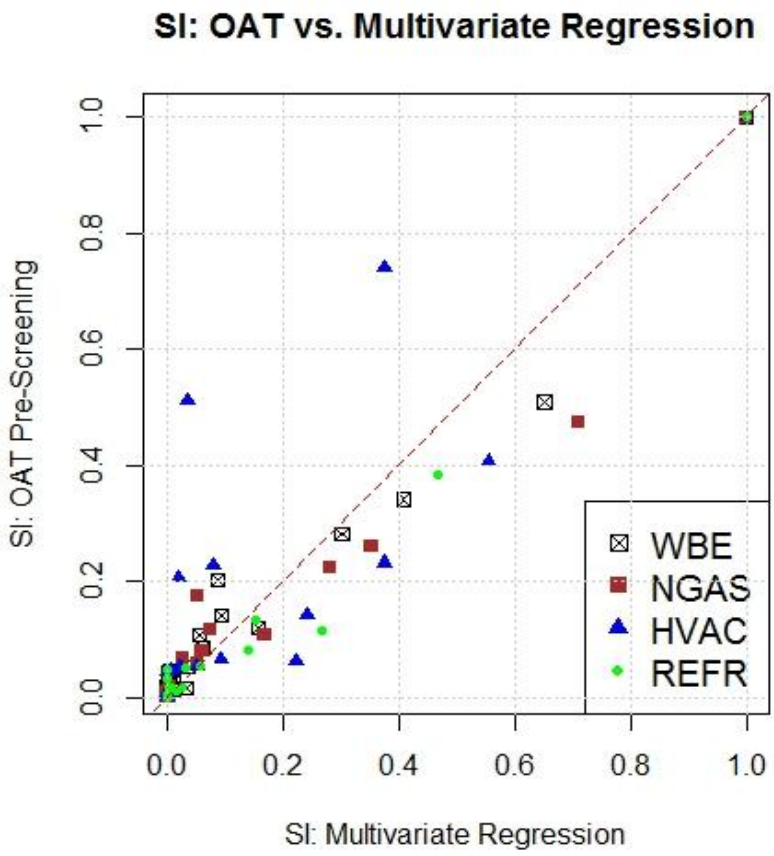


Figure 36: Significance index results for OAT and multivariate sensitivity analyses

As Loonen and Hensen (2013) have found, the response of a model to parameter perturbations varies not only with the sampled space of the other parameters but also over time. Since this work proposes the development of a dynamic performance assurance tool, the process of global sensitivity analysis shall be repeated on an ongoing basis throughout the duration of performance tracking. That is, as building controls or other model input parameters may change, an updated global SA process is needed. For example, it may occur that more information comes to light regarding an uncertain parameter (i.e., packaged DX rooftop AHUs are tested for operational cooling COP), shifting or narrowing the probable range of values and thus changing the influence of the range of uncertainty of this variable. This will ensure that the most significant parameters to the model are factored into objective function probable frequency distributions. Acknowledging the need for ongoing global SA, reducing computational expense for each SA is paramount.

5.5 Model Reduction Process Validation

It is important to make sure that model parameter reduction does not result in loss of descriptive information in the model. Model parameter reduction is valid to the threshold where it removing another parameter significnatly affects the distribution of expected energy consumption for any given end-use in any month. The process of model parameter reduction was validated at by comparing the distributions of a set of results from the model with larger parameter space with the set of results from the model with the new, smaller parameter space[11]. Reddy (2011) recommends a minimum of N*(k+2) model runs each; where *k* is the number of uncertain parameters, and *N* is the number of distinct values for each parameter. The target for N shall be 10 samples each.

37 parameters = 10∗39 = 390 runs

18 parameters = 10∗20 = 200 runs

Figure 37, below, gives a comparison of all end-use distributions for both the 18 and 37 parameter model. As evidenced visually by the minimal shifts in each end use, there appears to have been little information removed from the model by screening out the 19 parameters with least significance.

[11] It was discovered through this validation that the refrigeration energy end use was under-parameterized in both the 36-parameter and 17-parameter models. This conclusion was drawn from the narrow range of 95% confidence interval in the expected energy distribution. In response, another parameter influential to refrigeration was added to the model. The consequence of an under-parameterized end-use would be a narrow distribution that would cause many false alarms from the Energy Signal Tool – see further discussion in Chapter 8. It would also not have been compatible with the minimum risk tolerance range of +/- 3% measurement error that is possible to obtain from energy metering equipment.

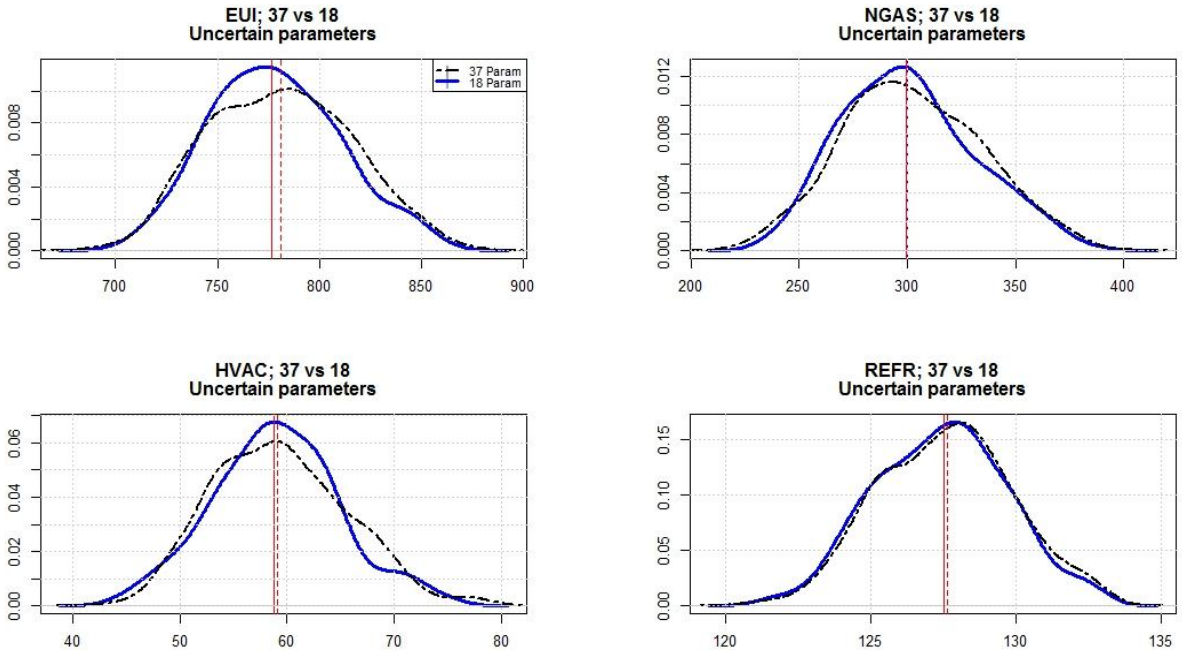


Figure 37: Comparison of the annual end-use distribution results from a 37-parameter model and 18-parameter model. Sample medians are shown with a vertical line.

Each set of distributions of results, shown in Figure 37 above, can be compared with formal statistical methods that test for a difference in the mean values of two independent populations (Montgomery, 2007). In general, this test for difference in population means is formulated as:

For population 1 being the results of the model with 37 uncertain parameters, and population 2 being the results from the model with 18 uncertain parameters. The actual variance of the population is unknown, since it would not be possible to conduct Monte Carlo sampling if all model parameters are uncertain.

Null Hypothesis:

Ho: $\mu 1 - \mu 2 = 0$ (both treatments of the model, with 37 or 18 uncertain parameters, should produce similar end-use results)

Alternate Hypothesis:

H1: $\mu 1 - \mu 2 \neq 0$ (reject the null hypothesis if the two means are different)

Confidence Level: 99% (α=0.01)

Test Statistic: $z_o = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$

Where,

n1 = 380

n2 = 190

Rejection Criteria: Reject H_o if $z_o > 2.326$ ($z_{0.01}$)

Computations:

End Use:	$\bar{x}_1$	$\bar{x}_2$	S_1^2	S_2^2	**n1**	**n2**	**z_o**	**P-value**
EUI	733.358	730.962	950.232	742.489	380	190	0.946	0.172
NGAS	252.354	250.989	716.080	616.712	380	190	0.602	0.273
HVAC	298.831	297.250	777.765	674.560	380	190	0.668	0.252
REFR	126.424	126.194	1.199	0.840	380	190	2.639	0.004

Test Conclusions:

For the first three end-uses, there is strong evidence to accept H_o; that the distributions are very similar for the 37-parameter model and the 18-parameter model. For the refrigeration end-use, we reject H_o, because the p-value is less than the confidence level. This is explained by the fact the influence rankings in the one-at-a-time pre-screening process selected for those variables that had greatest effect on model EUI, and froze some of those that had most effect on refrigeration energy end use. This being concluded, the variance observed for refrigeration end-use consumption in both the 37-parameter model and the 18-parameter model was of similarly small magnitude. Therefore, the model reduction shall be considered valid nonetheless.

5.5.1 The Case of an Under-parametrized End-Use

In order to apply the Energy Signal Tool to an end-use of interest, there should be at least +/- 5% magnitude of the 95% confidence interval of expected energy use in any given month. Without this minimum amount of uncertainty, the Energy Signal Tool becomes un-reliable – as the uncertainty in the values of energy measurement itself is no less than +/3% (NTSC 2011). When the uncertainty distribution is this narrow, the probability masses used to calculate the expected cost will spill over the boundary of the uncertainty range. This is discussed further in section 7.3. It may be that there are not sufficient uncertainties that affect an end-use even with the fully parameterirized model (to reach +/- 5% deviation from median value), and no more reasonable uncertainty can be introduced. In this case, the end-use is not suitable for monitoring with the Energy signal tool, and would be better served by monitoring with rule-based FDD.

In the process of model reduction validation, it was discovered that the refrigeration end-use was under-parameterized. Figure 38 shows that the refrigeration end use was under-parameterized since the absolute magnitude varied by less than +/- 5%, even with a fully parameterized model. Another uncertain parameter (describing refrigeration piping heat transfer) was discovered and added to the model. Figure 39 shows this improved variance for refrigeration end use. See further discussion in section on tuning. Similar figures showing expected end-use distribution for the other three end-uses are found in the appendix.

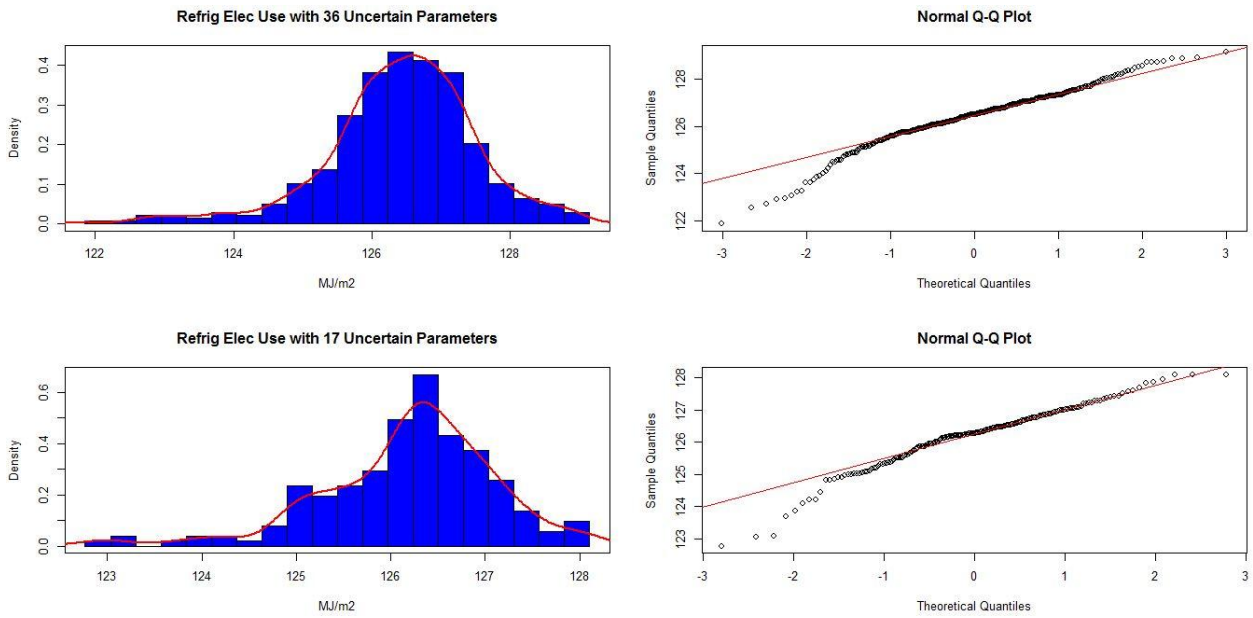


Figure 38: Refrigeration energy end use distributions and normality plots for full and pre-screened parameter set with 36 and 17 uncertain parameters

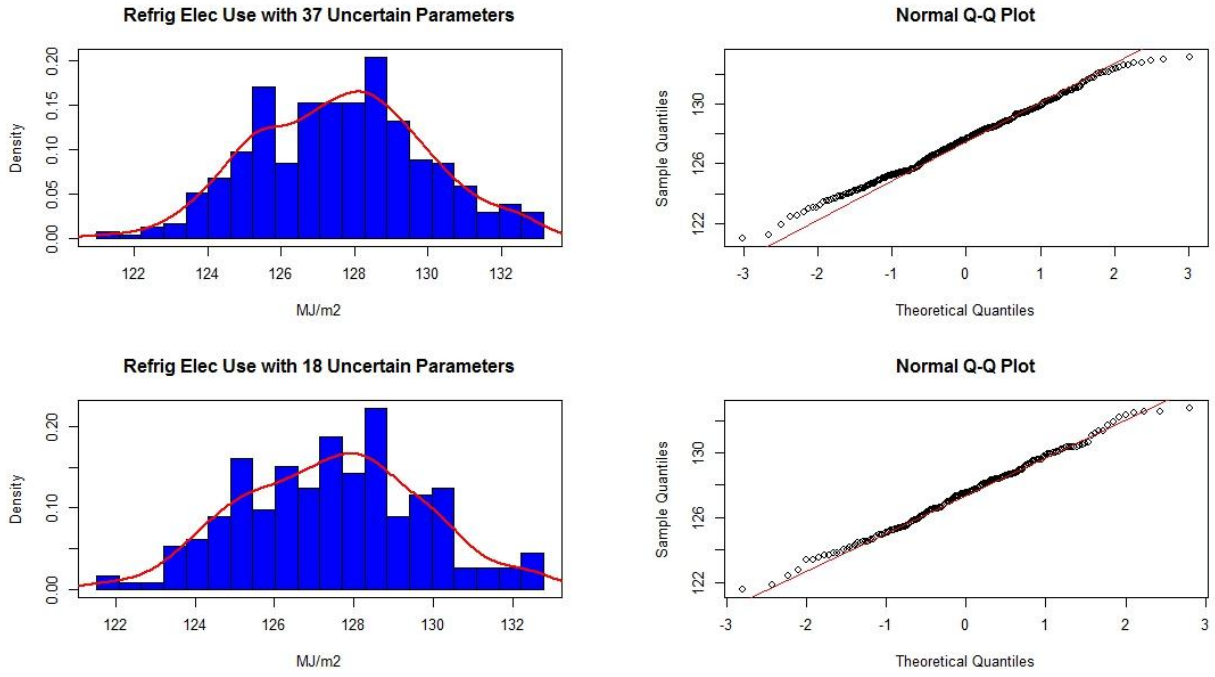


Figure 39: Refrigeration energy end use distributions and normality plots for full and pre-screened parameter set with 37 and 18 uncertain parameters

6 Model Calibration Methodology

The goal of model calibration is to obtain a detailed simulation model with output that most closely matches measured data from an operational building. This involves matching hourly or monthly simulated energy use to reported energy use from utility bills or sub-metered data. Typically, this measured data includes major category end-uses such as electricity, natural gas, steam, or district heating/cooling consumption. The process of model calibration involves the heuristic information gathering discussed previously, followed by trial-and-error tuning of uncertain parameters. The latter is discussed in this section. ASHRAE has published the best-practice standard for quantifying the quality of detailed building simulation calibration (ASHRAE-14, 2002). Reddy (2007a,b) has set forth a calibration process that includes additional metrics. Following these guidelines, one will result in obtaining multiple model solutions to calibration fit.

A well-calibrated building energy model has the capacity to predict the impact of building energy efficiency measures or to predict the expected energy consumption of a building given a certain set of environmental conditions. Parker *et al.* (2012[xviii]) compared the savings from an actual airport retrofit project to predictions made from an energy model at various stages of calibrations. They demonstrated that there can be a huge gap between retrofit savings predictions generated by a well calibrated model *vs*. a model prior to calibration.

For a testing study purely synthetic in nature, model calibration may be spurious; since test faults can be injected into any reasonable baseline model. After calibration of uncertain parameters, there is still uncertainty moving forward into future operational states. For the Energy Signal Tool presented in this work to be implemented in a building, calibration should be performed to produce a model that emulates measured data from the building in a state of *correct* operation. The calibrated model therefore represents a baseline of *operation as intended* against which to compare measured states of operation.

The purpose of going through the model calibration process in this work was to demonstrate and test the entire course of setting up the Energy Signal Tool, which includes extracting information from the results of a calibration process. It is also postulated that from the array of calibration solutions, a better idea of parameter distribution can be gathered. Since the Energy Signal Tool functions to assess building performance across a variety of discreet time periods, proper model calibration fit for *each period* is quite important. As discussed in a later section, this proves to be challenging for most buildings.

The over-parameterized and under-determined nature of detailed simulation models makes the calibration process challenging. These challenges are diminished by the parameter screening methods presented above, and by using statistical sampling to methodically search through every dimension of model parameter space. To do this, statistical calibration methods have many advantages over traditional hand calibration methods. These advantages include a more thorough search over possible parameter value combinations, less active time required, and minimization of user bias over parameter values that comes naturally with experience. The main pitfall of statistical calibration is that the process can lead to combinations of parameter values that are not plausible (Reddy, 2007). It is not possible to check for this manually, but EnergyPlus will sometimes crash if a model is described improperly. In this work, there were occasional infeasible parametric combinations that caused EnergyPlus to crash. With the parameter uncertainty distributions as described in the following section, this occurred at a rate of 0.5% of total simulations on average.

6.1.1 Latin Hypercube Sampling

Latin Hypercube parameter sampling (LHS) is one method within the Monte Carlo family of random sampling techniques. In this work, LHS was used to generate model samples for batch simulation. The LHS algorithm has been shown to result in the model parameter space being sampled in an agreeably uniform manner within a modeled multi-

dimensional parameter space with respect to computational power (Saltelli *et al.*, 2000; Reddy, 2007; Burhenne, 2013).

The algorithm for Latin-Hypercube sampling (LHS) works as follows (MacDonald, 2009[xix]):

1. For *n* samples, the distribution of each parameter is partitioned into *n* intervals of equal total probability
2. A point within each of the parameters' intervals is selected at random, to create a vector of parametric values comprising one model.
3. Step two is repeated over the entire number of iterations chosen for the analysis until *n* vectors are generated.

LHS leverages the assigned probability distributions as a stratification order for sampling, whereby more samples will be taken from the area of higher likelihood under the PDF curve. Therefore, LHS is not classified as a random sampling, but as *stratified* random sampling. According to the central limit theory of statistics, a greater size of random samples will inevitably result in a distribution of model outcomes that more closely matches reality. This is not true with LHS, because the sampling is actually quasi-random; in that not one combination of parameter values is repeated in the *n* vectors. Take for example the set of a uniform sequence of numbers from 1 to 10,000. If a random sample of size n=50 were to be taken from this set, the distribution would be much less than uniform from the range of 1 to 10,000. If the sample size were to be increased to n=500 and then n=5,000, the distribution of sampled values would get closer and closer to uniform. For LHS however, no matter the sample size, the sample distribution will match the sampled population (uniform in this case). Although the central limit theorem does not hold for LHS, increasing the sample size will yield a more robust set plausible solution models to choose from, since there will be more samples from the most likely ranges of each parameter.

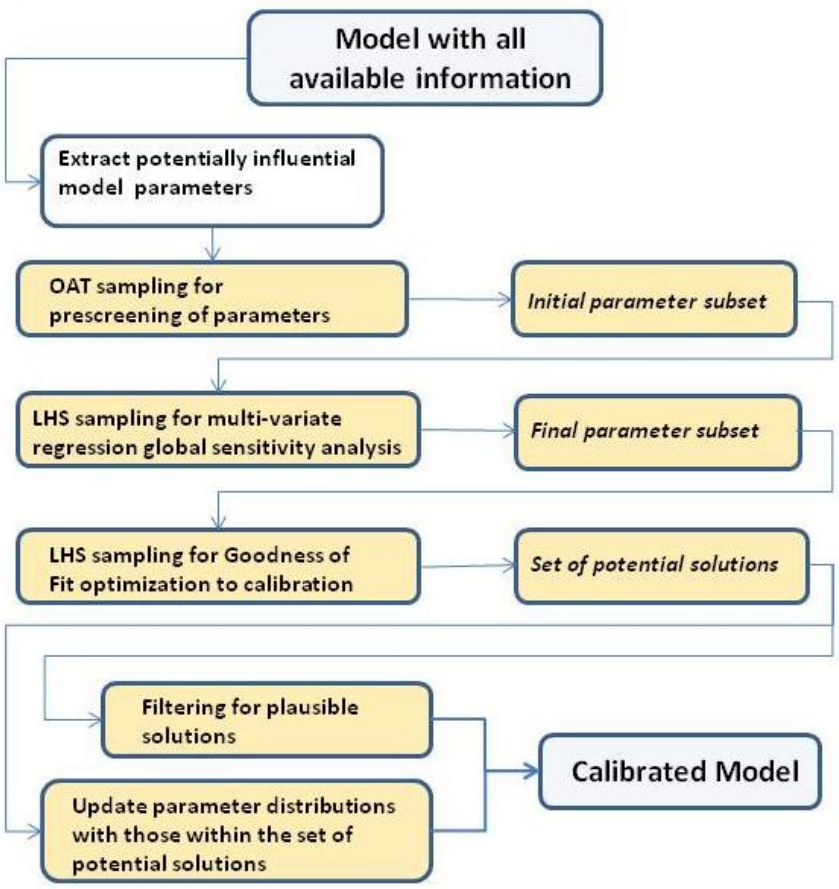


Figure 40: Workflow diagram for integrated calibration

6.2 Governing Equations for Calibration

Model calibration metrics seek a model that produces energy use matching measured data and minimizes the magnitude of matching error, as well as the variation in error, across all time periods. Per Reddy (2011), the model prediction fit to the observed data can be evaluated with the following equations:

Equation 9: Coefficient of variation in root mean squared error for calibration

$$CVRMSE = 100 * \frac{\left[\frac{1}{n-1} * \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{0.5}}{\bar{y}}$$

Equation 10: Normalized mean bias error for calibration

$$\text{NMBE} = \frac{\text{MBE}}{\bar{y}} = \frac{1}{n-1} * \sum_{i=1}^{n} \frac{y_i - \hat{y}_i}{\bar{y}} * 100$$

Where,

y_i = data used to calibrate against (utility bills or sub-metered energy data)

$\bar{y}$ = mean observed value of interest taken from n − number of observations

$\hat{y}$ = model predicted value of data of interest

n = number of observation sets; i.e., the number of months in the simulation process

The coefficient of variation of the root mean squared error (CVRMSE, Equation 9) describes the standard deviation of the residuals (modeled values less the observed values) weighted by dividing by the mean observed sample value. The normalized mean bias error (NMBE, Equation 10) describes the mean value of the residuals, normalized by dividing by the mean observed sample value. This is essentially the magnitude of the error with respect to the value being tested for.

Table 5-2 of ASHRAE Guideline 14-2002 (ASHRAE, 2002), provides guidelines for acceptable baseline model uncertainly thresholds defined by NMBE and CVRMSE metrics.

Table 8: ASHRAE Guideline 14 calibration metric thresholds

Resolution of energy consumption history	NMBE tolerance	CVRMSE tolerance
Hourly Calibration Data	+/- 10%	+/- 30%
Monthly Calibration Data	+/- 5%	+/- 15%

In this work, the model is calibrated to monthly utility bills, which have been normalized to calendar months. Some hourly sub-metered data was available, however the crucial component of gas usage is not being metered accurately. As a result, it was not possible to calibrate to hourly data. As mentioned previously, hourly sub-metered data was

used before calibration to the extent possible in defining load magnitudes and end-use schedules for lighting and plug loads.

6.2.1 Calibration Goodness of Fit Metric

A "Goodness of Fit" metric (GOF) is used in this work to determine which models are sufficiently calibrated by applying a weighting criteria to NMBE and CVRMSE that balances electricity and natural gas calibration results. Reddy (2011) gives the goodness-of-fit equations that quantify the closeness of the model to the actual building observed data for the dual objectives of matching end use fuels (electricity, gas, steam, etc.) consumed. Per the work of Henze and Harmer (2014), the different energy end-use metrics (electricity kWh, and gas therms) are combined with weighting ratios according to the percent cost to the owner of each during the calibration period, seen in Table 9 below.

Table 9: Energy cost weighting factors for calibration metrics

End Use	12-month cost	% of Total Cost
Electrical Energy	$*****	87.3%
Natural Gas	$*****	12.7%
Total	**$*******	**100%**

**Costs are redacted for retail partner privacy*

The GOF weighting equations for NMBE and CVRSME are as follows:

Equation 11: Goodness of fit for normalized mean bias error for calibration

$$GOF_{NMBE} = \left[\frac{w_{therms}^2 NMBE_{therms}^2 + w_{Elec}^2 NMBE_{Elec}^2}{(w_{therms}^2 + w_{Elec}^2)}\right]^{1/2}$$

Equation 12: Goodness of fit coefficient of variation for calibration

$$GOF_{CV} = \left[\frac{w_{therms}^2 CV_{therms}^2 + w_{Elec}^2 CV_{Elec}^2}{(w_{therms}^2 + w_{Elec}^2)}\right]^{1/2}$$

Where;

w_{therms} = *weighting factor of gas usage associated with the annual gas cost percentage of total energy cost*

w_{elec} = *weighting factor of electricity usage associated with the annual electricity cost percentage of total energy cost*

According to ASHRAE-14 2002, any unique parameter combination model solution is sufficiently calibrated to monthly utility data if GOF_{CV} and GOF_{NMBE} indices are less than 15% and 5%, respectively.

6.3 Model Calibration Algorithm: Objective Optimization

The batch runs generated with the OpenStudio Analysis Tool and OpenStudio server are automatically assessed for their NMBE and CVRMSE (Equation 9 and Equation 10) fit to electricity and gas consumption. This is as compared to a 12-month period of calibration utility data, and is reported for each simulation run in the batch. Additional scripts were written to process these figures into GOF metrics presented in Equation 11 and Equation 12. In this way, the large batch of runs was sorted into models with parameter combinations leading to results that did or did not meet ASHRAE-14 criteria for calibration fit. As discussed earlier, the total number of possible parameter combinations is nearly infinite. Larger simulation batches lead to a greater number of potential calibration solutions.

Increasing the sampling size comes at great computational expense; this being especially true for complex building models simulated on the EnergyPlus simulation platform. For this work, a large enough set of possible solutions was desired so as to make possible a cluster analysis and parameter shift analysis described in a later section. A solution set no smaller than 100 unique calibrated solutions would suffice to produce a significant set of clusters of models grouped by similarity for uncertain parameter input values. For any bounded grid search statistical calibration problem, the percentage of calibrated solutions realized from a set of sample runs will depend on how close the initial

model is to the desired calibrated result, how well parameter ranges are selected, and how tightly they are bound. To achieve the desired quantity of unique calibrated solutions, LHS sample batches were run in larger and larger sizes. It was found that due to the memory limits of OpenStudio Server, and large file sizes (> 40mb) created by each sample run, only batches up to 2800 samples in size were possible. To work within this limitation it was necessary to combine several smaller batches of sample runs to achieve the necessary results. It turns out that this method of combining small batches also increased the number of *unique* calibrated solutions by a higher amount than projected with simply increasing the batch size. This is due to the nature of LHS sampling, as discussed above. A summary of results is given in Figure 41 and Table 10 below.

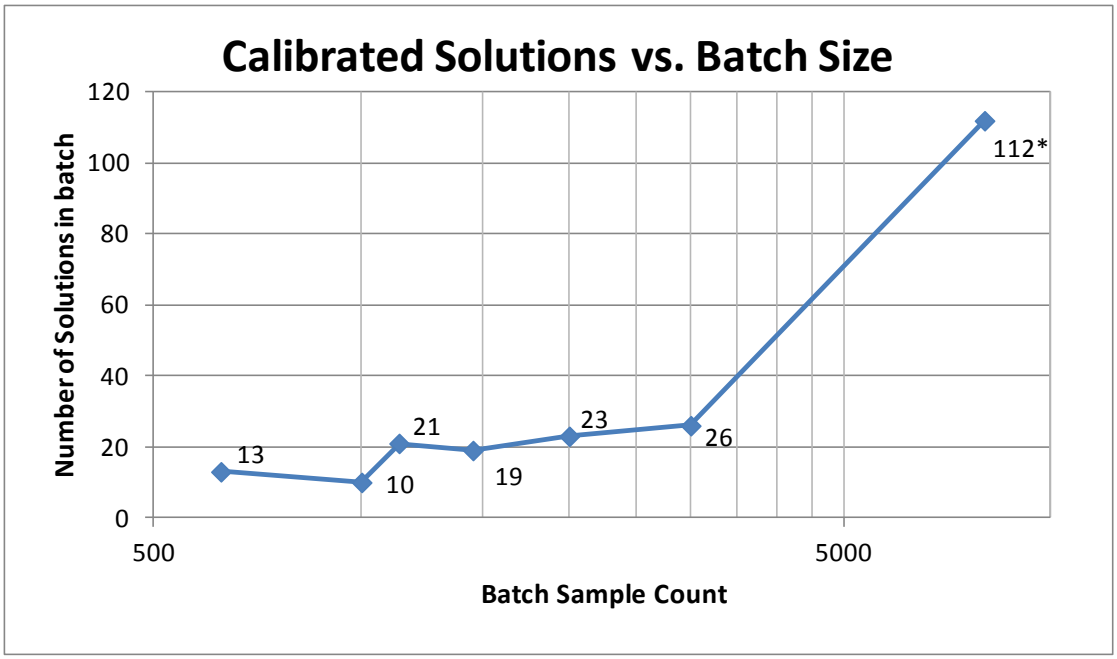


Figure 41: The number of calibrated solutions showed little increase with increasing batch size.

**112: this was a combination of all smaller batches*

Table 10: Batch size vs. number of calibrated solutions

Batch Size:	625	750	1133	1450	1667	2797	8422
Calibrated Solutions:	13	10	21	19	23	26	112

6.4 Model Calibration Results

From among the large set of samples, calibration solutions were uncovered with the metrics described above. The LHS sampling calibration process yielded 112 out of 8422 total samples that fell within the range of acceptability in terms of goodness of fit metrics. This fraction of total samples illustrates how difficult it is to match modeled data to observed data. Even after several hundred hours of effort were put into developing the detailed simulation model and expert judgment was applied to parameter uncertainty characterization, fewer than 2% of quasi-random samples resulted in a calibration solution. This is even more sobering when considering that calibration to ASHRAE-14 guidelines is based on annualized average error metrics, and a well calibrated model for an Energy Signal Tool should have consistent *monthly*, or better, *weekly* calibration fit to measured data. This will be discussed in more detail later on.

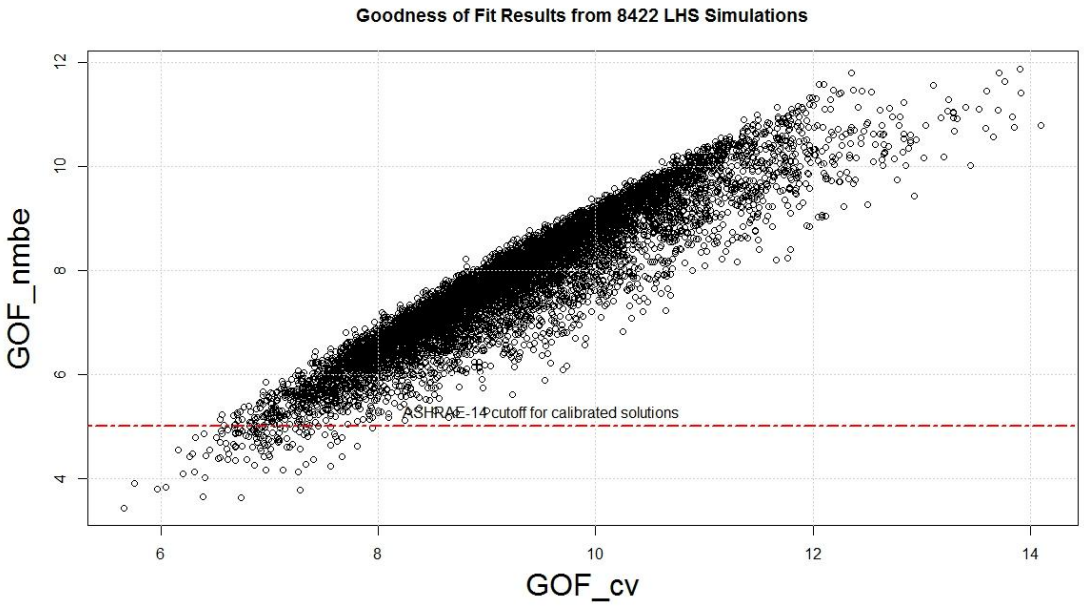


Figure 42: GOF results for primary calibration

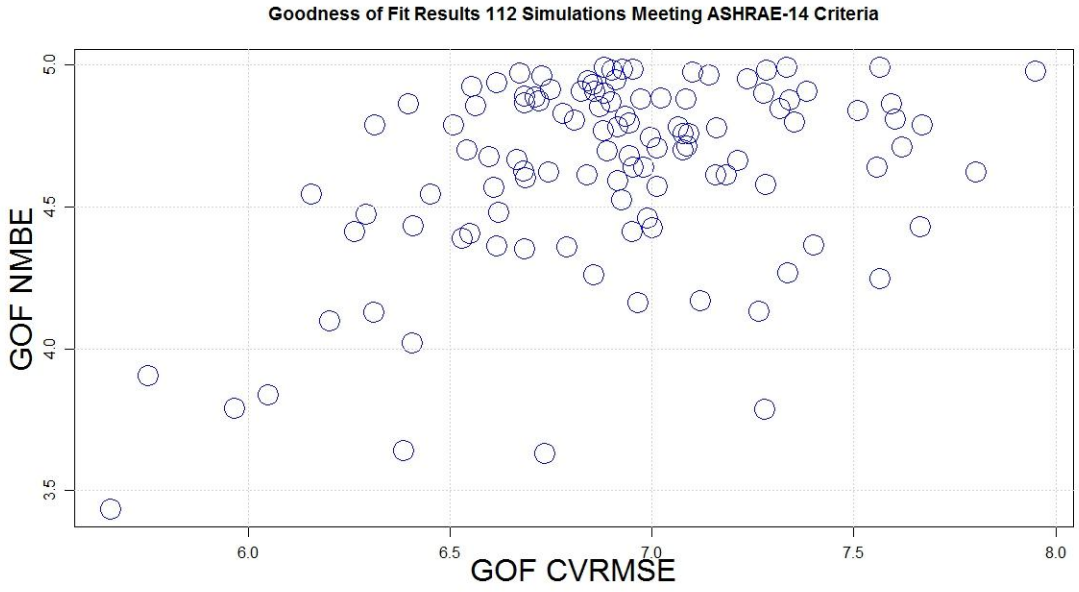


Figure 43: A closer look at calibrated solutions. Many distinct points indicate good sampling.

Reddy (2011) recommends proceeding with all acceptable models to create an ensemble of models, which are then run in parallel to create a distribution of possible results. However, in the parameterized approach to uncertainty analysis, we shall proceed by the set of possible solutions for generating a "posterior" distribution of each significant

parameter range. This distribution range produced by parameter sampling will in fact be wider than what would be generated by an ensemble of model solutions. Being able to generate hundreds or thousands of samples is more statistically robust than using multiple "calibrated" models.

6.4.1 Cluster Analysis Filtering for the Best Distribution

The uncertain parameter values generated by quasi-random LHS sampling are what make the model samples for calibration unique. A cluster analysis can be used to uncover *which* parameter value groupings were most likely to lead to calibration solutions, and extract more probable parameter uncertainty distributions from the largest group. As sensitivity analysis demonstrated, there are fewer than ten parameters that contribute significantly to model outcome in terms of total electricity or natural gas use. There are several possible combinations of shifts in these key parameters that can lead to a calibration solution. For example, a model that initially under predicts natural gas use could be corrected by increasing the infiltration rate, or also by decreasing gas burner efficiency if both parameters are uncertain. Due to this, the set of calibrated solutions is likely to contain natural groupings of models that share traits of having similar predicted values for one or more significant parameters.

Cluster groupings are not meant to contain identical models; rather, they are made up of models with close *proximity* in terms of parameter values. Some groups will be larger than others. Hierarchical clustering of the parameter values within the array of solution vectors is the method utilized in this work to identify the most plausible family of calibration solutions. It is proposed here that the largest group shall be considered most plausible set of solutions. The end goal of clustering calibrated solutions in this work is to extract information leading to an update of parameter uncertainty characterization.

Clustering can be used to partition a data set of n models into groups based on their similarity in terms of a set of k characteristics. Gao and Malkawi (2014) demonstrated that

clustering methods can be used for determining reasonable grounds of similarity among buildings to conduct performance benchmarking comparisons. The authors developed clusters of buildings by type, and then grouped these according to seven characteristic features. They chose the centroid (medoid) of each cluster to the *representative* for each group of buildings. In this work, the medoid of the largest cluster is examined for each parameter to verify the technical plausibility of the solution set it represents.

For calibrated building energy model solutions with a *k*-dimensional space of uncertainty, a combination of divisive hierarchical clustering and partitional clustering of the solutions' parameter spaces on the basis of similarity is appropriate (Reddy, 2011). Hierarchical clustering is useful for visualization and validation that the set of calibrated solutions is comprised of unique samples. It was elected to include all 18 potentially significant parameters in the *k*-dimensional space for the clustering analysis to describe uncertainty in each monitored end use. The clustering process was initiated with the use of tree visualization (dendrogram) from the "hclust" function within the 'cluster' package in the R environment (Maechler *et al.*, 2014). This algorithm makes use of the Euclidian distance metric for comparing the proximity of the parameter values of each calibrated model. A dendrogram visualization is constructed by starting with set of parameter values from individual models and moving upwards by grouping and connecting models with the closest distances between them. The completed dendrogram reveals a visualization of how many groups exist, their sizes, and what levels the vectors of parameter values representing each calibrated solution fit into (see example below in Figure 44).

Partitional clustering was used to decompose the larger set of calibrated solutions into related groups. Partitioning Around Means (PAM) is a partitional clustering algorithm, which determines a centrally located representative in each cluster (Halkidi *et al.*, 2001). CLARA (Clustering Large Applications) is an application of PAM on samples from a dataset (Halkidi *et al*., 2001). The function "clara", within the package "cluster" in R, is used to

identify the medoids of each cluster (Maechler *et al.,* 2014). A cluster medoid is the representative model of that cluster that has minimum dissimilarity to all other members of the cluster. Calling "clara" in R requires the user to input a quantity of significant clusters, *nc*, (identified visually from the dendrogram, and then checked for validity), and will then return a vector of *nc* cluster medoid (model tags). The largest cluster of models is taken, and the medoid of this cluster is inspected for plausibility. In small the example below, visual inspection revealed three main branches. The R function "clara" confirms the members of these branches, and finds that trajectory '4' is the medoid of the largest cluster.

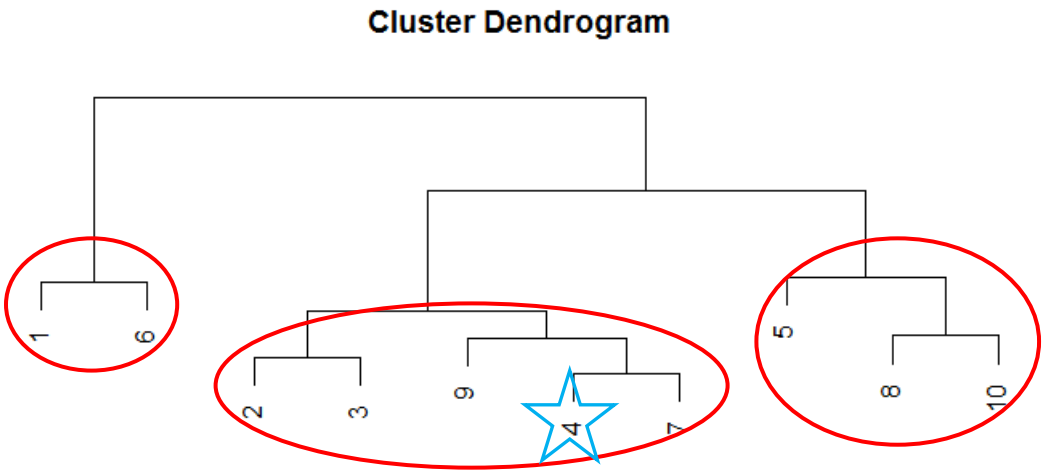


Figure 44: A simple hierarchical clustering example results diagram produced in R.

See appendix for a complete dendrogram of cluster results for the 112 calibrated solutions. The lines of natural division appear to separate the data into six crisp clusters.

6.4.2 Cluster Validation

The main challenge is processing the results of data clustering can be determining the optimal number of groups for the dataset. The partitional clustering algorithm results in clusters classified as "crisp" clusters. Visualization of the clustering results is a crucial first step to validation of groupings. The optimal grouping of clusters is one that best fits "the inherent partitions of the data set" (Halkidi *et al.*, 2001). As mentioned above, visual methods are used to identify the inherent partitions in the dendrogram results. According

to Halkidi *et al.*(2001), there are two basic criteria used in validating these visual selections of cluster partitions for crisp clustering:

1. Compactness: *is variance within each cluster minimized?*
2. Separation: *are clusters spaced apart sufficiently?*

Where both of the above can be based on a parametric analysis of number of clusters. For purposes of simplification, the metric for assessing the first criterion in this paper was simply the dissimilarity of the members internal to each cluster. The mean value of dissimilarity was then taken for all clusters. Figure 46 below displays the results of this parametric analysis for compactness. Mean dissimilarity is calculated in Equation 13 as:

Equation 13: Mean dissimilarity calculation for cluster partition analysis

$$For\ \left(k = j\frac{j-1}{2}\right);\ \sum_{i=1}^{nc}\frac{\frac{1}{2}\sum_{j=1}^{N}\left|X_j - Y_j\right|}{k}$$

Where *nc* is the number of clusters, *N* is the number of elements in the cluster, and *X* and *Y* are two elements being compared. Halkidi *et al.*(2001) state that the optimal solution in the cluster compactness analysis is found at the hinge of the elbow between two slopes, as demonstrated in Figure 46 below.

For measuring the second criteria, the Manhattan distance between representative models (medoids) of each cluster can be compared. A sensitivity analysis which compared the number of cluster divisions to the mean of the distances between all medoids of each cluster. This is accomplished with Equation 14 below.

Equation 14: Mean distance between cluster medoids for cluster partition analysis

$$For\ (i\ in\ 2{:}20)\left\{Dist_{mean}[i] = \frac{\sum Dist(medoids)}{i\frac{i-1}{2}}\right\}$$

Where the function $i\frac{i-1}{2}$ describes the number of medoid distances compared from among all medoids. The optimal number of clusters will have the maximum mean distance

between them (Halkidi *et al.*, 2001). The results were plotted for the set of 112 calibration solutions as follows in Figure 45:

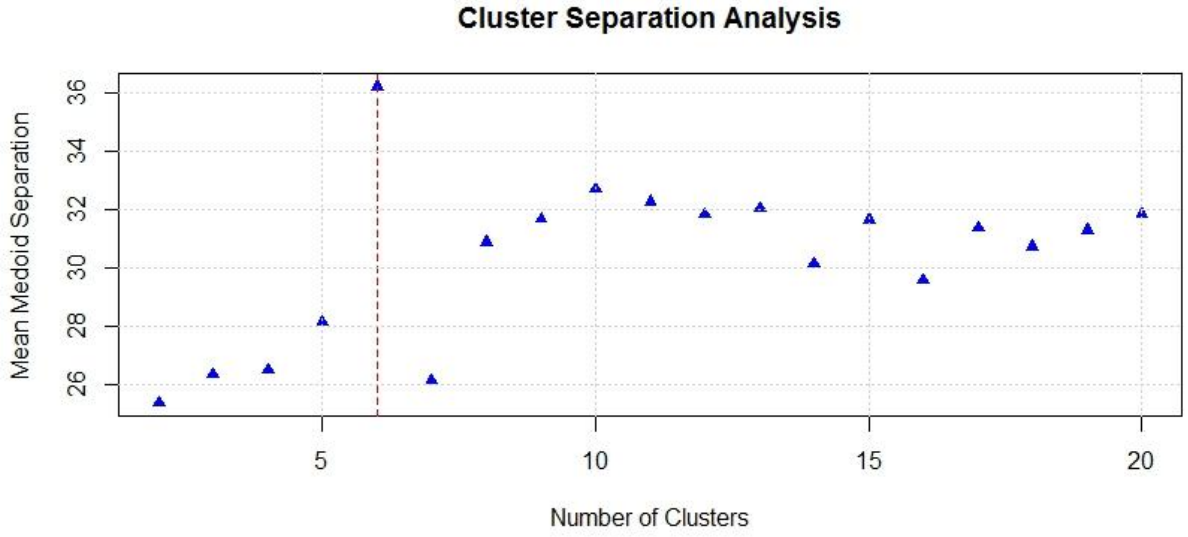


Figure 45: Cluster separation criteria metric revealing optimal number of clusters for medoids distance

The very clear solution of six cluster divisions given above in Figure 45 agrees with the crisp visual results shown in the appendix, which means that the visual solution of six partitions is valid.

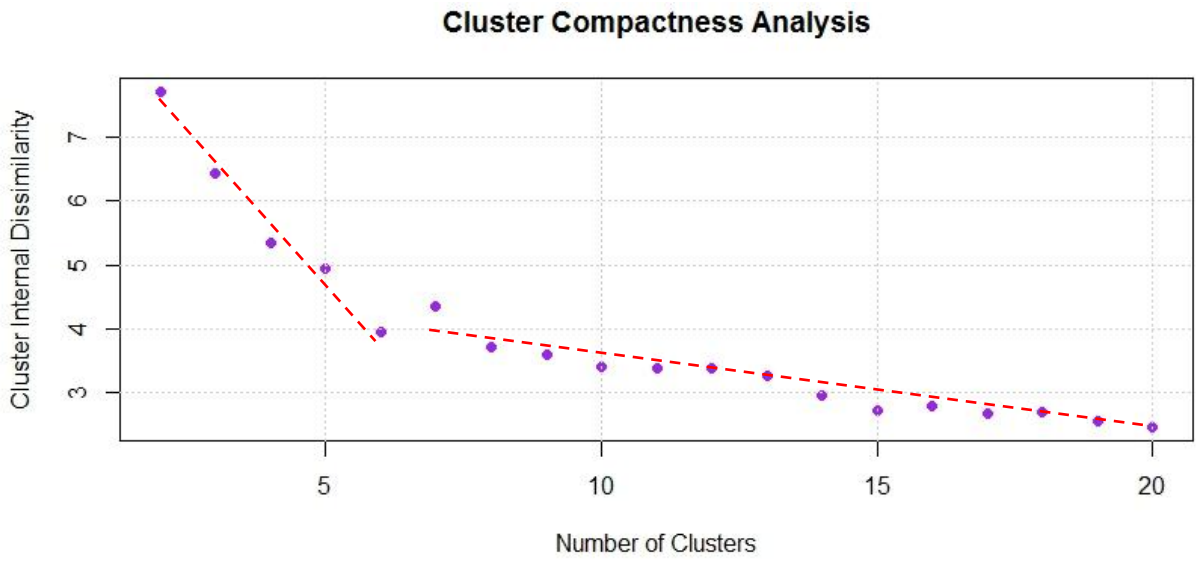


Figure 46: Cluster compactness metric (~ variance analysis) results reveals decreasing dissimilarity and an optimal elbow minimum value.

The Compactness analysis yielded an "elbow" inflection point at the cluster quantity (*nc*) of six. Thus, all validation tests confirmed the optimal number of clusters to be six.

6.4.3 Testing for Parameter Correlation and Multicollinearity

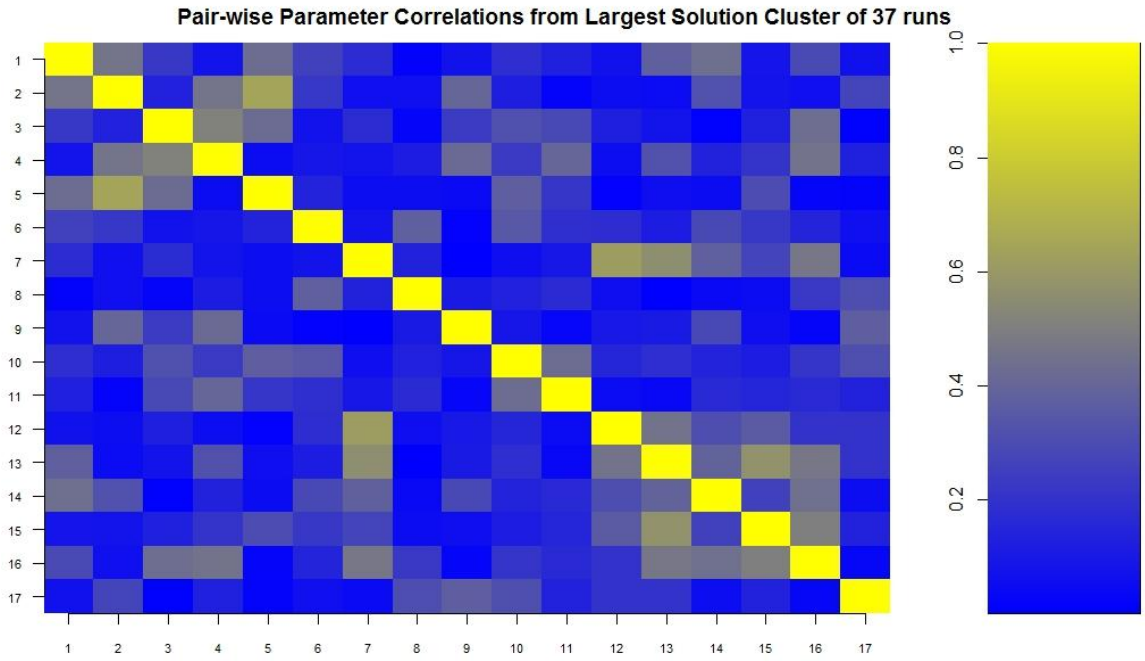


Figure 47: Plot of pair-wise parametric correlation among the parameters within the largest solution cluster

The figure above explores all pairs of parameters for correlation. (The plot is mirrored for plotting simplicity) There exists a moderate amount of correlation amongst some variables from the potential solution set. For screening purposes, a rule of thumb can be used to identify potentially correlated parameters, where a Pearson correlation coefficient of greater than 0.7 indicates a potentially moderate to strongly correlated pair (Reddy, 2011). When parameters are found to be correlated in the larger set of sampling, it is important not to separate them into sets of uncertain and frozen variables when reducing parameter space with sensitivity analysis. This would fail to incorporate their structure into the further uncertainty analysis used by the Energy Signal Tool. In this work, no significant parameter pairs were found to have Pearson correlations greater than 0.7

Another meaningful visualization of parameter behavior is a scatter plot comparison of output variables to check for visual covariance or colinearity. The figure below shows a plot of those variable pairings which were found to have Pearson correlations of greater than 0.45, as evidenced above in Figure 47.

In reference to the heat map of correlations (Figure 47 above), the figures plotted were:

{2:1, 2:4, 2:5}; {3:4}; {4:16}; {7:12, 7:13, 7:15}; {12:13}; {13:15, 13:16}; {15:16}

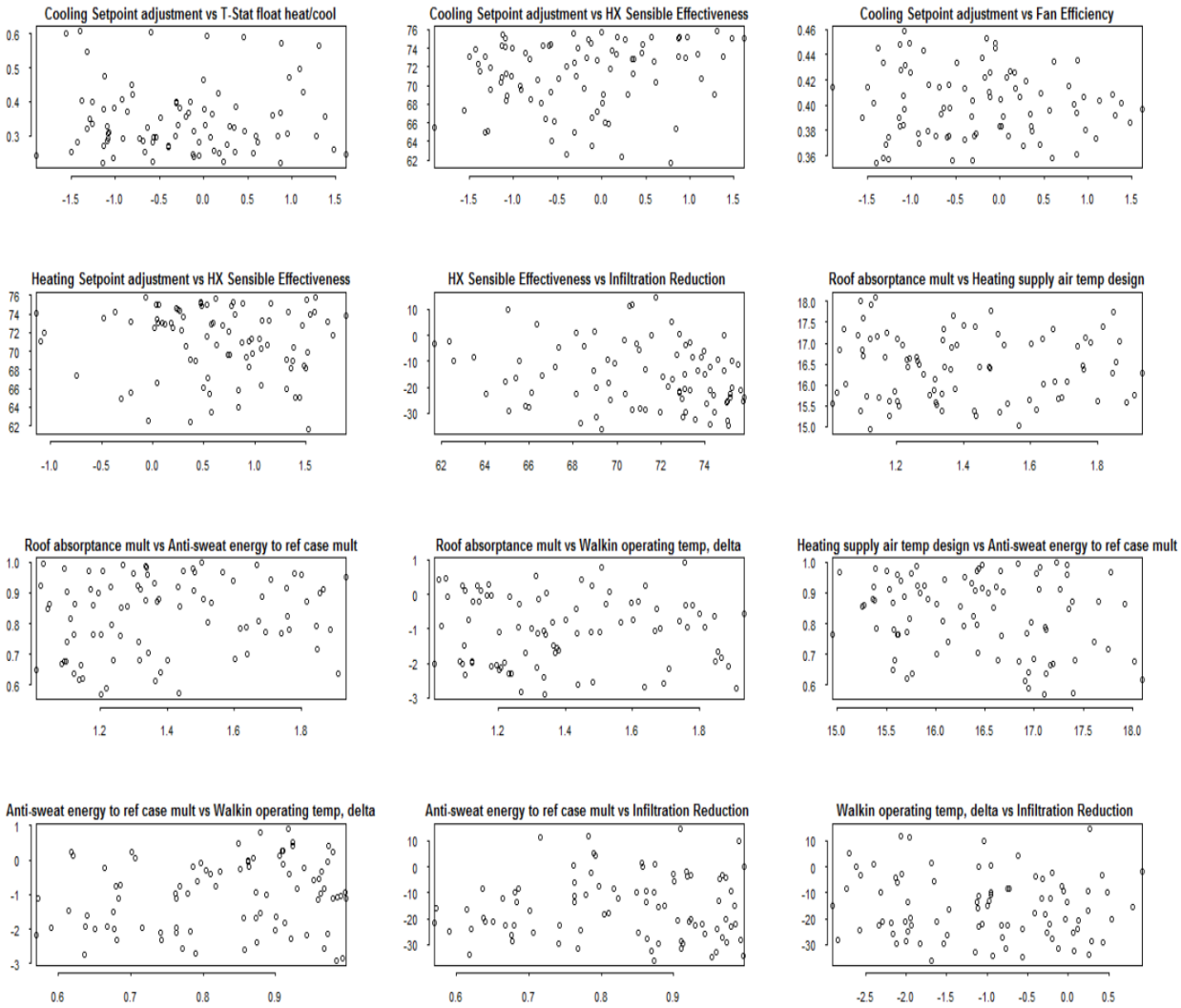


Figure 48: Scatter Plots of Parameters in largest solution cluster to examine for co-linearity

An examination of the above scatter plots does not show any especially prominent collinearities between variables. The distribution patterns are almost purely random; simply a translation of the sampled distribution for each parameter. There is some slight evidence in the relationship of the heat exchanger effectiveness and its impact on corresponding infiltration rate in the calibrated model solutions.

6.4.4 Checking for Non-Linear Behavior in the Model

Using data generated by a large number of LHS samples, and visualized in scatter plot diagrams, can reveal visual patterns in the relationships between input parameters and objective functions of interest. When samples are run by sampling parameters individually, a clear image of the effect of the entire range of this parameter on an objective function is visualized. This is similar to the OAT parameter prescreening test. The slope of the relationship is a rough indicator of how significant this variable is to the overall model. The advantage of the scatter plots is the ability to visualize non-linearities in the model output relative to some of the parameter uncertainty ranges. The disadvantage compared to the OAT SA process is a lack of ability to quantify parameter significance. Correlation plots of select parameters and end-uses can be found in the appendix.

Comparing correlation of each parameter to objective results can be a way to understand the non-linear response of the model to range portions of certain parameters (Saltelli *et al.*, 2000). It is important to note non-linear behavior of the model in response to certain ranges of parameter uncertainty. For example, refrigeration energy responds much more quickly to incremental changes at the high-range of minimum condensing temperature than at lower levels in the assumed range. Another example of note is that relationship between total energy use vs. heating system sizing factor exhibits two slopes – split at the assumed heating sizing factor of 120%. A lower than expected heating system sizing factor has a greater effect on the objective result. It is important to consider these

correlative relationships when drawing boundaries of "posterior" parameter distributions resulting from the calibrated solution sub-set.

6.4.5 Parameter Distributions Arising from the Calibrated Solution Set

The calibration process generates another set of parameter values; those which make up the models in the set of models meeting calibration criteria. From the vectors of parameters comprising each model, most probable distributions of parameter values can be derived. Bucking *et al.* (2014) demonstrated that more appropriately fitted PDFs, which characterize uncertain parameters, can be extracted from the set of results that demonstrate compliance with the building performance objectives. This allowed for the identification of parameter limits and most probable values. They used the following process:

1. Initial characterization of parameters was done by identifying a probable range and a discrete number of sampling steps in that range
2. Selecting the parameter combinations from the models that equaled or exceeded the objective of NZE performance
3. Counting the number of occurrences of each discretized interval in each parameter
4. Normalizing the sum of these counts to equal one and fitting with a kernel density function

In a similar manner, the work presented here proposes the extraction of parameter distributions from the sub-set of the *largest cluster* of sufficiently calibrated model solutions. By assembling the sampled values of the parameters in these models, it is possible to update the prior assumptions of significant parameter distributions.

Figure 49 and Figure 50 below illustrate the comparisons that can be made between prior assumed and posterior extracted parameter values. As discussed in Section 5.3.1 above, the relative shift in the distributions has little significance other than to inform the user how good the initial distribution assumptions were. These updated distributions can

be used to generate the expected energy use distributions for each end use. This will be the new benchmark reference used by the Energy Signal Tool, given the tests described in the following section return satisfactory results. The first figure below gives a sample of what prior assumed (grey) and posterior observed (various colors) will look like. Vertical dashed lines show the static values in the representative (calibrated) model.

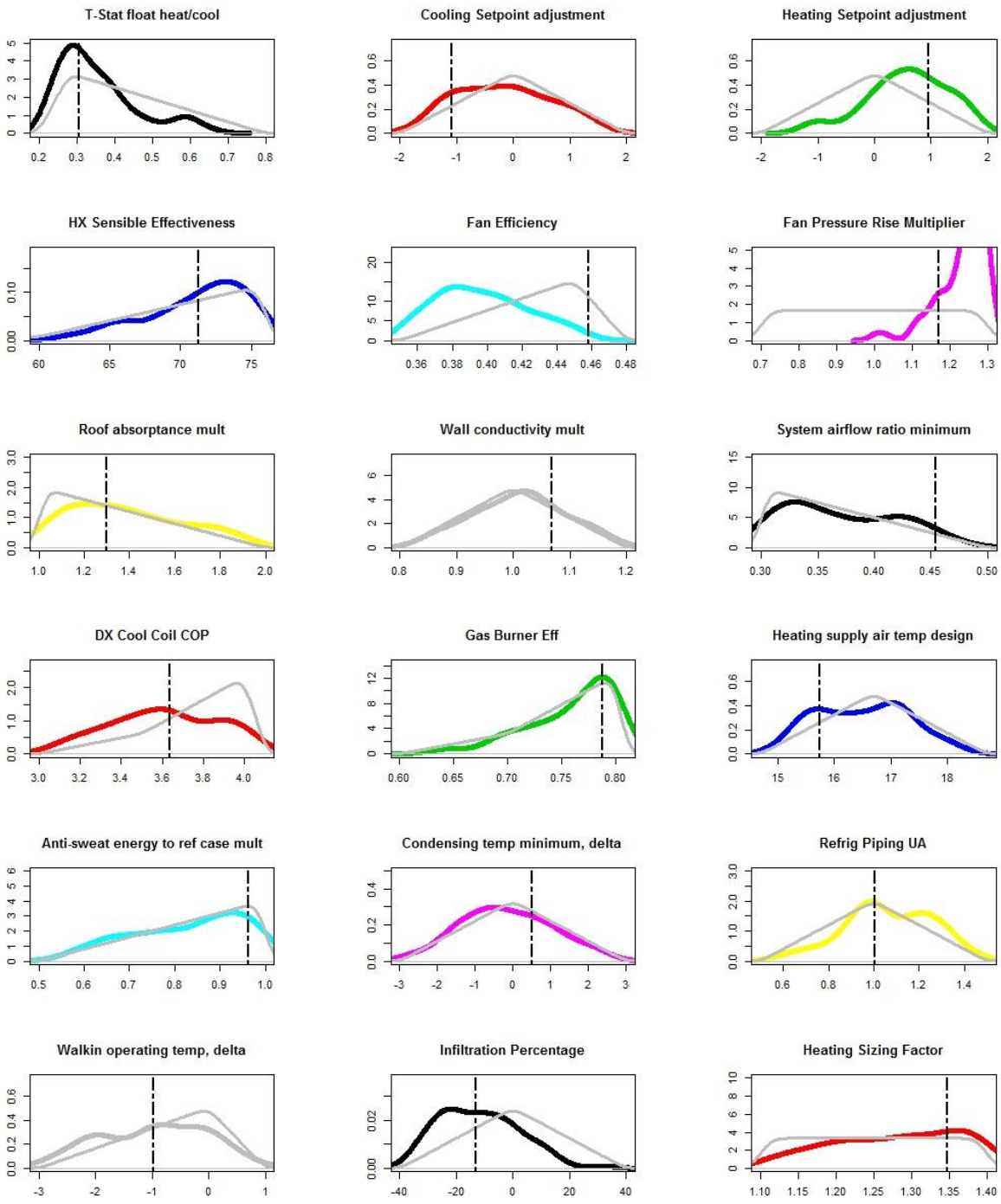


Figure 49: Posterior vs. prior parameter distributions, with posteriors from *all models with acceptable calibration results*, with representative model values in dotted line.

Figure 50 below compares the distributions observed in the entire set of calibrated solutions to those found in only the largest cluster. In this case, there is quite close

agreement between the two, indicating that the largest cluster is also still a good representative of the entire set.

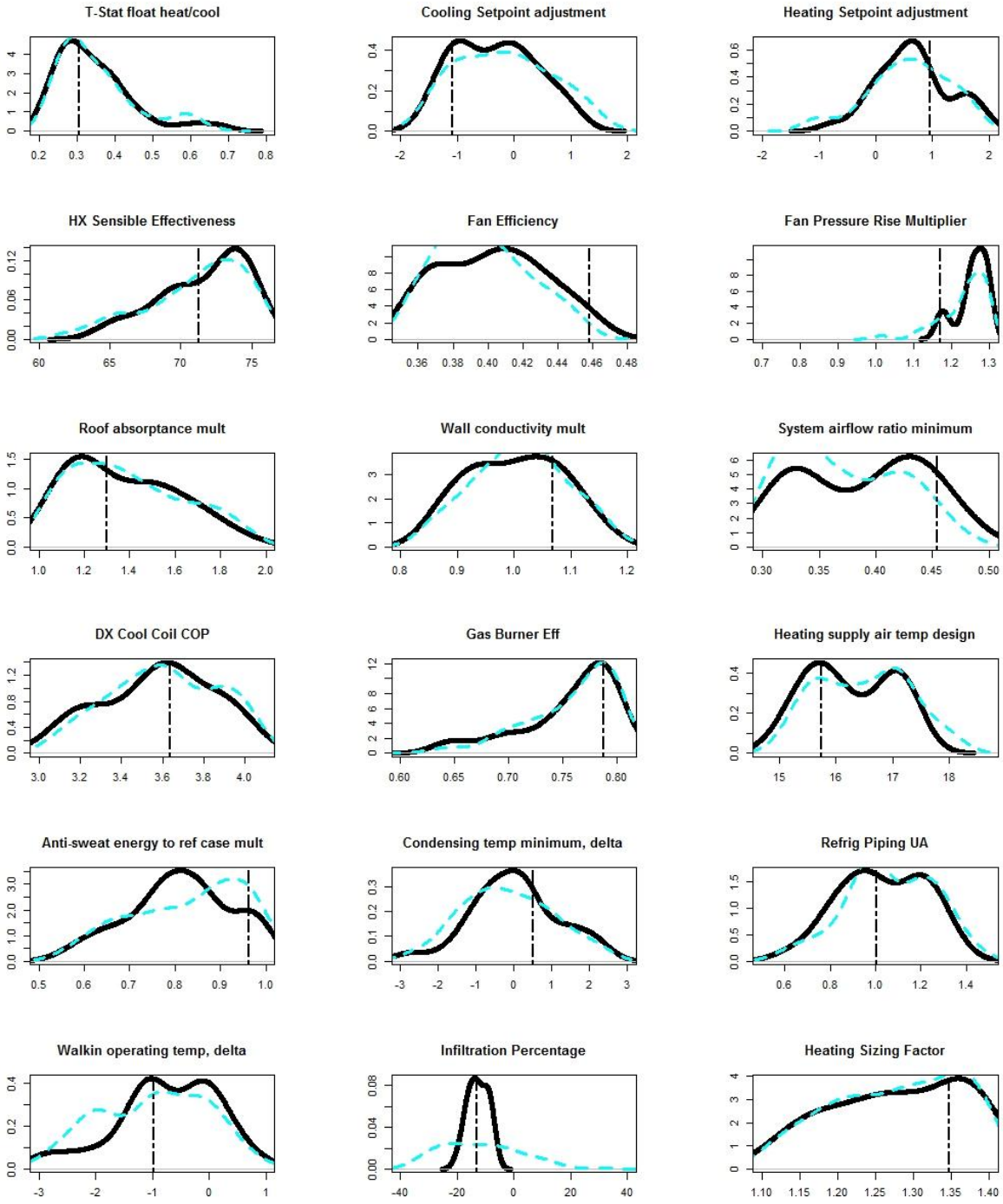


Figure 50: Comparison of posteriors from all models with acceptable calibration results (dotted line), to just those posteriors within the largest solution cluster (solid line).

With these new distributions that arose from those models forming the calibration solutions, it was possible to take new LHS model samples and re-define the expected ranges of consumption for each end use. The comparison of expected annual consumption resulting from old and new uncertain parameter distributions is shown below in Figure 51.

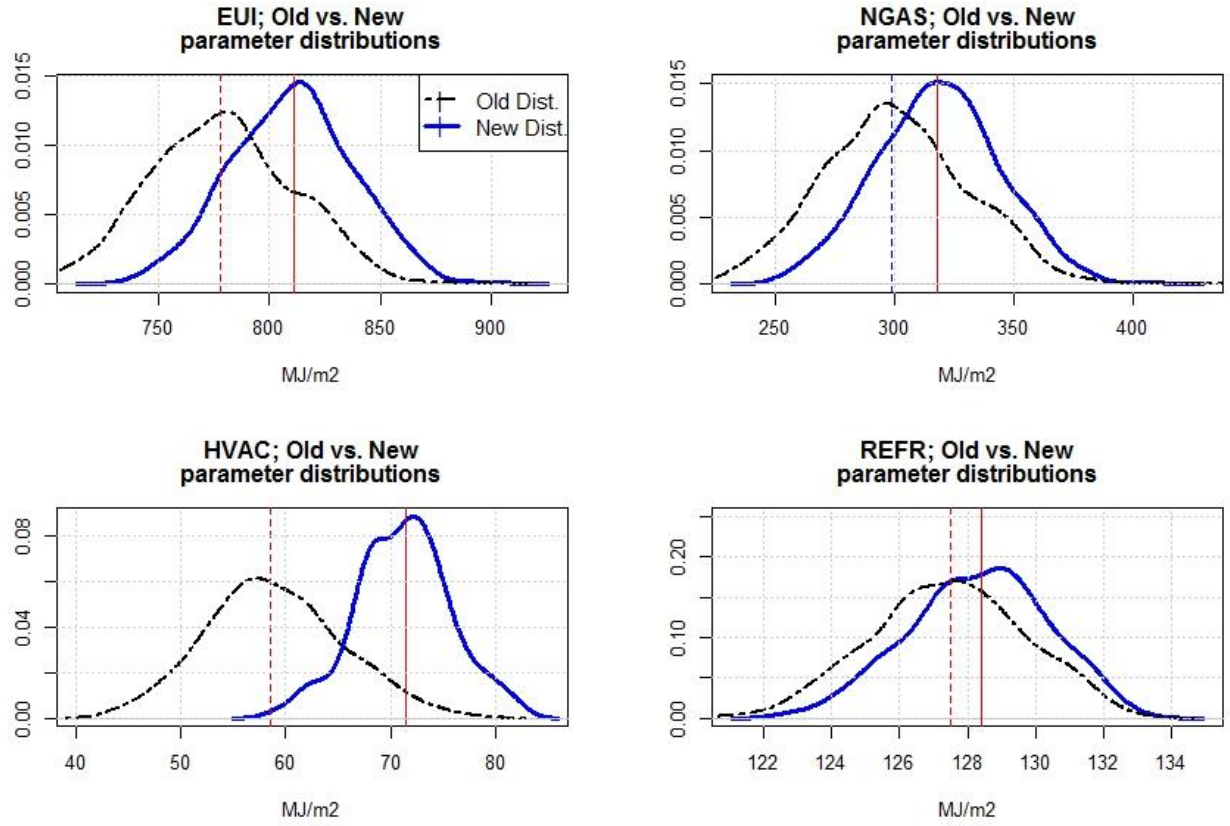


Figure 51: Comparison of the annual end-use distribution results from old parameter distributions and new parameter distributions derived from calibration solutions. Sample medians are shown with a vertical line.

This comparison shows an upward shift in the expected ranges of distribution. This reflects the difference between the energy use predicted by the un-calibrated model and that measured in the calibration period utility bills. Visually, we can see that there is very little change in observed variance in the end uses.

We can now examine the representative model (medoid of largest cluster) parameter values for reasonableness. A summary is given below in Table 11.

Table 11: Representative model parameter values and justifications

Parameter	Final Calib. Value	Initial Est. value	Units	Justification
T-Stat float heat/cool	0.3	0.278	°C	Little change
Cooling Setpoint adjustment	-1.08	0	°C	Manual override down is common
Heating Setpoint adjustment	0.96	0	°C	Manual override up is common
HX Sensible Effectiveness	71.35	76	%	76% is rated at optimal test conditions
Fan Efficiency	45.5	45	%	Little change
Fan Pressure Rise Multiplier	1.17	1		Air supply path may not be installed according to plans
Roof absorptance mult	1.27	1		Roof has collected some dirt, which is normal over time
Wall conductivity mult	1.07	1		Little change
System airflow ratio minimum	45	30	%	*No great reason; distribution in Figure 37 above was random.
DX Cool Coil COP	3.62	4		DX COP could be reduced due to installation in hot area
Gas Burner Eff	78	80	%	Little change
Heating supply air temp design	15.8	16.7	°C	Slight error in duct temperature sensor
Anti-sweat energy to ref case mult	0.96	1		Little change
Condensing temp minimum, delta	0.44	0	°C	Little change
Refrigeration Suction Piping sum UA mult	1.12	1		Installed piping could be longer than specified on plans
Walkin operating temp, delta	-1	0	°C	Actual refrigeration temperatures are likely to be
Infiltration Percentage	116.28	100	%	Infiltration is difficult to estimate
Heating Sizing Factor	1.35	1.2		Heating system sized based on uncertain loads

We have therefore accepted the parameter distribution arising from the largest calibration solution cluster as the representative model to form the basis of the Energy Signal Tool. Since the Energy Signal Tool relies on an *expected range* of end use consumption rather than a single point, the *representative* model is more of a culmination marker to the calibration process than it is useful. Regardless of parameter starting values,

the ESTool will still employ ongoing LHS sampling to produce consumption ranges, such as those in Figure 51 above.

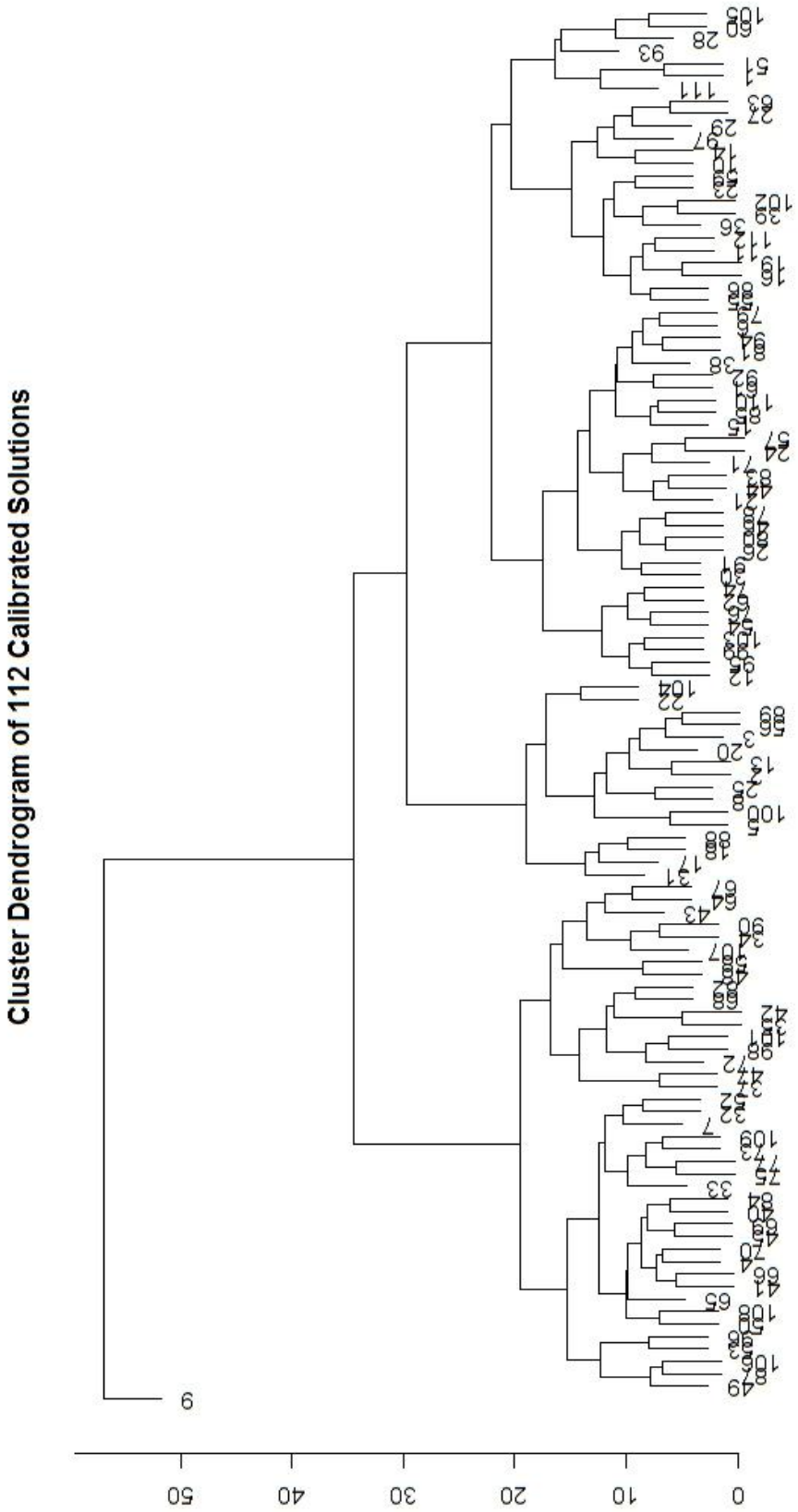


Figure 52: Resultant dendrogram of large LHS sample set

7 Energy Signal Tool Development

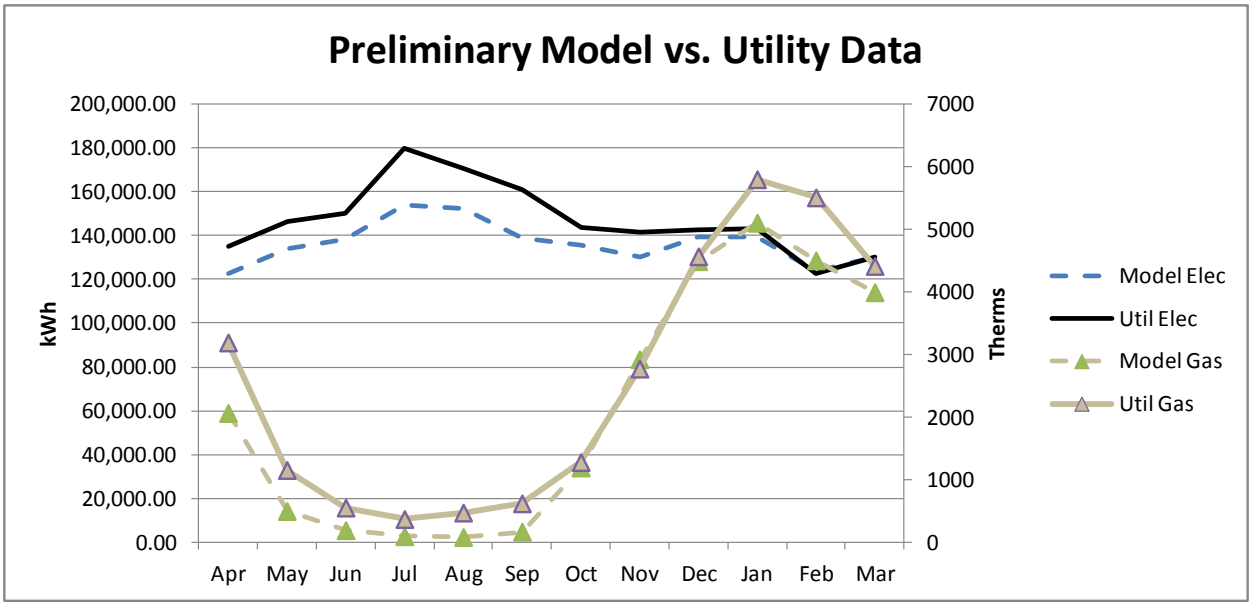
Facilities managers often make choices that affect energy use about how to operate their buildings. However, rather than waiting for utility bill analysis or feedback from building occupants, they need a more transparent view of how their actions affect energy use and building performance. Human understanding of vast amounts of complex metered data requires automated data simplification. The decision analysis component of the Energy Signal Tool transforms metered data and complex model output into a simplified visual output that supports energy-related decision making. This output cannot replace the judgment of the FM, but will greatly aid in a complex multiple criteria decision-making process. Comparing monitored data to a range of reasonable expectations will help the FM determine when corrective action (such as repair or commissioning) is cost effective.

Typically, the model calibration process (as discussed in the previous section) is carried out using data from advance of the monitored period, which is known to be generated by a building operating as expected. The calibrated model is a valid representation of the building in its environment so long as it is run with actual weather data and is updated with any major changes to building equipment, operational strategies, or use schedules. As described previously, it is not the calibrated model itself but the parameter uncertainty distributions arising from the largest cluster of calibration solutions that are useful. However, the benchmark model must accurately describe the measured data from the building. Comparisons of modeled to measured data over the calibration period, assessed for each of the *sub-periods* (such as monthly or weekly spans) should yield no more than small errors for *each end use.* If these comparisons give bad results, the energy signals generated in some or all periods will undoubtedly be made with an underlying bias. If the benchmark model is not sufficiently calibrated, more information must be obtained to improve it.

7.1 Benchmark Model Validation

The calibration process mentioned in the previous section is done to achieve a minimization in annual average monthly error, and variance in that monthly error. According to ASHRAE guideline14, variation in monthly error in *fuel use* modeled vs. predicted is allowed to *average* as much as 15%. This does not guarantee by any means that *end-use* error is minimized to these bounds. Even if it were, a 15% deviation would be a significant cost prediction error. It is desired that the Energy Signal Tool is capable of detecting faults which manifest in as little as 3 to 5% cost deviation in monitored end-uses. Therefore, it is necessary that the benchmark model be calibrated to higher standards than ASHRAE-14. With the calibration techniques available in this work, this was not a realistic goal. A sufficiently calibrated model would ideally result in end-use observations falling within modeled end-use distributions to produce all green signals, indicating operation as expected (refer to sections 7.5 and 8.2 below).

The following figures show the results of the representative model from the calibrated solution set as compared to the observed data with which the calibration was carried out. Particular attention should be paid to comparisons of monthly end-use predictions, and the magnitude of error relative to monthly observations.



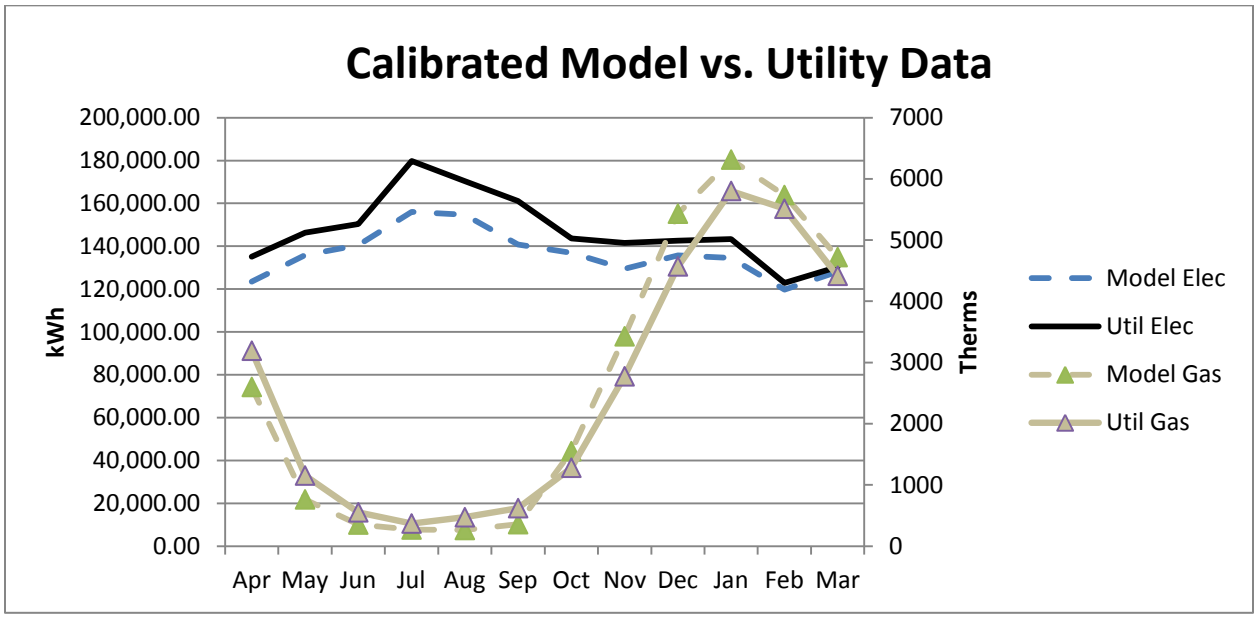


Figure 53 a,b: Monthly natural gas and electricity calibration comparison before and after sampling

Figure 53 shows that the calibration yielded slight improvement in bringing the model description closer to that suggested by the monthly utility data. Almost no change was affected in the electricity use profile, owing to the fact that the combination of the changes in the solution set parameters had little effect.

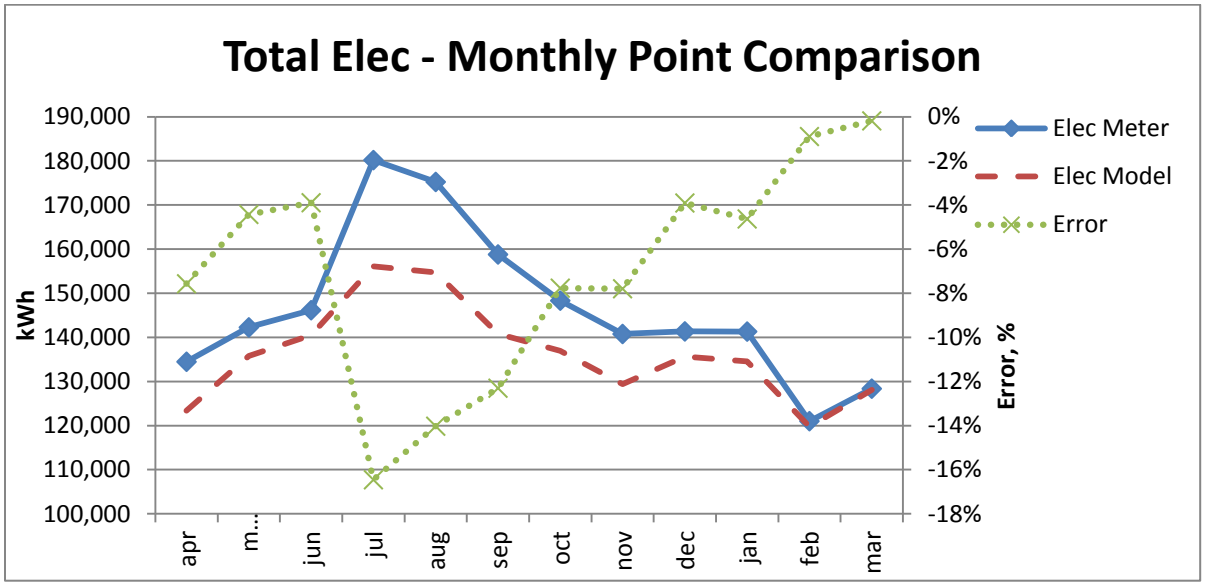


Figure 54: Monthly point comparison of measured and modeled total electricity use

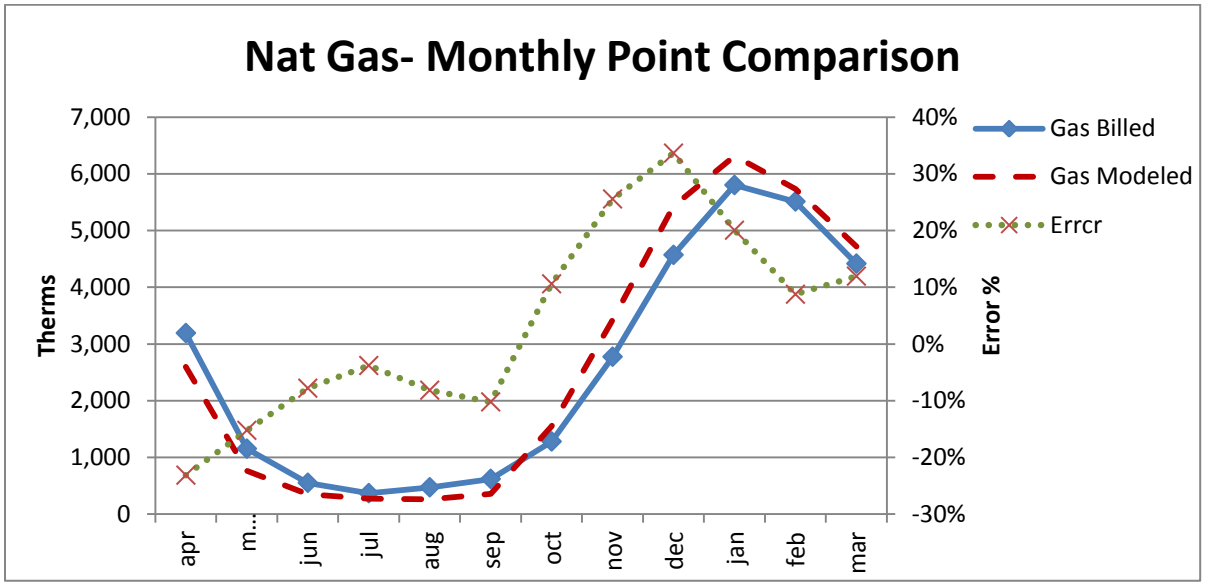


Figure 55: Monthly point comparison of measured and modeled natural gas energy use

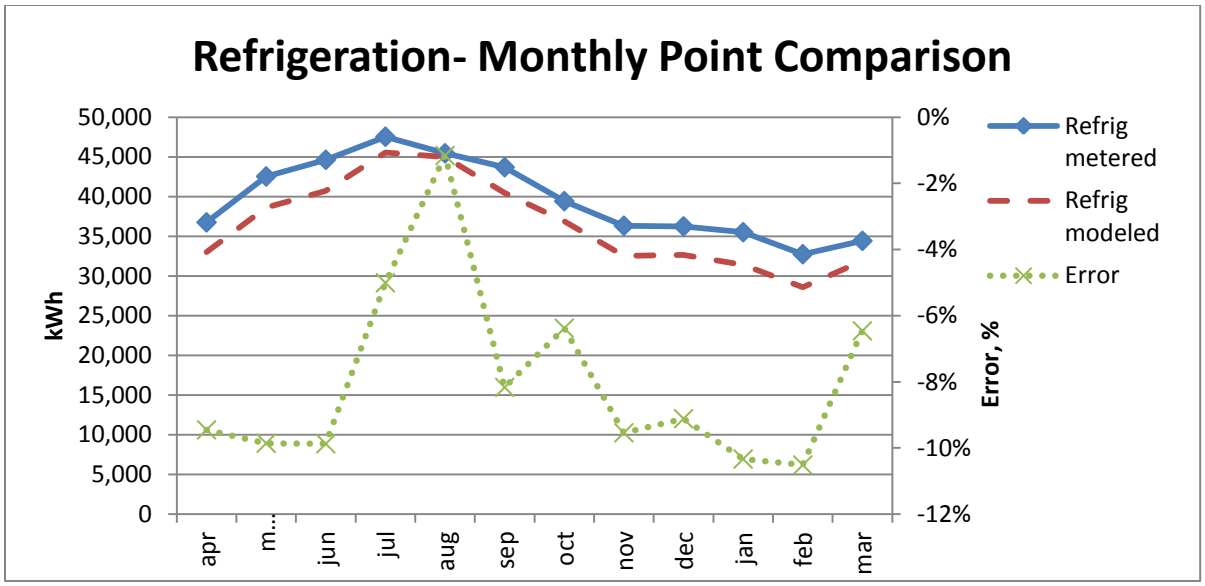


Figure 56: Monthly point comparison of measured and modeled refrigeration energy use

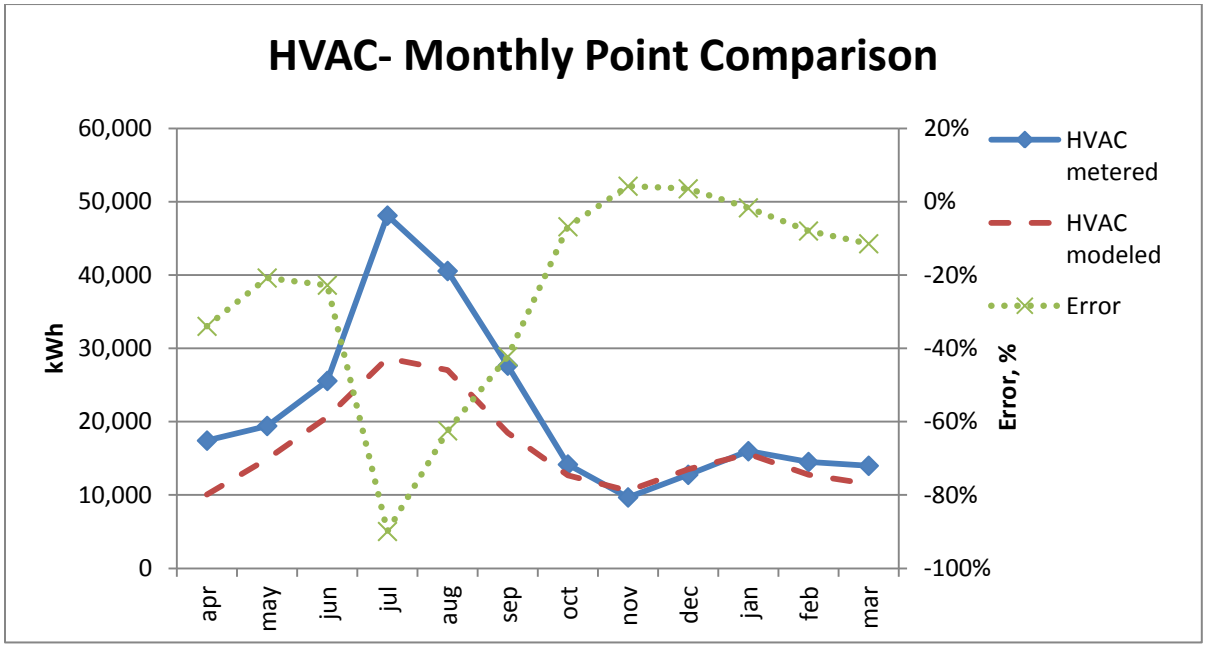


Figure 57: Comparison of modeled HVAC energy use to actual HVAC energy use over the calibration period

Figure 55 and Figure 57 above especially show that the case study model is not acceptably calibrated to several end-uses of interest. Natural gas consumption has an error associated with some seasonal shift not modeled. HVAC energy use clearly has an error related to cooling or dehumidification energy. Great efforts were made by the author to manually correct the model, but with little success. Additionally, there was little incentive

to calibrate the model perfectly when it was discovered that the metered end-use data was severely flawed and the retail partner was unavailable to give input (this is discussed below). If any of the sub-metered data can be trusted, the root cause of the error in HVAC energy is in the modeling of the dehumidification system. In the way the physical model was configured in EnergyPlus/OpenStudio, it was not possible to model a desiccant system which produced this usage profile.

The Energy Signal Tool is designed to report when a building is experiencing an energy-related fault. Therefore, benchmark model calibration should be done to data from a building that is free of major faults. Having relative errors that are consistently greater than a desired risk tolerance threshold is not acceptable. Although the Energy Signal Tool does not operate on a point-to-point basis, the results in the figures above would lead to an excessively skewed portrayal of the expected range as a context for observed consumption. This is because the entire probable range of consumption would be shifted in one direction (see Figure 59 below) and the mean expected consumption would not lie near the observed value. When a representative model fails the test of calibration acceptability, there are several possible courses of action to take:

1. Test representative models from other calibration solution clusters for their point-to-point comparisons with measured data.
2. Modify (extend) the bounds of parameter uncertainty and re-do statistical calibration with the hope of discovering a new unique combination of parameter values that meet the validation criteria.
3. Examine the comparison results to possibly uncover deficiencies in the model description. Collect missing information and re-do heuristic calibration. Then re-do statistical calibration.

The conclusion drawn from the above evidence is that statistical calibration alone is not sufficient to generate a baseline model for use in the Energy Signal Tool. Indeed, Figure

53 a, b above shows little improvement in fit to the measured monthly utility data. Statistical calibration only samples the model space defined by a set of uncertain parameters. This process may likely fail to eliminate an underlying model bias stemming from errors in monthly energy end-use consumption. In this work, it was not possible to write scripts to sample for things like seasonal or monthly variation in schedules, behavior patterns, or building set points. Using statistical methods to search through schedule variation or other possibilities is unwise considering the over-parameterized and under-determined nature of detailed building simulation models. To support an Energy Signal Tool in the form presented here, building stakeholders with additional knowledge must be involved in the model construction process. For example, operational schedules can be re-examined for monthly variations in usage. Additional studies can be done to determine if there are any seasonal factors affecting infiltration rates, such as customer volume, door settings, or window opening habits. Depending on the building type, all of this additional calibration work may take more effort than it is worth. Clearly one who is presenting the opportunity for such work to a client would need some idea in advance of the savings potential.

A decision was made to shift the focus of this work away from testing the Energy Signal Tool on the case study building in favor of more rigorous testing of the theory using synthetic data. A brief summary of the barriers in additional heuristic calibration steps needed to achieve a well-calibrated implementation of the Energy Signal Tool at the case study building is given below:

1. There were periods of missing or clearly erroneous data reporting delivered by the energy tracking service. This may have been due to some of the factors noted above in the section covering sub-metering data uncertainty, but nothing could be verified. This included VAV reheat data, where it was possible to make a regression model to

fill the gaps. See appendix for details. With additional time to apply advanced data repair techniques, it might have been possible to fix these errors.

2. Inconsistent operation of major equipment, especially the dehumidification unit and the VAV reheat, led to the inability to reconcile one year of data with the next. It is likely that manual overrides were applied to the BMS operational strategies. For example, Figure 58 below shows that the dehumidification unit (DHU) operation changed significantly from the end of July to the end of September. Meanwhile outdoor air conditions remained warm and humid consistently throughout this period – warranting continued store dehumidification. There exists no 12-month period in the observed data set where one can be certain of proper operation in the case-study building.

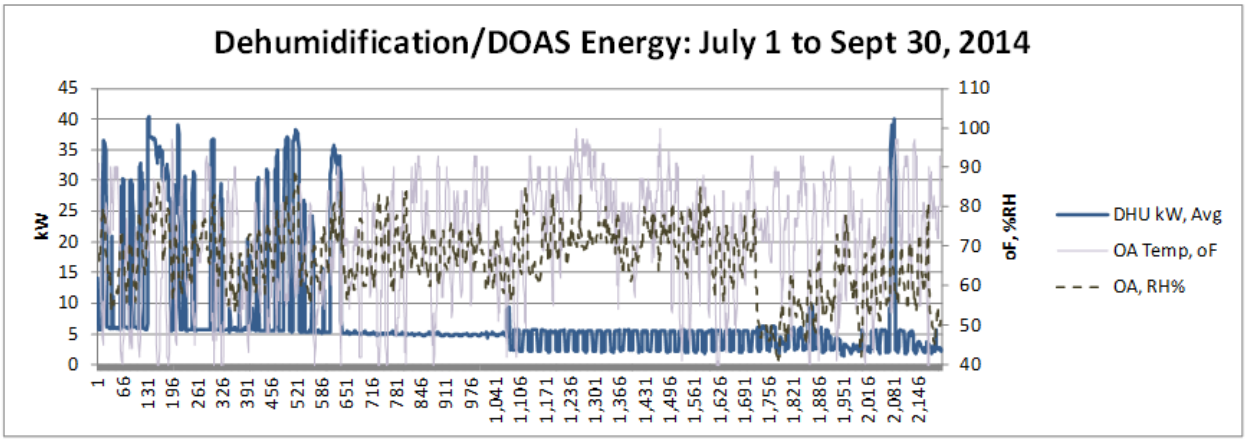


Figure 58: Hourly sub-metered data clearly demonstrates inconsistent operational strategies of the dehumidification unit

3. The retail partner was unavailable to answer questions about changes in operation strategy, or to verify proper installation of sub-metering equipment.
4. While calibration was successful by ASHRAE Guideline 14 standards (see above), the inconsistencies in operation, and lack of understanding of operational intent made it *virtually impossible* to calibrate the model to within low threshold tolerance

limits for each end use on a monthly (let alone weekly!) basis. Refer to Figure 55, Figure 56, and Figure 57 above.

It is reasonable to believe that these challenges are not unique to this work. Any project involving model-based performance assessment of a building could be faced with similar challenges of bad data, inconsistent operation, changes in facilities management, and lack of information. As we move into a time when sub-metering equipment is implemented with energy management in mind and the quality of data improves, some of these issues may diminish. However, these same problems are part of the justification behind the Energy Signal Tool concept. Due to uncertain performance risk variables, we will never be certain of how a building is supposed to operate, but a *probabilistic range of expectations* can offer some support as a lens to compare measured performance.

7.2 Energy Signal Tool Input

7.2.1 Overview

The Energy Signal Tool reports the action that is *most likely* to result in the lowest cost to the user for a given scenario of building performance. Recommendations are made in five levels of possible actions based on over or under-consumption. Each recommended action level is bounded by *risk tolerance thresholds.* It is important to think carefully about these thresholds – how much excess or under consumption can be tolerated – as well as what organizational objectives can be correlated to each energy end-use being monitored. Each organization may have a different tolerance level for risk in energy use deviation, environmental comfort, and other factors. At any given facility, there may be different acceptable risk tolerance thresholds for each end-use. The following section is a discussion of how risk tolerance thresholds are rationalized.

7.3 Defining Risk Tolerance Thresholds

This work addresses the issue of building energy diagnostics by incorporating a range of uncertainty in operation. Therefore, risk tolerance thresholds are defined on the basis of probability masses within an expected operational range. As described in the previous section, an expected (retrospective) range of energy consumption for each end use is generated by running batches of models with parameter values sampled from ranges of uncertainty. The probability masses are generated by placing an energy end use consumption observation in the context of the range of expected energy use, and then defining risk-tolerance thresholds at certain distribution quantiles on either side of the measured consumption point value. These thresholds shall receive the nomenclature M_1 and M_2. They are applied to bound either side of the energy consumption observation in the context of the expected consumption PDF. This yields two levels of risk tolerance and five distinct probability masses, as illustrated in Figure 59 below. Table 1, below, shows how the nomenclature for risk tolerance threshold is applied to several pieces of the Energy Signal Tool, which are explained in this section.

Table 1: Nomenclature for risk tolerance thresholds

Risk nomenclature		*State space nomenclature*	*Cost function nomenclature*
M_1	*yields*	**M_{low}** *slightly* lower and **M_{high}** *slightly* higher than expected	**(SL, SH)**
M_2	*yields*	**M_{vlow}** *much* lower and **M_{vhigh}** *much* higher than expected	**(ML, MH)**

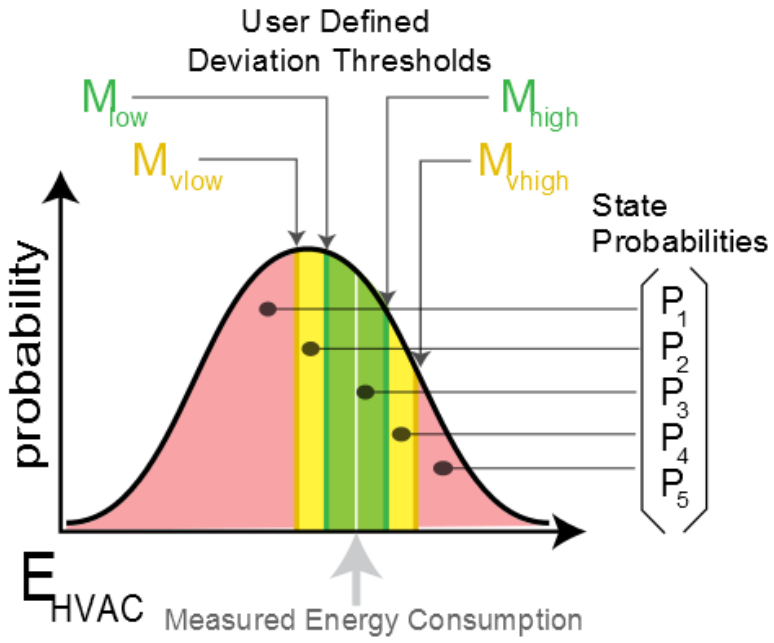


Figure 59: Basic probability distribution of expected energy end-use consumption (adapted from Henze *et al.*, 2015)

Figure 59 gives a graphical illustration of how the risk tolerance thresholds used to compare a measurement to an expected range of end-use consumption. The green region defines the probability range that the measured data is about the same, or acceptably close to the expected amount. This is bounded on the outer edges by M_{low} and M_{high}. The green region probability mass is maximized when observed consumption falls at the most likely expected consumption. The yellow regions define the probability ranges where the measured data shall be considered *slightly* higher or *slightly* lower than model expectations. The outer bounds of this region are M_{vlow} and M_{vhigh}. Finally, the red region defines the probability range where the measured data shall be considered *much* higher or *much* lower than model expectations. The notation SL, SH, ML and MH are used in the cost matrix sections in place of M_{low}, M_{high}, M_{vlow} and M_{vhigh}, respectively.

The state space probability masses, P_1 through P_5, are determined by the values of M_1 and M_2 thresholds for each end use. State space probabilities are defined by the following:

P_1 = Probability that observed energy use is *much higher* than expected

P_2 = Probability that observed energy use is *slightly higher* than expected

P_3 = Probability that observed energy use is *about the same* as expected

P_4 = Probability that observed energy use is *slightly lower* than expected

P_5 = Probability that observed energy use is *much lower* than expected

As illustrated in Figure 59 above, state space probability masses "slide" in the PDF with the observed consumption. For example, when measured consumption is higher than the mean of the expected distribution, P_1 and P_2 will grow larger, while P_4 and P_5 will become smaller. The energy signal is *not* simply computed based on the size of probability masses. As discussed later, utility theory allows the vector of probability masses to be combined with a cost function that incorporates more information about facilities management priorities.

Care must be taken in selecting the nominal M_1 and M_2 thresholds. As demonstrated in a later section, the ranges selected will determine how easily a deviation in observed energy end-use will trigger a signal to the energy manager alerting him or her that the building is slightly or greatly out of the expected range. Each end-use category may warrant a slightly different definition of threshold ranges. Some end-uses with more accurate sub-metering equipment could have narrower tolerance bounds. In this work, nominal values of 10% for M_1 and 20% for M_2 are used, and justification for this is given in Chapter 8. After nominal values are established, the risk tolerance threshold is further modified for end-use controllability, seasonal variation, and minimums for measurement uncertainty. This process is outlined in Figure 60 below, and described in the subsequent sections.

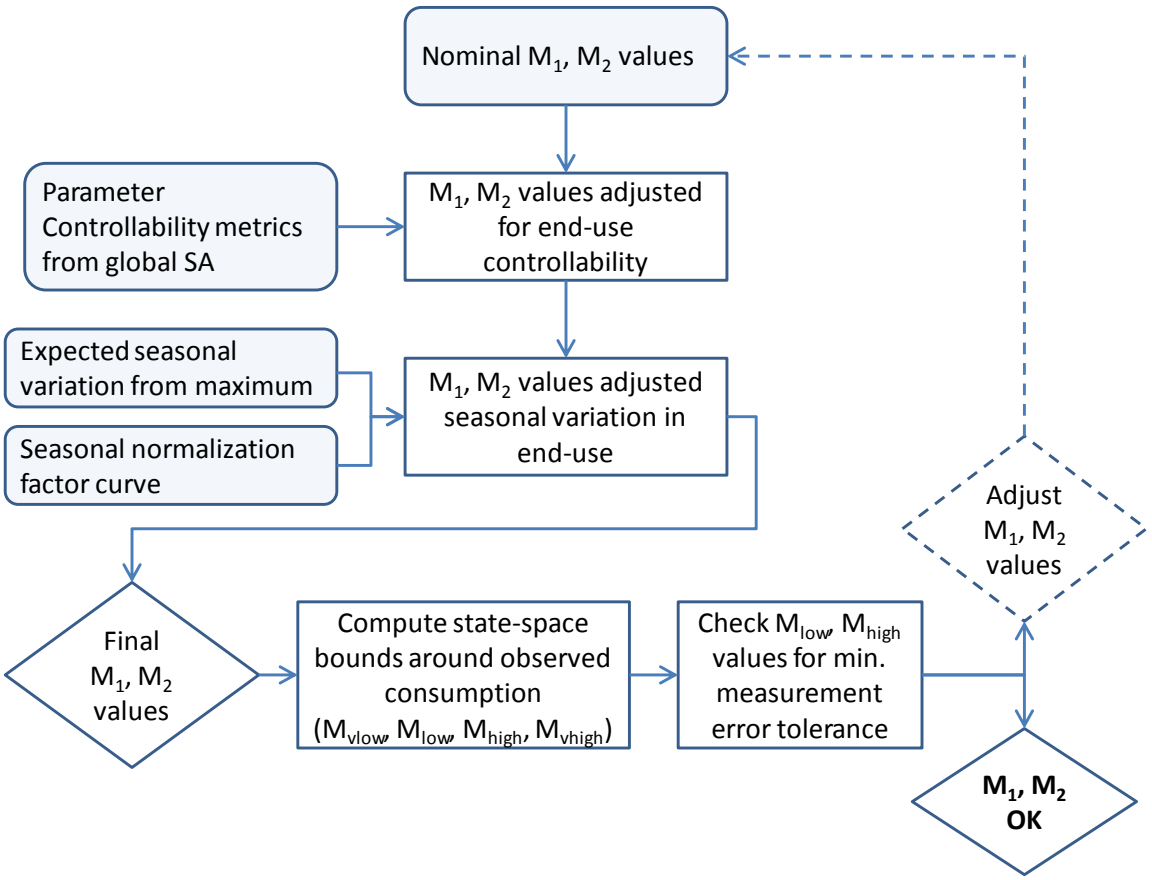


Figure 60: Flowchart for developing risk tolerance thresholds

7.3.1 Rule-based method to adjust the range of M_1 and M_2 for controllability

The uncertain parameters affecting the outcomes of the end use consumption distribution can be classified generally into two categories. These categories differentiate between parameters that can be acted on by the facilities manager at reasonable cost (controllable), and those that cannot be acted upon, or can only be acted upon at the scale of a building retrofit (outside the realm of normal operation and maintenance; or, uncontrollable). Depending on which types of parameters are most significant to an end use, the facilities manager may have more or less tolerance for deviation from expected usage. For example, if the range of expected energy consumption for heating use is due primarily to a range in possible infiltration values, which in turn depend on outdoor conditions, then the facility manager may have a greater tolerance for deviation before they

feel corrective action is needed. The definitions for controllable and uncontrollable parameters are given below:

Controllable Parameter: *If the value of this type of parameter is different from expected it is indicative of a system fault that the facilities manager needs to, and is able to address directly. The parameters relate to function and operations of equipment and the building.*

Uncontrollable Parameter: *When the value of this type of parameter is different from expected, it may be indicative of operation outside of normal range, but it is an abnormality that does not require maintenance or service attention. This parametric fault is not under the control of any building operator. These parameters can be physical traits of the building inherent to its construction, or characteristics of major equipment that change with age and cannot be fixed without a major investment which is outside the realm of energy management decision making.*

The values of M_1 and M_2 should reflect the percentage of uncontrollable variables acting upon (degree of uncontrollability for) each end use. In some cases, for example where it is known that uncontrollable variables such as customer volume are likely to affect end-use consumption on a regular or cyclical basis, the user may want to expand the green and yellow regions by increasing M_1 and/or M_2. In other cases, for example where it is known that only controllable variables such as set points or equipment efficiency are likely to affect end-use consumption on a regular basis, the user may want to shrink the green and yellow regions by decreasing M_1 and/or M_2.

The following method is proposed to determine the degree of uncontrollability and controllability for each end use, and subsequently assign the most appropriate threshold percentage values of M_1 and M_2. From the global sensitivity analysis, parameter significance index (SI) values are obtained that quantify the relative influence that each uncertain variable has on the end-use objective function. A descriptor of controllable (“c”)

or uncontrollable ("u") is assigned to each parameter. Table 12 below gives examples of variables that can be classified as controllable and uncontrollable.

Table 12: Examples of "controllable" and "uncontrollable" model parameters

Significant Uncertain Parameters	
Controllable	*Uncontrollable*
HX sensible effectiveness	Infiltration reduction
Gas burner efficiency	Wall conductivity mult
Roof absorptance mult	Ref suction piping UA
Heating setpoint adjustment	Fan pressure rise multiplier
Heating supply air temp design	Fan Efficiency
T-Stat float heat/cool	Heating sizing factor
DX Cool Coil COP	Anti-sweat multiplier
Walkin operating temp, delta	
System airflow ratio minimum	
Condensing temp minimum, delta	
Cooling setpoint adjustment	

The weighting of uncontrollable vs. controllable parameter significance is used to adjust risk tolerance thresholds. The following equation is used to determine the degree of controllability and uncontrollability of each end-use:

Equation 15: Degree of controllability

$$degree\ of\ controllability = C = \frac{\sum_{param}^{all="c"}(SI_{param})}{\sum_{param}^{all}(SI_{param})}$$

Equation 16: Degree of un-controllability

$$degree\ of\ uncontrollability = U = \frac{\sum_{param}^{all="u"}(SI_{param})}{\sum_{param}^{all}(SI_{param})}$$

Where SI_{param} is the significance of each parameter to each end use, as determined by global sensitivity analysis. And, M_2, M_1 are calculated as follows, where the nominal values of M_1 and M_2 are set at 5% and 10%, respectively:

Equation 17: Inner risk tolerance threshold adjusted for end-use controllability

$$M_1 = 10\% + 5\% * \left(\frac{\frac{U}{U+C} - 0.5}{0.5} \right)$$

Equation 18: Outer risk tolerance threshold adjusted for end-use controllability

$$M_2 = 20\% + 10\% * \left(\frac{\frac{U}{U+C} - 0.5}{0.5} \right)$$

In a brief explanation of equations above:

Standard $M_{1,\,nom}$ = 10%

Standard $M_{2,\,nom}$ = 20%

Expected value of $\frac{U}{U+C}$ = 0.5

Expected value of $\left(\frac{\frac{U}{U+C} - 0.5}{0.5} \right)$ = 0

For an end use that is composed of a greater number of uncontrollable parameters, the risk margins will grow wider, up to a maximum value of 30% for M_2. Table 13 below gives an example of calculating risk tolerance threshold values adjusted for controllability. Here, it is determined that whole building energy has a high degree of uncontrollability due to heavy influence from parameters such as infiltration, wall conductivity and heat exchanger effectiveness.

Table 13: Example of end-use sensitivity analysis results; controllable and uncontrollable uncertain parameters

Whole Building Energy			**Refrigeration**		
Variable	**μ***	C/U?	**Variable**	**μ***	C/U?
Infiltration Reduction	1.000	u	Ref Suction Piping UA	1.000	u
Gas Burner Eff	0.652	c	Condensing temp minimum, delta	0.467	c
Heating Setpoint adjustment	0.408	c	Heating Setpoint adjustment	0.269	c
Wall conductivity mult	0.302	u	Heating supply air temp design	0.153	c
HX Sensible Effectiveness	0.157	u	Walkin operating temp, delta	0.139	c
Heating Sizing Factor	0.095	u	T-Stat float heat/cool	0.057	c
Heating supply air temp design	0.087	c	Wall conductivity mult	0.033	u
System airflow ratio minimum	0.062	c	Infiltration Reduction	0.027	u
Fan Pressure Rise Multiplier	0.056	u	Fan Pressure Rise Multiplier	0.015	u
T-Stat float heat/cool	0.037	c	System airflow ratio minimum	0.008	c
Roof absorptance mult	0.034	c	HX Sensible Effectiveness	0.003	c
DX Cool Coil COP	0.011	c	Heating Sizing Factor	0.002	u
Ref Suction Piping UA	0.007	u	Fan Efficiency	0.001	u
Fan Efficiency	0.005	u			
Condensing temp minimum, delta	0.003	c			
Walkin operating temp, delta	0.002	c			
Cooling Setpoint adjustment	0.001	c			
Σ C:	**1.298**		**Σ C:**	**1.096**	
Σ U:	**1.623**		**Σ U:**	**1.077**	
M1	**10.6%**		**M1**	**10.0%**	
M2	**21.1%**		**M2**	**19.9%**	
where μ is calculated with AIC regression coefficients*					

7.3.2 Adjustment of M_1 and M_2 for Seasonal Variation

It is proposed that a second adjustment to M_1 and M_2 tolerance thresholds is needed to account for significant seasonal variation in *expected* energy consumption. This is common to some end-uses such as natural gas in climates with distinct seasonal changes in weather. The purpose of this adjustment is to shield the user from unnecessary signals during times of expected low seasonal consumption. This type of use is difficult to model, as well as low priority to act on. For example, the signal tool should not be as sensitive to an error in summer natural gas use, which may be very large in percentage but small in terms of magnitude. For the case study building, this rule is only applicable to the natural gas, and HVAC (fans, cooling, and heating electricity) end-uses – as shown in Figure 61 below. For other types of buildings, it could also apply to end uses such as cooling energy, heating

electricity, refrigeration, and other seasonal processes that the user may wish to monitor as end-uses.

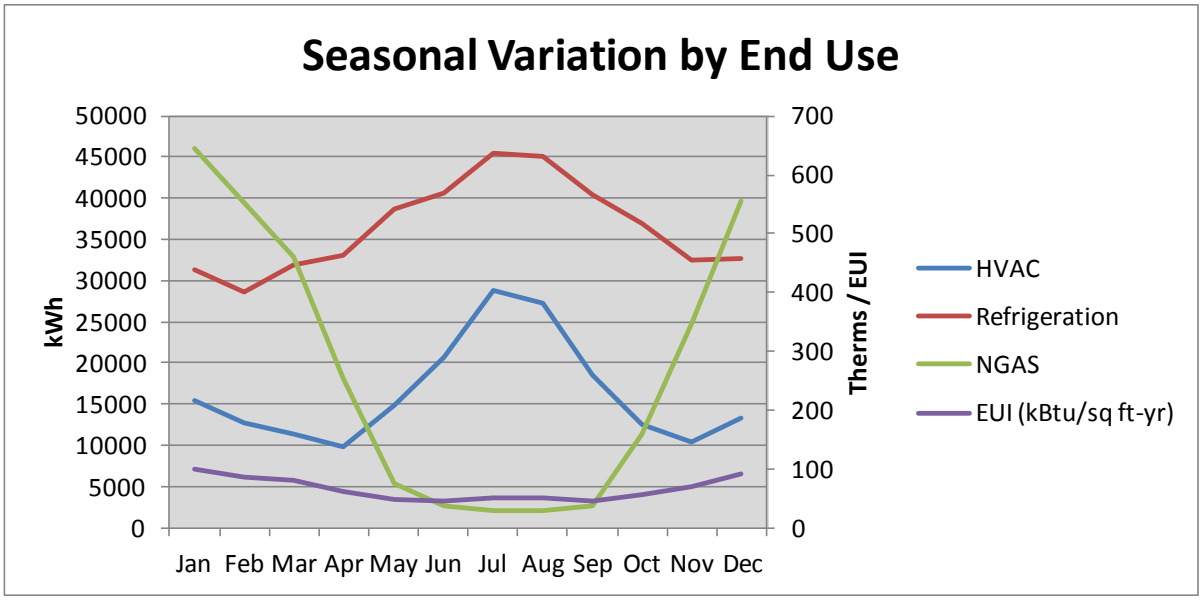


Figure 61: Seasonal variation in end-use consumption for the case study building. The only major seasonal shift is found in Natural Gas use.

A seasonal normalization factor for risk tolerance must, when used as a multiplier, have a shape that widens risk tolerance thresholds in non-critical seasons. Equation 19, below, describes a normalization factor curve, which varies with the ratio of the expected energy consumption of the period to the expected annual maximum of any period of equal length. This equation was derived to produce a curve that the author felt would best meet the needs of seasonal normalization, and would be applicable to any building.

Equation 19: Seasonal threshold range normalization factor

$$F_n = 1 + 1.1 * \left(-0.15 * ln\left(\frac{\bar{X}_{per}}{\bar{X}_{max}} \right) \right)^{0.9}$$

Where, X_{per} is the annualized average amount of consumption *expected* in the period, compared to X_{max} the annualized amount *expected* for any peak period of equal length throughout the year. It is important to base this factor on the *expected* use for any period. Basing the factor on the observed use would defeat the purpose of the tool by expanding

risk tolerance to what could encompass a major fault. Equation 19 produces the following logarithmic curve, in Figure 62 below, with minimum at 1 and maximum (where normalization factor will rise near vertically and asymptotically) to a maximum of close to *2*. The normalization factor, F_n, rises sharply as the ratio of $\frac{X_{per}}{X_{\max}}$ drops below ¼; where an expected periodic consumption so small would indicate a major seasonal shift.

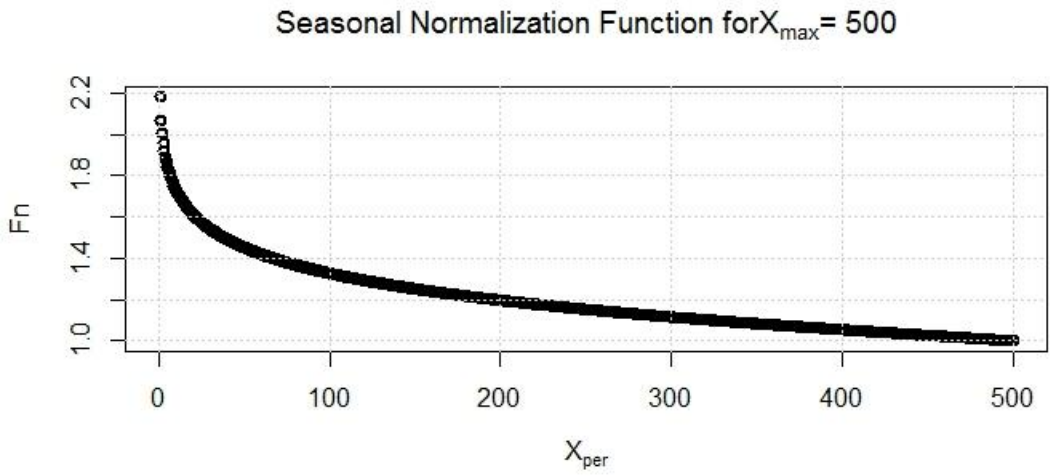


Figure 62: Example of seasonal normalization factor curve

The curve of this normalizing factor was carefully shaped to add significant leniency to the modeled expected range of results for periods such as summer gas usage and for shoulder seasons. The risk tolerance threshold is altered very little for periods when expected consumption is at or near critical levels. Now, M_1 and M_2 are defined, finally, as follows:

Equation 20: Risk tolerance probability mass thresholds corrected for seasonal variation

$$\boldsymbol{M_{1,adjusted} = F_n * M_1}$$

$$M_{2,adjusted} = F_n * M_2$$

7.3.3 Calculating the State Space Boundaries

Once M_1 and M_2 are obtained, an additional computation is needed to determine the *bounding values* for the risk tolerance thresholds of each end use. The proposed method of

calculating these ranges is as follows, where M_{low} and M_{high} are a function of the *probability masses* about the *mean of the expected data*:

Equation 21: Defining the outer bounds of the probability range quantifying the probability of operation as expected

$$M_{low} = X_{meas} - ((quantile(data, 0.5 - M_1)) - (median(data)))$$

$$M_{high} = X_{meas} - ((quantile(data, 0.5 + M_1)) - (median(data)))$$

Equation 22: Defining the outer bounds of the probability range quantifying the probability of operation slightly lower than expected.

$$M_{vlow} = X_{meas} - ((quantile(data, 0.5 - M_2)) - (median(data)))$$

$$M_{vhigh} = X_{meas} - ((quantile(data, 0.5 + M_2)) - (median(data)))$$

Where;

M_1 = Inner risk tolerance threshold percentage, as calculated in the previous section

M_2 = Outer risk tolerance threshold percentage, as calculated in the previous section

X_{meas} = Measured energy consumption data for an end-use summed over a given time period

data = the expected range of distribution for the energy end-use consumption over a given time period

And, the "quantile" function (in R, base package) produces the quantile of sample (data) corresponding to the given probability of a quantity M_1 or M_2 above and below the median expected value. M_{vlow}, M_{low}, M_{high}, and M_{vhigh} are the values of energy end-use magnitude that represent the risk tolerance thresholds. These are similar in function but different in definition from the values of $E_{0(low,high)}$ and $E_{1\ (low,\ high)}$ as defined by Henze *et al.*(2015).

In other words, there is a conversion from probability mass to the end-use values that would bound it if X_{meas} happened to equal the median expected value on the PDF. This is made even clearer by illustrating with a cumulative distribution function, as in Figure 63

below. In reference to Figure 63, Equation 21 describes the outer bounds of the green region, and Equation 22 describes the outer bounds of the yellow region.

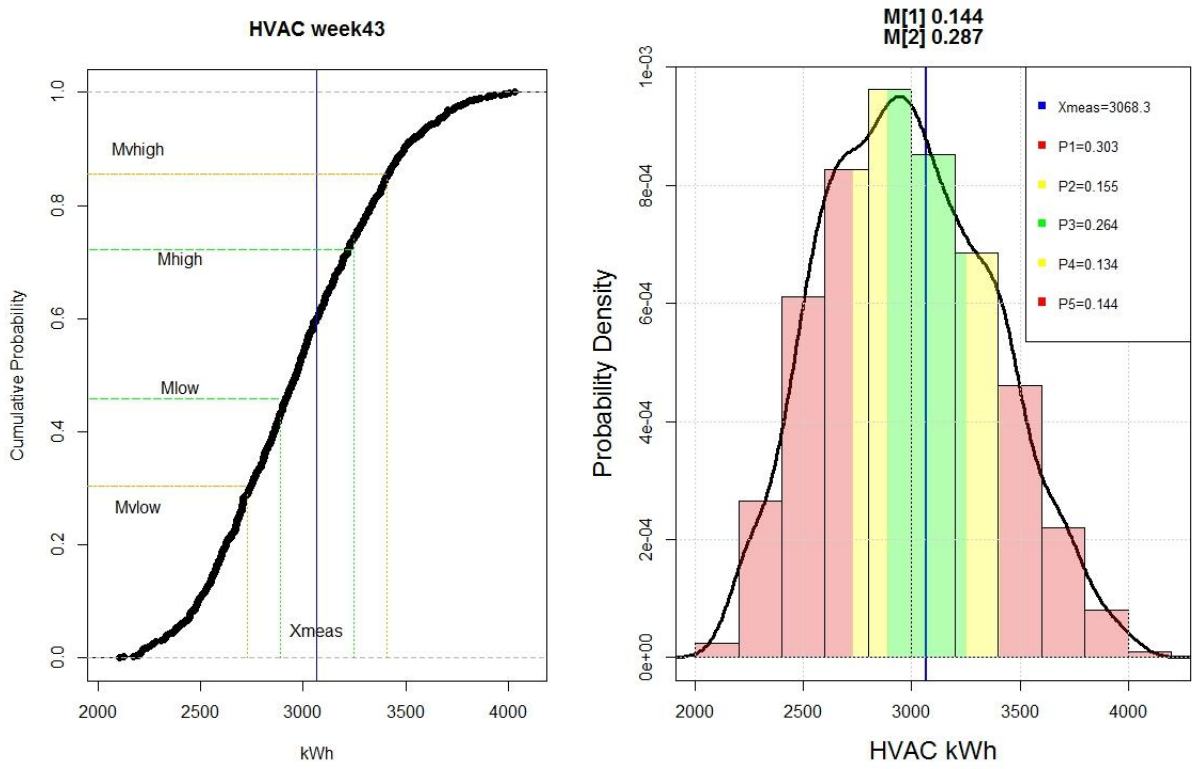


Figure 63: Sample output of PDF and CDF resulting from expected, measured data, and risk tolerance thresholds

Thus, computing the M_{low}, M_{high}, M_{vlow}, and M_{vhigh} values is done in reference to those distances bounding symmetrical distributions of probability masses about the mean value of expected consumption. This method incorporates the amount of expected uncertainty existing in the modeled end use calculation. The signal is thus generated based truly on the position of the measured value within the expected distribution. This way, the user will not fall into the trap of reserving too great of space for signals indicating "as expected", or slightly above/below expected values for end-uses with low modeled uncertainty. The final resultant state space probability masses can be computed graphically by the differences between the CDF values of the risk tolerance thresholds, as shown at left in Figure 63.

7.3.4 Constraints of measurement uncertainty

There is of course also some uncertainty in the sub-metered energy use measurement itself. Current transducers, used to monitor electricity consumption, have an accuracy of +/- 3% (for reasons discussed in an earlier section) of a real value obtained in a reading. To account for this, the width of the middle range of the signal output is defined by ensuring that, even for distributions with high kurtosis, the following is true:

$$X_{high,1} \leq X_{meas} * 0.97$$

$$X_{high,2} \geq X_{meas} * 1.03$$

In other words, it must be allowed for the range of the yellow signal reading to encompass potential measurement error. Since a yellow signal indicates the bounds of questionable (but not definite) action, this provision does not significantly impact recommended action. With substantial improvement in sub-metering technology, this condition for minimum risk tolerance could be omitted if needed.

7.4 Defining the Cost Matrix

The signal displaying the recommended course of action is generated by multiplying a [5x1] vector containing the state space probabilities, {P} (as described in the previous section), with a [5x5] cost matrix. The cost matrix, {C}, defines the cost of taking each of the five possible actions given each of the five possible observed energy performance scenarios.

In any energy management or maintenance scenario, the right approach to solving a problem is the approach that minimizes the marginal cost of operation. In facilities management for buildings, there are several components to the "marginal cost" of any end use, including:

- Cost of energy
- Variable cost of maintenance and service

- Value that the end use provides to the building function (e.g., thermal comfort)
- Value the end use provides to company image or branding (e.g., lighting)
- Value of the end use to providing inventory stability (e.g., refrigeration)

The cost matrix is an opportunity to incorporate the energy cost and other cost considerations for each end use into the energy signal tool. For facilities management, the cost matrix can be characterized with two basic dimensions: the costs of having problems, and the costs of solutions to those problems. Table 14, below, illustrates the concept of the cost function for the energy signal tool. **Bold** cells in red indicate a false signal (more than one state space away from the matching signal).

Table 14: Cost matrix definition

	State Probability of Observed vs. Expected Energy:				
Action Signal:	Much Higher	Slightly Higher	About Same	Slightly Lower	Much Lower
High Red	C(RH, MH)	C(RH, SH)	**C(RH, S)**	**C(RH, SL)**	**C(RH, ML)**
High Yellow	C(YH, MH)	C(YH, SH)	C(YH, S)	C(YH, SL)	C(YH, ML)
Green	**C(G, MH)**	C(G, SH)	C(G, S)	C(G, SL)	**C(G, ML)**
Low Yellow	**C(YL, MH)**	**C(YL, SH)**	C(YL, S)	C(YL, SL)	C(YL, ML)
Low Red	**C(RL, MH)**	**C(RL, SH)**	**C(RL, S)**	C(RL, SL)	C(RL, ML)

False High Signal

False Low Signal

Whereas the upper right and lower left regions of the cost matrix tend to contain signals classified as "false positive", the upper left and lower right regions contain "true positive" signals. Any green signal, is considered a negative signal in that is calls for no action. Green signals at the left or right ends of the cost matrix will tend to be "false negative", since the observed state is not the same as expected. A green signal in the middle of the cost matrix will tend to be a "true negative". It can only be worded that signals *tend* to be classified a certain way because this cannot be determined outside the context of the entire scenario including the state space probabilities. The classification skill section of chapter 8 goes into more detail on this topic.

Equation 23: Detailed cost function

Observed Energy Probabilities →

Action signals ↓

$$\begin{vmatrix} RH,MH & RH,SH & RH,S & RH,SL & RH,ML \\ YH,MH & YH,SH & YH,S & YH,SL & YH,ML \\ G,MH & G,SH & G,S & G,SL & G,ML \\ YL,MH & YL,SH & YL,S & YL,SL & YL,ML \\ RL,MH & RL,SH & RL,S & RL,SL & RL,ML \end{vmatrix}$$

Table 14**Error! Reference source not found.**, above, is a detailed look at the contents of the cost matrix. Where, [MH, SH, S, SL, ML] and [RH; YH; G; YL; YL; RL] are defined in Equation 25 and Table 14 above, respectively. The values in each cell of the cost matrix indicate the cost of taking false action, or no action, given the state observed. As an input to the Energy Signal Tool, the cost function captures the response that the facility manager would recommend for each unique scenario of observed over or under consumption. It compares how important it is to act to address high energy consumption, low energy consumption, and to take no action depending on what state is observed. For example, if natural gas used is observed to be much higher than expected, and if this is a significant operational expenditure and the cost to fix it is comparatively low, immediate action to address high gas use would be recommended.

7.4.1 The Neutral Cost matrix

Equation 24: The Neutral Cost Matrix

$$\begin{vmatrix} 0 & 2 & 4 & 6 & 8 \\ 2 & 0 & 2 & 4 & 6 \\ 4 & 2 & 0 & 2 & 4 \\ 6 & 4 & 2 & 0 & 2 \\ 8 & 6 & 4 & 2 & 0 \end{vmatrix}$$

Equation 24, above, shows the neutral cost matrix, or that which would be used in the absence of other information about cost considerations for the building end-use. The scalar quantities in the cost matrix penalize sending a signal for false action. The highest

"costs" are associated with false signals as shown above in Table 14. For example, a cost value of 8 is assigned to displaying a "high red" signal when energy use is much lower than expected, or for showing a "low red" signal when energy use is much higher than expected. A cost of 0 is assigned to a decision to display the proper signal corresponding to the observed vs. expected energy use; since right action should minimize cost. Figure 59 above has shown that there is some probability of being in each state of observed consumption with respect to the expected range. The values in the cost function are highly subjective to end-use application, and are presented merely as scalar values in this work. This begs future quantitative research that would help improve signal prioritization (see section 7.6),

By adding across each *row* of the cost matrix, the total probable cost (scalar) of each action is determined. Lacking information about the scenario, the total cost of taking action (row RH, or RL) will always be higher than not taking action, since taking action means the facilities manager is initiating a service call. By adding down each *column* of the cost matrix, the relative cost of each state is determined. Cost of states in lower than expected consumption are assumed to be the same as those states with higher than expected consumption because, in addition to energy cost considerations, there is inherent value assumed in systems operating as expected. In the example above of the neutral cost matrix, the cost values are assigned and distributed in a symmetric fashion, without a bias to a false high or false low signal.

7.4.2 Secondary Cost Matrices

Assigning a unique cost matrix to each end use adds some intelligence to the signal generation process. Secondary cost matrices will introduce some signal bias as appropriate. For example, actions to address lower than expected energy use for a refrigeration system would have a different priority level than actions to address low energy use for a cooling system in a cool climate. Therefore, to further ensure that the cost matrix meets the

operational and maintenance cost expectations of the facilities manager, the user provides an answer to two simple questions for each end use:

1. **Q:** Is the ratio of $\frac{cost\ to\ maintain}{cost\ to\ operate}$ high or low?

   A: High, Low, N/A

   (high: chiller in warm climate; low: heating in warm climate)

2. **Q:** Does a high or low energy consumption reading indicate especially serious problems relating to customer comfort, product stability or otherwise especially damaging to operation?

   A: High, Low, N/A

   (high: equipment sensitive to overuse; low: refrigeration or ventilation fans)

The user may choose to answer each question with "High", "Low", or "N/A" (not applicable) which can be generally applied to any monitored end-use. By answering the questions, the user has the chance to modify the pre-defined cost matrix with their preferences about this end-use. This can help to incorporate knowledge of the relative maintenance and repair costs for each end use. If the response to both questions is "NA", then the pre-defined cost matrix is left unchanged. If just one question is answered "High"/ "Low", and the other is answered "NA" then the predefined matrix is changed less than if both questions are answered (see appendix for calculation examples). For example, answering the questions sequentially with "High, High" will result in a cost matrix that penalizes false signals (lower left and upper right corners) heavily, and that especially penalizes false low signals (columns [1,2]). General examples of transforming the cost matrix are given in the figure below.

Decision Cost Matrices

Bias against false low signal

0	1	2	4	6
2	0	1	3	4
4	2	0	1	2
7	5	2	0	1
10	7	4	2	0

Neutral matrix

0	2	4	6	8
2	0	2	4	6
4	2	0	2	4
6	4	2	0	2
8	6	4	2	0

Bias against false high signal

0	2	4	7	10
1	0	2	5	7
2	1	0	1	4
4	3	1	0	2
6	4	2	1	0

Figure 64: Examples of neutral and biased cost matrices

The matrix on the left in Figure 64 above would be a result of answering "High, High" to the two questions above. The figure in the middle would be the result of answering "N/A, N/A". The matrix on the right would be the result of answering "Low, Low". The cost matrix definition process can be seen as a decision tree, where the user starts with a "neutral" cost matrix, and by answering the two questions above, will end with one of nine possible secondary cost matrices. Figure 65 below depicts this flow.

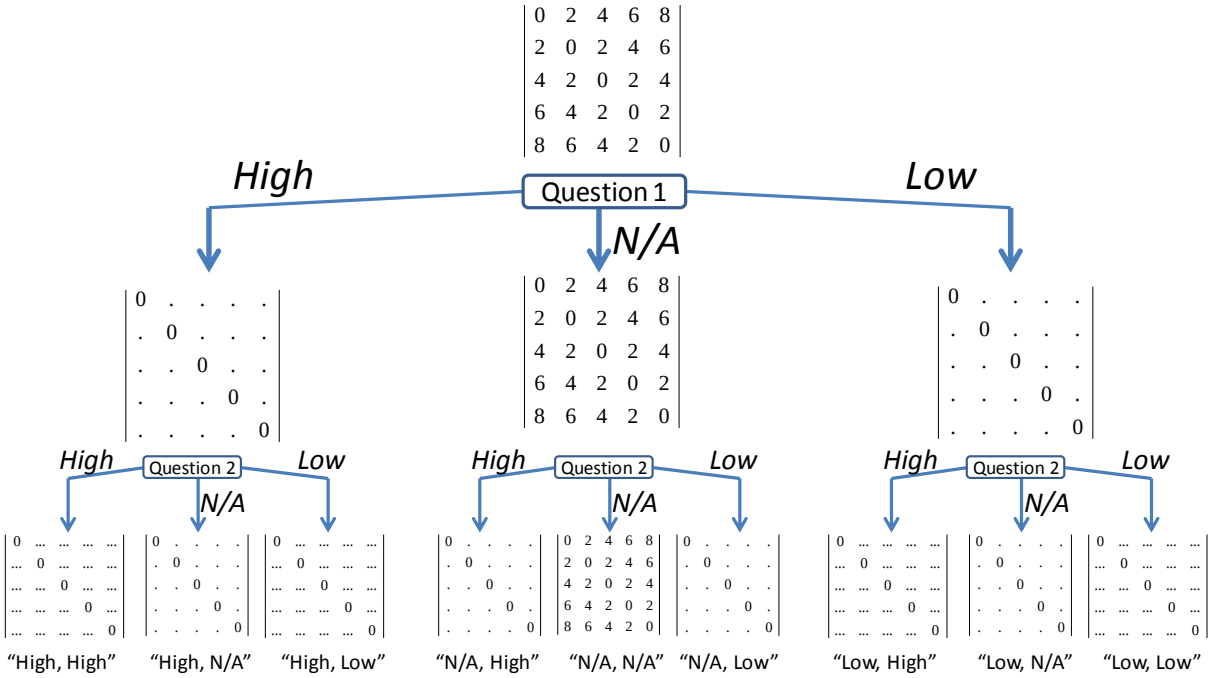


Figure 65: Cost matrix decision tree

The energy signal tool proceeds with secondary cost matrix as defined by the user. See appendix for additional calculation details and the full array of possible cost matrices.

7.5 Energy Signal Tool Output

So far, we have described the background information needed for a data processing script that takes a dataset of expected (modeled) energy end-use consumption ("data") and compares it to a measured point ("X_{meas}"). The following section describes how this data is transformed into action signals that give a quick view of how a building is performing relative to expectations, and what (if any) action is recommended. The five state-space probabilities of X_{meas} relative to the distribution of expected consumption were calculated as the probability masses bounded by the risk tolerance thresholds. The state probabilities form a vector as follows:

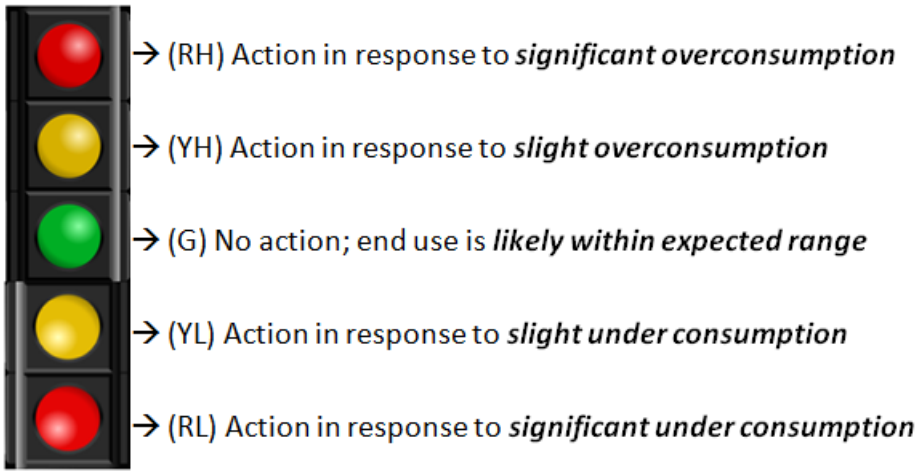
Equation 25: State probabilities vector

$$P=\begin{bmatrix}P_1\\P_2\\P_3\\P_4\\P_5\end{bmatrix}=\begin{bmatrix}MH\\SH\\S\\SL\\ML\end{bmatrix}$$

Where, P_1, P_2, P_3, P_4, P_5 are defined above and illustrated in Figure 63. Expected energy use is a probabilistic range created by LHS sampling runs of the calibrated model over the range of each uncertain parameter. The cost matrix has been defined above. These three items encompass the required gathering of external data. The following steps illustrate how this external information is processed into a signal for decision support.

7.5.1 Defining the Action Signal

The traffic light as applied to giving signals on roadways has been refined to give drivers a quick understanding of what action to take when approaching a road intersection. If a traffic light had five levels, the added lower set of yellow and red signals could alert drivers of actions they need to take in abnormal traffic situations; such as emergency vehicle passage, construction, or inclement weather. Similarly, the extended traffic signal can be a way to visualize recommended responses to the state of building end-uses in terms of observed energy consumption versus the range of expected consumption.



Other types of visualizations, such as circular gauges similar to a speedometer, could also be appropriate for this application. One could imagine a gauge that is calibrated to an expected range of energy use, and has green, yellow, and red regions bounding the mean value that indicate thresholds of energy consumption outside the tolerance range for normal use. One such example in operation is the Energy Tracker dashboard implemented at the NREL Research Support Facility (Henze *et al.*, 2014[xx]).

It is quite clear that higher than expected energy consumption is a concern, but both high and low consumption have significance. For example, if refrigeration energy were to be observed at much lower than expected (low red) in a building, the facilities manager might be concerned about the stability of food products and would want to investigate the problem soon. Or, if HVAC energy were observed at slightly higher than expected (high yellow), a reasonable response would be to look in the maintenance database for any other recent complaints relating to HVAC in that store, and then decide what, if any, action to take. When used to prioritize retrofit measures across a portfolio of buildings, lower signals would indicate better returns on investment. In this case, an appropriate response would be to implement the retrofit measures in those buildings with the most severe low signals.

7.5.2 Computing the Expected Cost Vector

The facilities manager is interested in taking action that will minimize cost for the organization based on several criteria as discussed previously. Utility theory is used in the

Energy Signal Tool to advise the best action, given the scenario of measured observations, in the context of uncertain expectations for the scenario (Henze *et al.,* 2015). Taking the dot product of the cost matrix and the state probability vector, yields a [5x1] vector, E_c. This describes the expected cost of the five possible courses of action as defined above (per Henze *et al.,* 2015).

Equation 26: Expected cost vector

$$E_c = C \bullet P = \begin{bmatrix} C(RH) \\ C(YH) \\ C(G) \\ C(YL) \\ C(RL) \end{bmatrix}$$

And,

$$signal = \min[E_c]$$

The expected cost vector reports probable costs as scalar values. The type of response to the current scenario (*i.e.,* recent past observation) that minimizes expected cost should be recommended to the facilities manager. The five-level traffic light level is designed to align with the E_c vector. The light corresponding to the minimum term in the E_c vector is "illuminated" to signify that some action should be made after diagnosing the problem. For example, if the vector *P* is taken from the above illustration in Figure 63, and *C* from Equation 24 above (neutral cost matrix) is used for the cost matrix, then E_c and its corresponding signal would be calculated as follows:

$$\begin{vmatrix} 0 & 2 & 4 & 6 & 8 \\ 2 & 0 & 2 & 4 & 6 \\ 4 & 2 & 0 & 2 & 4 \\ 6 & 4 & 2 & 0 & 2 \\ 8 & 6 & 4 & 2 & 0 \end{vmatrix} \bullet \begin{bmatrix} 0.667 \\ 0.0654 \\ 0.1430 \\ 0.0321 \\ 0.0924 \end{bmatrix} = \boldsymbol{E_{c,\,HVAC}} = \begin{bmatrix} \mathbf{1.635} \\ 2.303 \\ 3.233 \\ 4.734 \\ 6.365 \end{bmatrix} =$$

Cost Matrix **State Probabilities** **Expected Cost** **Signal**

Here, since the minimum term in the expected cost function is the first term in E_c (likely cost of issuing an RH signal) it will be best to give a "high red" signal. If, hypothetically, the state-space probabilities were to be distributed symmetrically (where the observed consumption falls right at the mean of expected), then the middle term would be minimized and a "green" signal would be produced.

$$\begin{vmatrix} 0 & 2 & 4 & 6 & 8 \\ 2 & 0 & 2 & 4 & 6 \\ 4 & 2 & 0 & 2 & 4 \\ 6 & 4 & 2 & 0 & 2 \\ 8 & 6 & 4 & 2 & 0 \end{vmatrix} \bullet \begin{bmatrix} 0.226 \\ 0.133 \\ 0.28 \\ 0.134 \\ 0.226 \end{bmatrix} = E_{c,\,HVAC} = \begin{bmatrix} 4.054 \\ 2.932 \\ \mathbf{2.342} \\ 2.872 \\ 3.938 \end{bmatrix} =$$

Or, if the end-use were to warrant a cost matrix very heavily biased against a false low action signal, such as in the example below, we can see that it is possible to obtain a yellow signal even with a symmetrical state space probability vector[12].

$$\begin{vmatrix} 0 & 1 & 2 & 4 & 6 \\ 2 & 0 & 1 & 2 & 4 \\ 5 & 3 & 0 & 1 & 2 \\ 7 & 5 & 2 & 0 & 1 \\ 10 & 7 & 4 & 2 & 0 \end{vmatrix} \bullet \begin{bmatrix} 0.226 \\ 0.133 \\ 0.28 \\ 0.134 \\ 0.226 \end{bmatrix} = E_{c,\,HVAC} = \begin{bmatrix} 2.585 \\ \mathbf{1.904} \\ 2.115 \\ 3.033 \\ 4.579 \end{bmatrix} =$$

In another example, if it is known to be especially expensive to call maintenance for a particular end use, a bias could be set that adds cost for higher level signals, and would thus only recommend definite action when the problem is quite urgent. The first example below shows how normally a red signal for definite action might be generated by higher than expected energy consumption with a neutral cost matrix.

[12] Such a heavily biased cost matrix is not recommended, and is only shown for purposes of illustrating the impact of the cost matrix.

$$\begin{vmatrix} 0 & 2 & 4 & 6 & 8 \\ 2 & 0 & 2 & 4 & 6 \\ 4 & 2 & 0 & 2 & 4 \\ 6 & 4 & 2 & 0 & 2 \\ 8 & 6 & 4 & 2 & 0 \end{vmatrix} \bullet \begin{bmatrix} 0.53 \\ 0.21 \\ 0.16 \\ 0.04 \\ 0.06 \end{bmatrix} = E_{c\,\text{HVAC}} = \begin{bmatrix} \mathbf{1.78} \\ 1.90 \\ 2.86 \\ 4.46 \\ 6.22 \end{bmatrix} =$$

The next case illustrates how adding a scalar bias to the upper and lower rows of the cost matrix changes this signal from one calling for immediate action to one calling for monitoring the problem.

$$\begin{vmatrix} 0 & 2 & 4 & 7 & 12 \\ 2 & 0 & 2 & 6 & 7 \\ 4 & 2 & 0 & 2 & 4 \\ 7 & 6 & 2 & 0 & 2 \\ 12 & 7 & 4 & 2 & 0 \end{vmatrix} \bullet \begin{bmatrix} 0.53 \\ 0.21 \\ 0.16 \\ 0.04 \\ 0.06 \end{bmatrix} = E_{c\,\text{,HVAC}} = \begin{bmatrix} 2.06 \\ \mathbf{2.04} \\ 2.86 \\ 5.41 \\ 8.55 \end{bmatrix} =$$

7.6 Signal Prioritization

The facilities manager will undoubtedly have multiple signals (red and yellow) calling for action or monitoring across a large portfolio of buildings. Their job is to minimize the risk of operating buildings improperly and minimize the cost of resources used to address problems. Or, perhaps they wish to determine the best opportunity for implementing a minor energy efficiency retrofit. In either case, they need to be able to prioritize their actions according to the expected cost / benefit. There are two methods proposed here to do this: prioritization based on signal urgency, and prioritization based on measured deviation from expected cost.

Taking the dot product of the cost matrix and the state probability vector, yields a [5x1] vector, *Ec*. This describes the expected cost of the five possible courses of action as defined above (per Henze *et al.,* 2015).

Equation 26 above) over the minimum term. This ratio effectively indicates the strength of, or confidence in, the recommended course of action. In other words, this ratio expresses the risk of the recommended action versus the risk in taking no action or taking the wrong action.

The primary benefit of incorporating uncertainty in decision support is the ability to quantify the urgency of responding to a situation (Henze *et. al.*, 2015). Measured consumption that falls near the outer tails of the expected use distribution is likely require more urgent action. If costs are defined carefully, the expected cost vector can be used to rank and prioritize actions according to urgency. Taking the dot product of the cost matrix and the state space probability vector yields a [5x1] vector, E_c (per Henze *et al., 2015)*. The quantitative urgency of an action can be determined by taking the ratio of the maximum term in the expected cost vector (E_c, see Equation 26 above) over the minimum term. This ratio effectively indicates the strength of, or confidence in, the recommended course of action.

Equation 27: Signal Strength Ratio

$$SR = \frac{(\max(E_c, end\ use))}{(\min(E_c, end\ use))}$$

Since the expected cost vector is made up of scalar values, the signal strength ratio is valid *within* each end-use category to prioritize signals from across the entire portfolio. The signal ratio values are scalars themselves; heavily influenced by those values used in the cost matrix – also just scalar representatives. An urgent signal for action to fix a problem with the heating system may have more energy cost associated with it, but may not be as important as a signal to fix refrigeration equipment.

It is up to the facilities management to decide how to prioritize action among multiple end-uses, and reporting energy cost deviation ranges can help with this. Energy cost deviation is one possible way to rank faults in order of priority. The expected energy

use distribution range can be used to define an expected range of cost deviation. Equation 28 below describes how cost deviation is calculated.

Equation 28: Probable range of cost deviation

$$C_{d,j} = (unit\ cost)(X_{meas} - X_{97.5}X^{97.5} : X_{meas} - X_{2.5}X^{2.5})$$

Where;

$C_{d,j}$ = 95% confidence interval for the range of deviation of observed from expected cost

$unit\ cost$ = Unit cost of energy in $/kWh or $/Therm, etc.

X_{meas} = Measured energy use

$X^{97.5} = 97.5\ percentile\ of\ expected\ energy\ use$

$X^{2.5} = 2.5\ percentile\ of\ expected\ energy\ use$

This range expresses the 95% confidence interval estimate of the monetary value of cost deviation of each end use when comparing the observed consumption to the expected range of consumption. Positive values indicate how much money is potentially being spent in excess. It can be seen that the signal priority ration is highest when the 95% confidence interval has either all negative or all positive monetary values. Throughout the remainder of this report, *mean cost deviation* is referenced. This is computed simply by taking the arithmetic mean of the resultant cost bounds computed in Equation 28 above.

7.7 Sample Outputs from Developer interface

The figures below give outputs from the R environment user interface developed in this work. As described in the next section, all synthetic faults were tested and processed in this environment. This is the functioning interface at the moment, and its use requires some skill with R and data set manipulation. This does not fulfill the vision for the final ESTool interface (as proposed conceptually in Figure 68 and Figure 69 below), but is the closest working approximation given the programming skills of this researcher.

7.7.1 Signal Output

The signal output shows recommended actions for each end-use, as calculated by the minimum of the expected cost matrix. The numbers below each signal indicate the signal strength ratio (indicating the level of urgency), and second the 95% confidence interval of expected cost deviation, as calculated above in Equation 28.

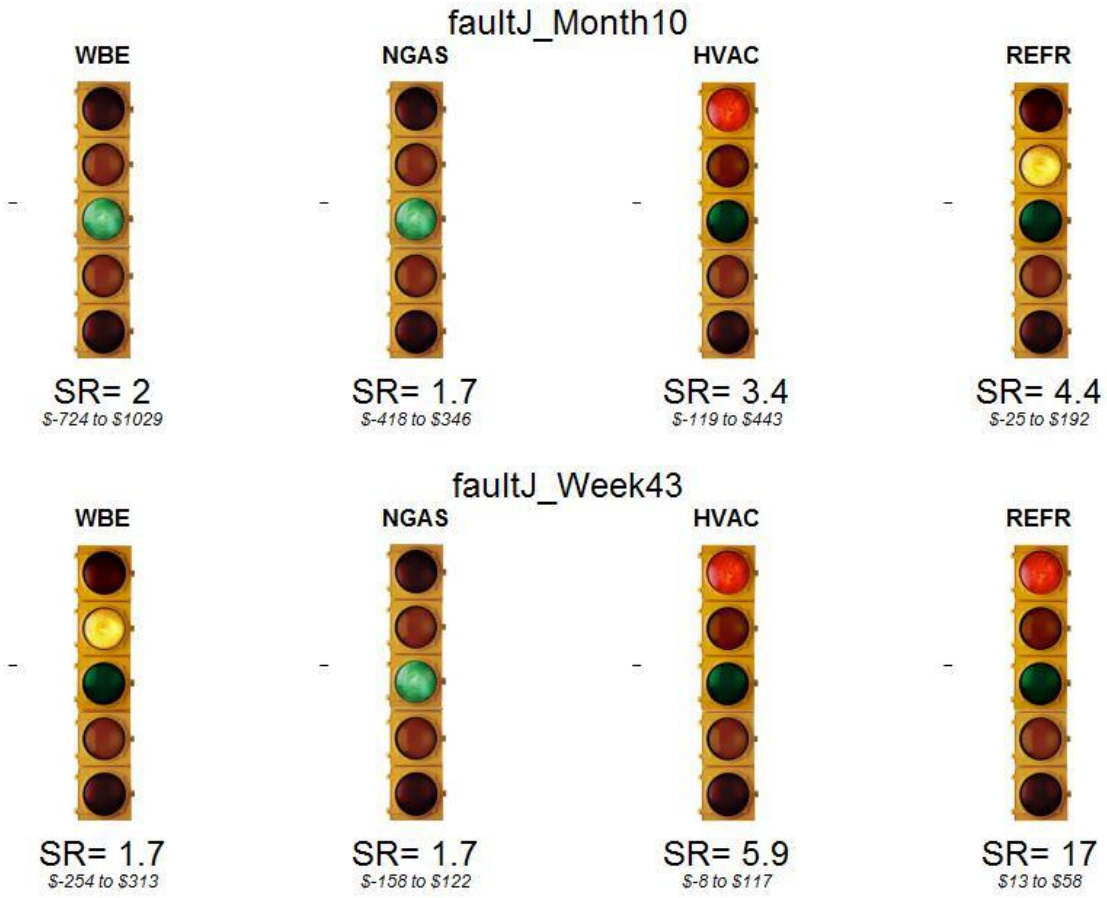
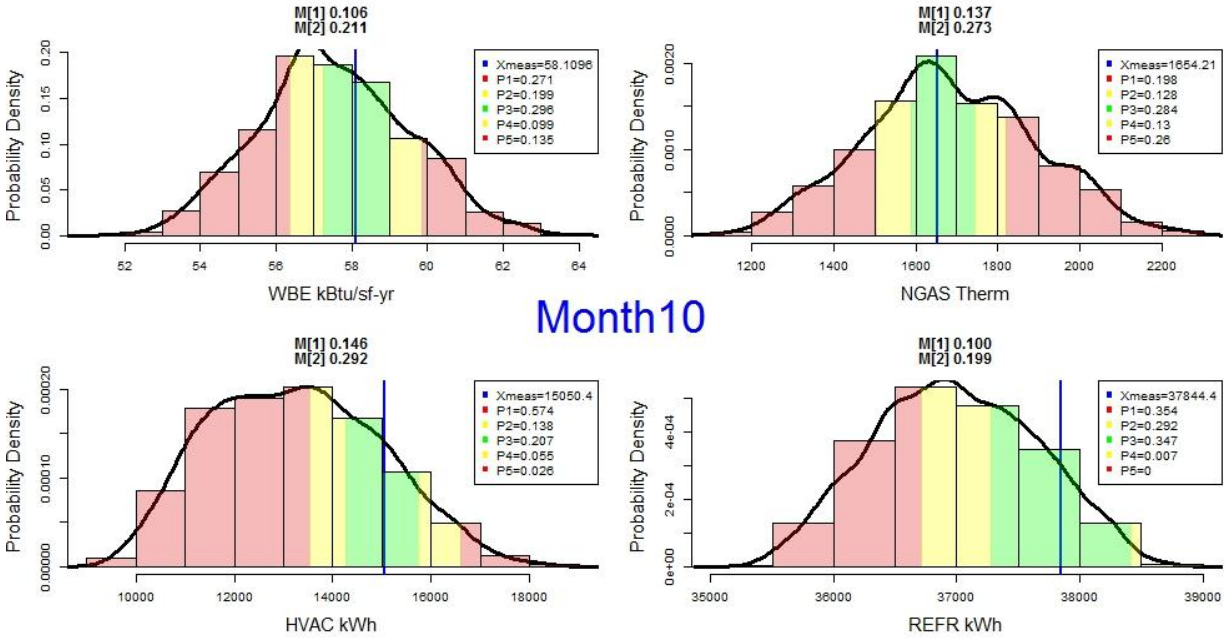


Figure 66: Sample outputs of end-use signals for the month of October and the last week in October. The fault tested was an increase in the thermostat float range

7.7.2 PDF Output

The second meaningful output of the ESTool R environment user-interface are Probability Density Function graphics for each end-use, at each time period. These PDFs help the user to visualize the state-space probability masses, as well as the position of the observed consumption within the expected distribution range. At the top of each PDF, the

calculated risk tolerance thresholds (M_1 and M_2) are given. At the right of each graph, a legend displays the sizes of the probability mass regions for each of the five state spaces. These regions are shaded in appropriate colors. A vertical blue line shows the position of the observed data (X_{meas}). Sometimes, X_{meas} may fall completely outside of the expected distribution range. This just means that a red signal will definitely be generated for this period. The signal strength ratio will be infinite as a result of some state space probabilities having mass of zero. In the remainder of this work, any signal priority ratios with infinite value were changed to a value of 6000, which is still greater than any others observed for purposes of ranking.



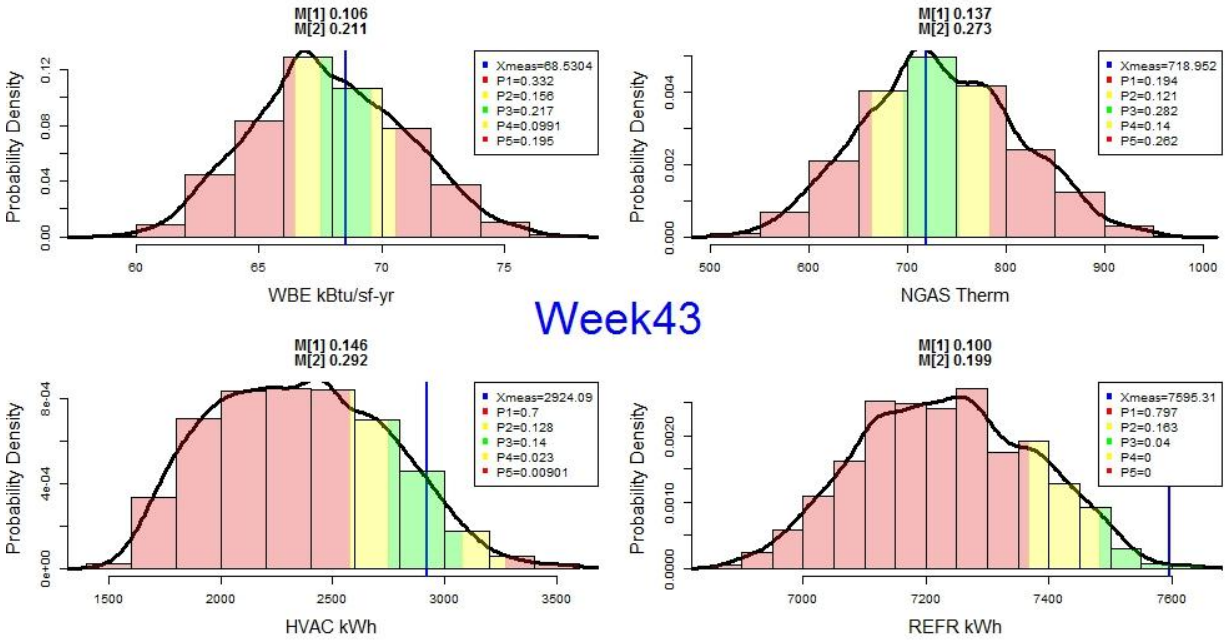


Figure 67a,b: Sample PDFs graphic outputs for the month of October and the last week in October. These correspond with the signal visualization output in Figure 66 above.

7.7.3 Conceptual User Interface

The vision for the Energy Signal Tool is that it is something that a portfolio energy manager could look to on a weekly or monthly basis for decision support. It would be an improved way to obtain an overview of building performance assessment. Performance information is summarized succinctly, rather than having to dig through mountains of sub-metering data to extract it. With input from the retail partner, a concept for a building portfolio management user interface to the Energy Signal Tool was developed. This interface would contain basic information about each building, display visual signal output, as well as give prioritization metrics. Such an interface could be coded so that it would be available online in HTML format to allow for remote access at multiple locations. The following figures show one type of view that would display only mean deviation from expected energy cost, and another that would give full results of the energy signal.

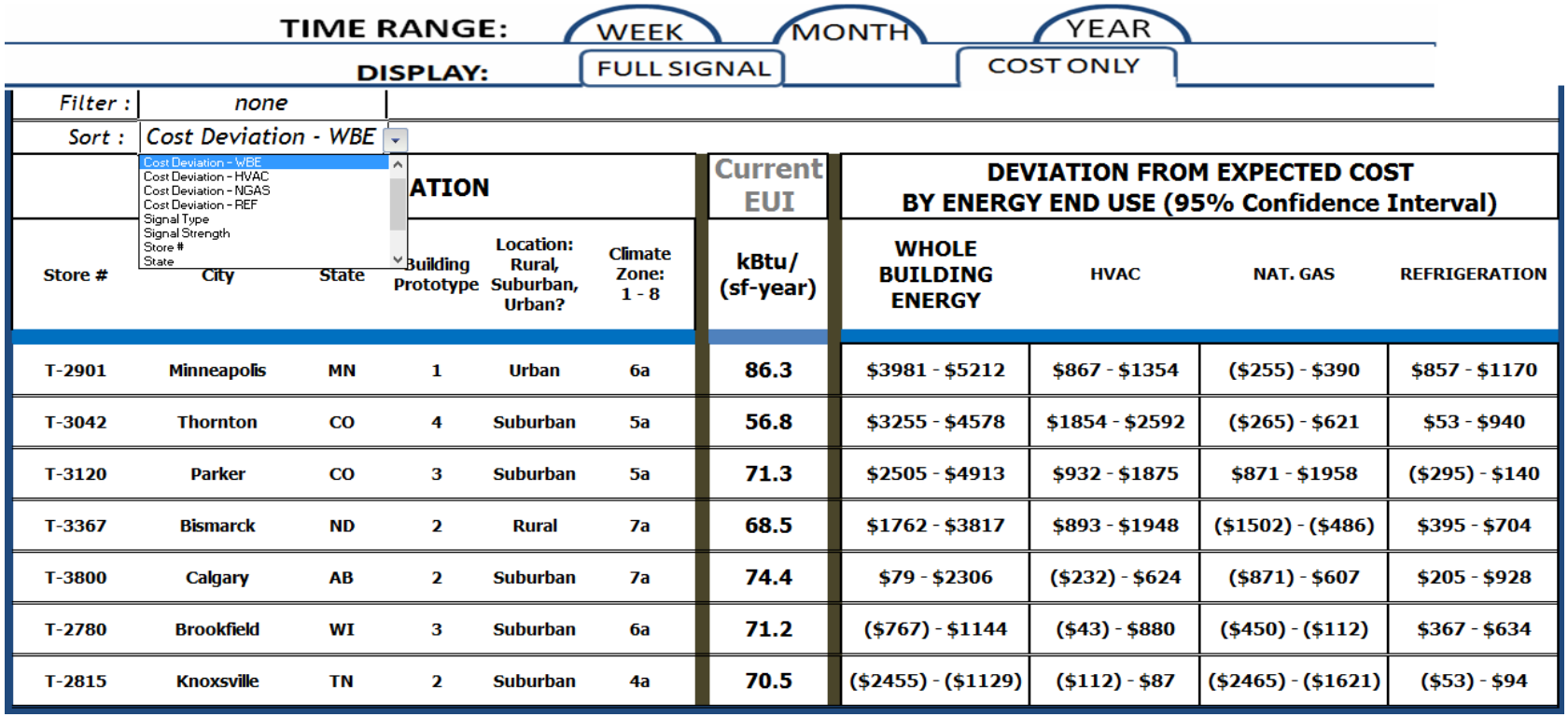
The detailed modeling of the unique qualities of each building, as well as actual weather data, allows the Energy Signal Tool to produce self-benchmarked results. Unlike

peer benchmarking, any sub-set of self-benchmarked results is a relevant basis on which to compare building performance. This is a much more customized and credible decision-making platform than the EPA Portfolio Manager can provide. With the ESTool, maintenance or retrofit prioritization can be done directly from any comparison. The user interface would allow the energy manager multiple ways to sort and filter performance assessment results. Filtering would produce sub-sets for comparison, and sorting would help indicate action priority. For example, a user might be interested in filtering to view only buildings of a certain store prototype (model), and then sort those by action signal priority ratio. This would give the energy manager a quick view of top priorities to address based on overall organization goals. Or, the user could filter buildings by climate zone, and sort by deviation from expected cost of natural gas use. This would give the energy manager a quick view of best opportunities for energy cost savings. Displaying building energy use intensity (EUI) in the user interface; this can help the energy manager track building performance relative to broad energy performance targets, such as the AIA 2030 challenge[13] or other sustainable building operation standards.

[13] The Architecture 2030 challenge proposes that continuous improvements in building design and operation be made such that by the year 2030, buildings will not need to rely on any fossil fuel or GHG emitting sources to operate. This can only happen from a combination of increased renewable energy penetration and better building energy efficiency. A building portfolio manager with on-site renewable energy generation could track both energy generation at each building and energy consumption, and report these as net EUI.

Figure 68: Conceptual ESTool user interface; view of full-signal output information

Figure 69: Conceptual ESTool user interface; view of cost only output information



TIME RANGE: WEEK MONTH YEAR

DISPLAY: FULL SIGNAL COST ONLY

Filter : none

Sort : Cost Deviation - WBE

Cost Deviation - WBE
Cost Deviation - HVAC
Cost Deviation - NGAS
Cost Deviation - REF
Signal Type
Signal Strength
Store #
State

ATION						Current EUI	DEVIATION FROM EXPECTED COST BY ENERGY END USE (95% Confidence Interval)			
Store #	City	State	Building Prototype	Location: Rural, Suburban, Urban?	Climate Zone: 1 - 8	kBtu/ (sf-year)	WHOLE BUILDING ENERGY	HVAC	NAT. GAS	REFRIGERATION
T-2901	Minneapolis	MN	1	Urban	6a	86.3	$3981 - $5212	$867 - $1354	($255) - $390	$857 - $1170
T-3042	Thornton	CO	4	Suburban	5a	56.8	$3255 - $4578	$1854 - $2592	($265) - $621	$53 - $940
T-3120	Parker	CO	3	Suburban	5a	71.3	$2505 - $4913	$932 - $1875	$871 - $1958	($295) - $140
T-3367	Bismarck	ND	2	Rural	7a	68.5	$1762 - $3817	$893 - $1948	($1502) - ($486)	$395 - $704
T-3800	Calgary	AB	2	Suburban	7a	74.4	$79 - $2306	($232) - $624	($871) - $607	$205 - $928
T-2780	Brookfield	WI	3	Suburban	6a	71.2	($767) - $1144	($43) - $880	($450) - ($112)	$367 - $634
T-2815	Knoxsville	TN	2	Suburban	4a	70.5	($2455) - ($1129)	($112) - $87	($2465) - ($1621)	($53) - $94

8 Testing and Tuning the Energy Signal Tool with Synthetic Faults

As noted above, the many anomalies found in the sub-metered data made model calibration impractical. In addition to these issues, the research team lost all contact with the project sponsor as of October, 2014; which meant that unexpected changes in operational strategy could not be investigated.

The combination of these problems led to there being insufficient information with which to complete a model that accurately characterized a fault-free building. It was decided that using a more stable source of building performance data, such as synthetic data, would be best for testing and tuning of the ESTool and the associated theory. Benefits of testing the ESTool on synthetic faults in an ideal data collection scenario include the following:

- Isolates the fault detection capabilities from the unknowns that exist in the building model
- Tests the sensitivity if the tool on a wide variety of typical building faults, rather than waiting for faults to arise
- Tests the new theory developed for defining customized risk tolerance thresholds for each end use
- Tests the new cost matrix definition methodology, and could lead to refinement of this methodology with examination of preliminary results
- Can test signal prioritization (ranking) algorithm against known, ranked, cost deviations
- Can help gain clarity in understanding how the secondary effects of faults can affect end-use energy signals.

Testing of the Energy Signal Tool with synthetic fault data was undertaken to answer the following questions:

1. Does a test of the base-case model with no faults yield all green lights over various time periods throughout the year?
2. Which faults are significant enough for the Energy Signal Tool to display an action signal?
3. When an action signal is displayed for a faulted case, how consistently does it correspond to the end use associated with the root cause of the problem?
4. How closely correlated are the ranks of signal priority ratio and magnitude of the mean expected cost deviation across all faulted cases?

8.1 Synthetic Testing Methodology

Ten faults common to commercial buildings were applied to the model in OpenStudio or EnergyPlus, and the model was run with actual weather (for the period 04/01/13 to 03/31/14) to generate one year of hourly end-use results of "observed" data. Conversations with retail partner facilities management, as well as general knowledge of common building faults, led to the creation of a list of faults. "Expected" data (or conditional end-use probability distributions) were generated by sampling the ranges of the uncertain parameters in the fault-free baseline model. Just as in reality, a signal can be generated by placing the "observed" end-use consumption in the context of the expected range of consumption for a given time period.

The following are the details of the process to test and refine the theory of the energy signal tool.

1. Create as expected energy model based on apparent design and operational intent
2. Incorporate probable ranges of uncertainty into the parameters – no faults
3. Determine significant parameters to each end use (done already with regression)
4. Demonstrate the definition of risk tolerance thresholds with adjustments for end-use controllability and seasonal normalization.
5. Demonstrate the proper definition of cost matrices using the questions and established matrices for each end use being monitored (NGAS, REFR, HVAC, WBE)
6. Compute expected range of performance for nine time periods (Year, January, April, July, October, and the last week in each of these months) by perturbing uncertain parameters, with 1000 LHS simulation samples for each time period.
7. Test the point estimate of the "as expected" model by treating this output as "observed" consumption. The as expected model should yield all green signals. Tune the base model, the signal tool, and its inputs appropriately so that signals are green for the base model.
8. Inject the following ten typical faults into the model and save models separately:
    a. Outside air dampers stuck open (= 3x normal OA)
    b. Economizer broken (no economizer)
    c. Fans broken (belts snapped – 2 fans off)
    d. Refrigeration compressors operating at low efficiency
    e. Scheduling of operation errors (AHUs turn on at 3 a.m. instead of 7 a.m.)
    f. Thermostat setbacks not working (in all zones)
    g. Cooling set point overrides, set down an extra 1.8 ^{o}C (in all zones)
    h. Heat recovery is being bypassed
    i. Humidity controls fail or are shut off
    j. Temperature sensor float increases from 0.278 ^{o}C to 1.5 ^{o}C
9. Use these faulted models as measured data points for all four end uses and put these in the context of expected consumption to calculated state-space probabilities.
10. Generate signals for week/month/year in each scenario.
    a. Compute signal priorities – rank signals.
    b. Compare signal priorities to expected consumption deviation ranks.
11. High level analysis of the results.
    a. What was the skill of the signal?
    b. Were there secondary effects which made for noisy signal generation?
    c. Did priorities match with cost deviation ranks?

8.2 Tuning the ESTool with the As-Expected Energy Model Benchmark

To begin the synthetic testing process, the signal was first generated based on "measured" data coming from a completely fault-free building (which is also the base model for all fault injections). It was expected that the base model would produce end-use results near the mean of expected consumption distributions, thus yielding all "green" light action signal results. This was not the case at first. This led to iterations in defining the

"base" model until all signals gave "green" light results for all time periods and end uses. This iterative tool tuning approach is described below.

Table 15: Initial results of Baseline model signals for baseline performance compared baseline expectations

Iteration #1: Base model with all uncertain parameters set to likely values and nominal risk tolerance thresholds (M_1, M_2) set to (0.05, 0.10)

	WBE	NGAS	HVAC	REFR
Year Signal	Low	Low	high	
Month 1 Signal	Low	Low		high
Month 7 signal	low	Low	high	low

The initial model run yielded a large number of false signals, as seen in Table 15 above. It is a problem to have all parameter values set to expected values, because they have combined effects of influencing the building to consume much less natural gas than the mean of the expected distribution range. These parameters, specifically were:

1. Gas burner efficiency (can be lower but not higher than rated)
2. Air-to-air heat recovery effectiveness (can be lower but not higher than rated)
3. Minimum airflow rate (likely to be higher than intended, not lower)

The parameter distributions were characterized as being nearly right triangles, and it was only a matter of adjusting the baseline mode to represent a building as it is *likely to operate,* rather than a building under *ideal* operation. Therefore, the next iteration utilized a base model with uncertain parameter values selected at the *mean value* within the characterized distribution.

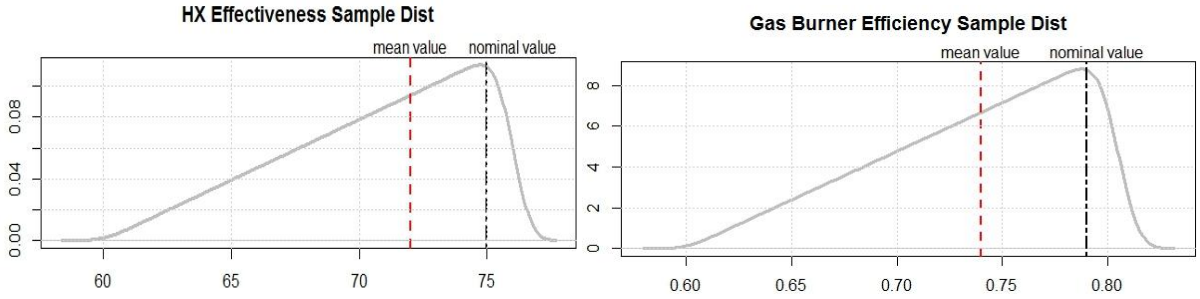


Figure 70: Examples of parameter distribution ranges and the difference between *expected* and *mean* values

Table 16: Baseline signals after second iteration of tool calibration

Iteration 2: Parameter values set to median values

	WBE	NGAS	HVAC	REFR
Year Signal	high	high		
Month 1 Signal	high	high		high
Month 7 signal	high	high	high	low

This had an effect on the signal results, and this parameter adjustment process was repeated several times more. Even a model with parameters set to mean values will exhibit end-use result bias resulting from parameter interactions. This became a manual optimization problem; with the lesson being that gathering as much descriptor data as possible before tuning the ESTool is paramount.

The next step in tuning the tool was to re-examine the cost matrices for each end use to adjust the signal bias. The natural gas end-use cost matrix had too strong of a bias for showing a signal calling for action in response to a "higher consumption than expected"

state. It was also determined that there had been a slight error in interpreting the minimum tolerance thresholds of 3% - which should have been a minimum of 3% of the "measured" values, not 3% probability mass. (The purpose of this minimum threshold was to account for normally occurring measurement error). The refrigeration end-use PDF had state space boundaries that had been too tightly constrained. Correcting this calculation error made excellent improvements to this end.

Table 17: Baseline signals after third iteration of tool calibration

Iteration 3: Cost matrices adjusted and minimum tolerance corrected to 3%

	WBE	NGAS	HVAC	REFR
Year Signal	high	high		
Month 1 Signal	high	high		high
Month 7 signal	high	high	high	low

This still resulted in a large number of action signals. A great deal of time had already been invested in adjusting uncertain parameter values in an attempt to obtain all green signals, without success. The next step was to calibrate the tolerance range of the tool; that is, raise the values of the M_1 and M_2 risk tolerance thresholds until all signals for each time period were green. Henze *et al.* (2015) had arbitrarily defined nominal risk tolerance threshold values of 5% and 10%. Adjusting M_1 up to 10%, and M_2 up to 20% achieved the goal of eliminating false alarms[14]. In other experiments, the user could still begin with lower nominal values for risk tolerance thresholds and test these first.

[14] Recall that M_1 and M_2 are later modified to reflect the degree of controllability of the parameters that influence each end use, and also modified further for any needed seasonal normalization factor (natural gas).

Table 18: Baseline signals after fourth tuning iteration

Iteration 4: M_1, M_2 set up to (0.10, 0.20)

	WBE	NGAS	HVAC	REFR
Year Signal				*SR = 12
Month 1 Signal				*SR =19
Month 7 signal				*SR = 54

This iteration produced all green lights; the baseline model was now operating within its innermost expected range for all end-uses. However, the consumption bounds of the extended risk tolerance thresholds are near the limits of the refrigeration energy distribution for all time periods. This resulted in green lights for the refrigeration end use, but with high signal ratios. The signal ratio is meant to be a rank-able indicator of signal severity to allow for prioritizing action within each signal type. It is the ratio of the largest to smallest term in the expected cost matrix: essentially showing the skewness of the tolerance thresholds superimposed upon the expected energy consumption range. Any value above ten is a high signal priority ratio. The high signal ratios associated with green signals (see

Table 18, above) indicated a problem with the scale of the minimum threshold regions as it compared to the overall magnitude of expected operation. It was concluded that the range of expected operation, which was between +/- 2.5 to 4% in magnitude for each time period, was too narrow. To correct this, another uncertain parameter was uncovered and added to the model (heat loss in the refrigeration gas cycle). This parameter addition widened the range of expected refrigeration energy use (also refer back to section 2.5).

Table 19: Baseline signals and signal priority ratios after the fifth round of tool calibration

Iteration 5: Parameter added to refrigeration end use

	WBE	NGAS	HVAC	REFR
Year Signal	SR = 1.7	SR = 2.1	SR = 2.2	SR = 3.2
Month 1 Signal	SR =2.1	SR = 1.9	SR = 1.8	SR =3.6
Month 7 signal	SR = 2	SR = 2.2	SR = 2.5	SR = 2.9

Finally, parameter uncertainty and seasonal normalization factors could be applied to each end use.

Table 20: Baseline model tolerance levels (yellow or green signal) for each end use after final tool calibration

	Whole Building Energy		Natural Gas		HVAC (Cool + Fans + Elec Heat)		Refrigeration	
Period:	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
Year	0.106	0.211	0.106	0.212	0.124	0.247	0.1	0.199
Month 1	0.106	0.211	0.106	0.212	0.141	0.282	0.1	0.199
Month 4	0.106	0.211	0.127	0.254	0.152	0.304	0.1	0.199
Month 7	0.106	0.211	0.169	0.339	0.124	0.247	0.1	0.199
Month 10	0.106	0.211	0.137	0.273	0.146	0.292	0.1	0.199

Table 20, above, gives the final magnitudes of tolerance interval for each end-use of the tool for the application of testing synthetic faults with the baseline model. These values were calculated including adjustments for controllability and seasonal variation in each

end use. These values of M_1 and M_2 were carried through into all the fault testing cases for climate zone 6a. Risk tolerance thresholds are unique for each building and location the Energy Signal Tool is applied to.

8.3 Fault Testing Results

Table 21 below gives a summary description of the faults tested. The nine test periods (Year, Month 1, Month 4, Month 7, Month 10, Week 4 [Jan 22-28], Week 17[Apr 23-29], Week 30 [Jul 23-29], Week 43 [Oct 22-28]) were selected to be exemplary of the four seasons in the Wisconsin climate. The testing included adjustments to the M_1 and M_2 thresholds for parameter controllability and seasonal variation. The results showed that the tool is quite capable of recognizing eight out of the ten faults tested. The results also demonstrate the value in separating action signals by end use, rather than just examining whole building energy. However, the signals generated often point to the secondary effects of the fault rather than pointing directly to the primary end use which contains the root cause of the fault. See especially fault J.

Table 21: Fault testing descriptions and modeling assumptions

Fault	Description	Modeling Assumptions
a	OA dampers stuck open (50% OA)	50% OA instead of 17% OA for DOAS System
b	Economizer broken (no ecomomizer)	No economizer operation - "Office" AHU only
c	Fans broken (belts snapped)	"Sales1" AHU and "Restroom" AHU fans off
d	Refrigeration compressors broken	Refrigeration rack "A" UA suction piping to 600 W/k
e	Controls scheduling errors: AHUs	All AHUs turn on at 3am instead of 7am
f	Thermostat setbacks not working	No night setback for heating thermostats
g	Cooling setpoint overrides	Cooling thermostat setpoint is set down 3oF
h	Heat recovery is being bypassed	Delete air-to-air heat exchanger on one AHU
i	Humidity controls fail or are shut off	EMS code for humidity control is disabled
j	Thermostat float temp increase	Increase float to 1.5 oC for all zones

Base: As expected base building									
End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 signal	Month 7 signal	week 30 signal	Month 10 signal	week 43 signal
Buildng Energy	1.9	1.5	1.4	1.9	2.3	4.1	6.2	2.2	low ; 2.1
Natural Gas	1.5	1.4	1.4	1.7	1.5	2.1	2.1	1.7	low ; 2.2
HVAC	1.6	1.7	1.9	2	2.1	1.6	1.7	2	2
Refrigeration	4.8	4	3.4	4.3	4.5	4.2	3.9	4.7	high ; 5.2

Comments: There was some noise in the base case, giving false signals in 3 of the 36 periods. A green signal means that the observed energy consumption is quite close to the mean of the expected consumption distribution. A model composed of parameters set to their mean values is not guaranteed to produce this result. In fact, there were some adjustments made to the parameter input values, within the given uncertainty ranges, in an attempt to meet the expected consumption for the maximum number of periods. This is important for the base case, since the signal results of all other faults will carry any bias produced by comparing the as-expected building to the probable distribution range of output. This process was time consuming. In a real building modeling scenario, there would be the same, necessary, challenge of matching the output of the base model to every period in the calibration period. As discussed above, this requires more than statistical parameter perturbation; it requires gathering more information from the facilities stakeholders.

Fault A: OA Dampers Stuck Open (50% OA instead of 17% OA)									
End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 signal	Month 7 signal	week 30 signal	Month 10 signal	week 43 signal
Whole Building Energy	high ; inf	high ; inf	high ; 1700	high ; inf	high ; inf	high ; 3.6	high ; 4.9	high ; inf	high ; inf
Natural Gas	high ; inf	high ; inf	high ; 1200	high ; inf	high ; inf	2.2	high ; 3.9	high ; inf	high ; inf
HVAC	high ; 17	high ; 580	high ; 1700	high ; 3.1	2.1	high ; 5.1	high ; 5.8	high ; 2.5	high ; 3.3
Refrigeration	4.8	low ; 14	low ; 17	high ; 4.4	high ; 5	4.1	high ; 5.1	high ; 4	high ; 4.3

Comments: Signals show strong indications of faults, as expected. A signal ratio of "Inf" signifies the measured value is completely outside the expected range generating by sampling. HVAC (Cooling, fans, VAV reheat) energy is affected most during peak heating or

peak cooling periods. Signals are not as strong in shoulder seasons, when OA is close to supply air temperature in this climate. Refrigeration gives a strong low signal in heating season because the heating system is not sized for this additional load, thus the indoor air temperature drops significantly, reducing the load on the refrigerators. It is clear that action should be taken right away.

Fault B: Economizer broken (no ecomomizer for one AHU)									
End Use	**Year Signal**	**Month 1 Signal**	**Week 4 Signal**	**Month 4 Signal**	**Week 17 signal**	**Month 7 signal**	**week 30 signal**	**Month 10 signal**	**week 43 signal**
Whole Buildng Energy	**1.9**	**1.5**	**1.4**	**1.9**	**2.3**	**4.2**	**6**	**2.2**	**low ; 2.1**
Natural Gas	**1.5**	**1.4**	**1.4**	**1.7**	**1.5**	**2.1**	**2.2**	**1.6**	**low ; 2.3**
HVAC	**1.6**	**1.7**	**1.9**	**2**	**2.2**	**1.6**	**high ; 2**	**2**	**1.9**
Refrigeration	**4.8**	**4**	**3.4**	**4.3**	**4.5**	**4.2**	**4**	**4.6**	**high ; 5.2**

Comments: The economizer would provide free cooling given the right set of outdoor air conditions. Only one AHU in the building operates with an economizer. This climate zone also has relatively few hours of outdoor air conditions during which economizer mode would activate. Thus, there is little change in energy use and it is permissible that action signals are not generated. It is certainly not expected that there are action signals during months when cooling demand is low (month 1, parts of months 4 and 10). The facility manager would be better off waiting for more issues to arise than to take action here.

Fault C: Fans broken (belts snapped) (Sales1 AHU and Restroom AHU fans off)									
End Use	**Year Signal**	**Month 1 Signal**	**Week 4 Signal**	**Month 4 Signal**	**Week 17 signal**	**Month 7 signal**	**week 30 signal**	**Month 10 signal**	**week 43 signal**
Whole Buildng Energy	**low ; 1300**	**low ; 430**	**low ; 580**	**low ; 350**	**low ; 73**	**low ; 5.4**	**low ; 6.3**	**low ; 85**	**low ; 3500**
Natural Gas	**low ; 2600**	**low ; 690**	**low ; 700**	**low ; 120**	**low ; 30**	**2.2**	**low ; 3.8**	**low ; 32**	**low ; 870**
HVAC	**low ; 91**	**low ; 4.5**	**low ; 3.8**	**low ; 210**	**low ; 580**	**low ; 120**	**low ; 270**	**low ; 4500**	**low ; 250**
Refrigeration	**4.8**	**4.1**	**3.4**	**3.9**	**4.5**	**4.2**	**3.9**	**4.7**	**high ; 4.8**

Comments: One major and one minor AHU have fans that are disabled. It was also assumed that outdoor airflow to the zones served by these two AHUs was eliminated. Fan energy is a relatively small component of total HVAC energy, even in shoulder seasons. A change in outdoor air flow rate is much more significant than fan energy, thus a signal is stronger for the natural gas end-use in the heating season. This fault has strong, consistent signals, and merits action.

Fault D: Refrigeration compressors broken (Refrigeration Suction UA up by 300% on two racks)									
End Use	**Year Signal**	**Month 1 Signal**	**Week 4 Signal**	**Month 4 Signal**	**Week 17 signal**	**Month 7 signal**	**week 30 signal**	**Month 10 signal**	**week 43 signal**
Whole Buildng Energy	**2.1**	**high ; 1.6**	**high ; 1.5**	**1.9**	**high ; 2.9**	**high ; 4.1**	**high ; 4.9**	**high ; 2.8**	**1.7**
Natural Gas	**1.5**	**1.4**	**1.4**	**1.7**	**1.5**	**2.1**	**2.2**	**1.6**	**low ; 2.3**
HVAC	**1.6**	**1.7**	**1.9**	**2**	**2.2**	**1.6**	**1.7**	**2**	**1.9**
Refrigeration	**high ; Inf**	**high ; Inf**	**high ; Inf**	**high ; Inf**	**high ; Inf**	**high ; Inf**	**high ; Inf**	**high ; Inf**	**high ; Inf**

Comments: The increased refrigeration energy use is a larger component of whole building energy during the cooling months; thus, this end use begins to affect whole building energy use in July. A continuous signal of "Inf" for the refrigeration energy use is a rightful call to immediate action.

Fault E: Scheduling of operation errors: (All AHUs turn on at 3am instead of 7am)									
End Use	**Year Signal**	**Month 1 Signal**	**Week 4 Signal**	**Month 4 Signal**	**Week 17 Signal**	**Month 7 Signal**	**Week 30 Signal**	**Month 10 Signal**	**Week 43 Signal**
Whole Buildng Energy	**high ; 15**	**high ; 11**	**high ; 8.6**	**high ; 17**	**high ; 52**	**high ; 17**	**high ; 12**	**high ; 28**	**high ; 11**
Natural Gas	**high ; 9.3**	**high ; 7.9**	**high ; 6**	**high ; 9.6**	**high ; 26**	**high ;7.3**	**high ; 3.6**	**high ; 13**	**high ; 5.2**
HVAC	**high ; 39**	**high ; 39**	**high ; 37**	**high ; 32**	**high ; 17**	**high ; 47**	**high ; 28**	**high ; 15**	**high ; 24**
Refrigeration	**high ; 4.9**	**high ; 4**	**high ; 4.1**	**high ; 3.9**	**high ; 3.9**	**3.8**	**high ; 4.4**	**high ; 3.8**	**high ; 14**

Comments: This operational error would have all systems running as if the store were occupied 4 hours earlier than intended. This includes thermostat adjustments, but does not affect outdoor air quantity, since this is already delivered at a constant rate.

Signals across the year were in the quite high. The presence of heating and cooling related signals indicates simultaneous heating and cooling was expected. This fault had been present in the case study building, as discussed in a previous section, and was only discovered through a retro-commissioning effort. The consistent signals generated in all months by the Energy Signal Tool would have prompted action to correct this problem sooner than actually occurred.

Fault F: Thermostat setbacks not working : (no night setbacks or setups)									
End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal
Whole Buildng Energy	high ; 200	high ; 74	high ; 43	high ; 170	high ; 270	high ; 29	high ; 13	high ; 230	high ; 94
Natural Gas	high ; 95	high ; 48	high ; 26	high ; 83	high ; 130	high ; 7.9	high ; 5.5	high ; 140	high ; 33
HVAC	high ; 76	high ; 80	high ; 66	high ; 33	high ; 20	high ; 120	high ; 28	high ; 17	high ; 41
Refrigeration	3.9	high ; 4.6	high ; 5.7	3.7	3.7	3.7	3.9	4	high ; 13

Comments: Strong signals were generated for every month of the year. Once again, the presence of heating and cooling related signals indicates simultaneous heating and cooling was expected. This fault would be detected very quickly.

Fault G: Cooling setpoint overrides: (Cooling T;stats set down 1.8oC)									
End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal
Whole Buildng Energy	high ; 5.6	high ; 3.8	high ; 4	high ; 2.9	high ; 4.1	high ; 180	high ; 500	high ; 5.3	high ; 1.9
Natural Gas	high ; 2.8	high ; 3.2	high ; 3.1	high ; 2	high ; 2.2	high ; 52	high ; 10	high ; 2.6	1.6
HVAC	high ; 370	high ; 18	high ; 13	high ; 32	high ; 23	high ; 250	high ; Inf	high ; 43	high ; 21
Refrigeration	4.2	3.7	3	3.6	3.7	low ; 4.4	4.2	4.1	high ; 3.6

Comments: This fault would have a high ratio of savings to action cost. The energy signal tool was quick to send a strong signal detecting cooling setpoint overrides. This confirms that the store has winter cooling.

Fault H: Heat recovery is being bypassed: (Delete Air;to;Air HX on one AHU)									
End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 signal	Month 7 signal	week 30 signal	Month 10 signal	week 43 signal
Whole Buildng Energy	high ; 62	high ; 23	high ; 11	high ; 1700	high ; 580	3.8	high ; 4.3	high ; 870	high ; 500
Natural Gas	high ; 160	high ; 26	high ; 11	high ; Inf	high ; Inf	high ; 2.6	high ; 3	high ; Inf	high ; 870
HVAC	high ; 1.9	high ; 2.4	high ; 2.5	1.9	2.1	high ; 1.8	high ; 2.1	1.8	1.8
Refrigeration	4.6	3.9	3.4	high ; 4	high ; 4.6	4.1	high ; 4.9	4.3	high ; 6

Comments: The AHUs with heat recovery have them because they are economical and impactful on the energy consumption of the building. These results show an expected sharp increase in natural gas use when lacking a major heat recovery element. This is especially true in months when OA-Indoor temperature difference dominates the heating load (shoulder seasons). This fault would result in needed action after a short monitoring period.

Fault I:Humidity controls fail or are shut off: (EMS code for humidity Control is disabled)									
End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 signal	Month 7 signal	week 30 signal	Month 10 signal	week 43 signal
Whole Buildng Energy	low ; 2.4	1.5	1.5	1.9	2.3	low ; 56	low ; 6.3	low ; 3	low ; 2.1
Natural Gas	low ; 1.9	1.4	1.4	1.7	1.5	low ; 51	low ; 140	low ; 2.3	low ; 2.3
HVAC	low ; 24	1.7	1.8	2	2.1	low ; Inf	low ; 71	low ; 26	2
Refrigeration	high ; 4.9	3.9	high ; 3.5	4.2	4.4	high ; 6.6	high ; 55	high ; 4.5	high ; 5.2

Comments: Humidity control is a huge load on the cooling system during warm months. Without it, HVAC energy use drops significantly. Without humidity controls, space humidity levels are allowed to drift up, which puts a greater load on the refrigerated cases. Examining the end-use signals in this case might lead to confusion that the root cause lies partially in the refrigeration system. However, when multiple time periods are considered and the energy manager observes that refrigeration energy use increases only in the warmest months, it would be quite clear that the refrigeration fault is due to an interactive

effect between systems. This fault is a case for having distinct signals by end use; especially at the level of the entire year. The signal priority for action on whole building energy is quite low, but only because the SR values for HVAC and refrigeration are strongly opposing each other. It was certainly not expected that humidity control failure would be a fault in the periods of month1, week 4, or week 43 according to the AMY weather data (see appendix). Many false positive signals for the natural gas end use (being too low) indicate that overcooling of the building was an expected side effect of dehumidification.

Fault J:Thermostat Float temp increase: (Increase float to 1.5 oC for all zones)									
End Use	**Year Signal**	**Month 1 Signal**	**Week 4 Signal**	**Month 4 Signal**	**Week 17 signal**	**Month 7 signal**	**week 30 signal**	**Month 10 signal**	**week 43 signal**
Whole Buildng Energy	**2**	**1.5**	**high ; 1.6**	**1.9**	**low ; 3.3**	**low ; 4.2**	**low ; 4.2**	**low ; 2.7**	**low ; 2.2**
Natural Gas	**low ; 1.5**	**1.5**	**high ; 1.6**	**1.7**	**low ; 3.2**	**low ; 2.8**	**low ; 4.3**	**low ; 2.2**	**low ; 2.4**
HVAC	**low ; 5**	**1.8**	**1.8**	**low ; 2.9**	**low ; 6.3**	**low ; 36**	**low ; 11**	**low ; 5.5**	**low ; 2.5**
Refrigeration	**4.6**	**3.8**	**high ; 3.6**	**high ; 4**	**high ; 4.1**	**4.1**	**high ; 4.4**	**4**	**high ; 10**

Comments: A higher tolerable range in indoor temperature float would decrease heating and cooling energy. Conversely, the refrigeration systems are overworked as the cases go through unnecessary temperature swings that engage added defrost energy. Thus, opposite end-use signals are generated. Once again, this fault is a great case for breaking the building energy signal into end-use components. The priority of *combined* whole building energy action is comparatively low due to the combination of energy increase in one end use and energy decrease in another. With only whole building energy metering, this fault might have been placed low in priority.

8.4 Signal Classification Skill Metric

Classification skill tests the aptitude of the Energy Signal Tool as an application of fault diagnostics. An advantage of using synthetically generated fault data is that the accuracy of signal classification can be assessed easily. This application has five possible

classifications for the energy signal, as explained above. Rather than assessing the skill of the tool here based on the signal magnitude, classification skill shall be based on whether or not a signal corresponds to a fault. Florita (2014) proposed a way to classify each signal in one of the following four categories, based on known conditions from the synthetically generated faults:

TP: *True positive;* a signal is given and it corresponds to the root cause end-use fault

FP: *False positive;* a signal is given which does not correspond to the root cause

TN: *True negative;* no action is prescribed, and the end-use is not faulted

FN: *False negative;* no action is prescribed, when actually the end-use is faulted

From the enumeration of these classification categories, two useful metrics arise:

1. Frequency of Detection (FOD); *what percentage of faults are identified?*
2. False Alarm Ratio (FAR); *what percentage of signals call for un-needed action?*

Where,

Equation 29: Frequency of detection

$$FOD = \frac{TP}{TP + FN}$$

Equation 30: False alarm ratio

$$FAR = \frac{FP}{FP + TP + TN}$$

In this work, skill metrics shall be applied to all four end uses of whole building energy, natural gas, HVAC, and refrigeration as described previously. In terms of fault detection capability, FOD is significant for its measure of sensitivity to faults. A false negative signal may be produced when the range of risk tolerance is wide enough to absorb the anomaly generated by a fault. A false negative in whole building energy may be produced when two end-uses see opposite changes in energy consumption and cancel out. To a lesser degree, false negatives may also be a product of the fault having effects on components of the building which negate each other. FOD is not applicable in the case of a

fault-free building. FAR is essentially estimating the percentage of spurious actions prescribed by the tool. For fault detection in facilities management, this can be compared to the percentage of maintenance calls that result in no further action. In some cases, this comes in the form of actions prescribed for too many end-uses; which is why FP is expressed as a ratio compared to all true signals.

This skill metric was assessed in each of the nine time periods, for each of the fault testing cases. The following is an example of how skill metrics would be tallied for each end use in each period of analysis. Since the example is of an outdoor air quantity fault, *true positive* signals for action are expected in all end uses except refrigeration. Outdoor air quantities are related equally to HVAC and natural gas end-use energy. Any action taken on systems related to refrigeration would certainly not address the root cause of the problem. Additionally, this fault is expected to generate signals in all months of the year.

Fault A: OA Dampers Stuck Open (50% OA instead of 17% OA)										
End Use	**Year Signal**	**Month 1 Signal**	**Week 4 Signal**	**Month 4 Signal**	**Week 17 signal**	**Month 7 signal**	**week 30 signal**	**Month 10 signal**	**week 43 signal**	
Whole Building	high - inf	high - inf	high - inf	high - inf	high - inf	high - 13	high - 260	high - inf	high - inf	
Natural Gas	high - inf	high - inf	high - inf	high - inf	high - inf	high - 28	high - inf	high - inf	high - inf	
HVAC	high - 15	high - 308	high - 436	2.8	2.6	high - 10	high - 7.7	2.7	2.5	
Refrigeration	4	low - 24	low - 27	3.5	high - 3.7	3.8	high - 4.2	3.9	high - 4.3	**Σ:**
TP	*3*	*3*	*3*	*2*	*2*	*3*	*3*	*2*	*2*	***23***
FP	*0*	*1*	*1*	*0*	*1*	*0*	*1*	*0*	*1*	***5***
TN	*1*	*0*	*0*	*1*	*0*	*1*	*0*	*1*	*0*	***4***
FN	*0*	*0*		*1*	*1*	*0*	*0*	*1*	*1*	***4***

Figure 71: Example of classification skill calculation

In this case, since it is known in advance what the signals should be, the results of the ESTool shall be compared to known truth values for signals. For several faults, as noted above in the signal commentary for each, it was not reasonable to expect an action signal for certain months of the year. These faults do not affect the operational energy in some seasons, depending on climate. To summarize, these were:

Fault B, Economizer Control: *only during cooling season*

Fault G, Cooling T-stat set points: *only during cooling season*

Fault H, Heat exchanger failure: *only in seasons with outdoor/indoor air temperature differences greater than 5 ^{o}C*

Fault I, Humidity Control: *only during cooling season*

Accordingly, there were seasons with no *false negative* classification possibilities for faults with significant seasonality, such as those listed above. If there is no chance of observing a fault in certain seasons, then its skill should not be penalized for lack of detection in these seasons. Thus the denominator for the FOD metric was less than other faults. The exception for fault H did not apply to the case study location, but could apply to a climate with extended periods of mild weather. Care was taken to address this for each climate zone.

8.4.1 Summary of Skill Testing Results

Testing results for the skill metric are given below when nominal values of M_1 and M_2 are set to 0.1 and 0.2, respectively. All metrics should be compared to those of the base case, since there was a small bias of false alarms in the base model already – as discussed above.

Table 22: Testing results summary giving classification skill metrics with [M_1, M_2] = [0.1, 0.2]

Climate Zone 6a	BASE	Fault A	*Fault B*	Fault C	Fault D	Fault E	Fault F	Fault G	Fault H	Fault I	Fault J
FOD	N/A	0.85	*0.11*	0.96	0.83	1.00	1.00	0.90	0.81	0.67	0.74
FAR	0.08	0.16	*0.11*	0.03	0.03	0.22	0.08	0.28	0.13	0.38	0.17

All faults except fault "B" had a high frequency of detection. While there may be false alarms as a result of monitoring multiple end-uses, this is better than having a single indicator of whole building performance that is diluted by opposing changes in energy consumption. False alarms were mainly due to interactive effects that occur naturally in a building. For example, when humidity controls are disabled, it is normal that refrigeration energy use increases. However, acting directly on refrigeration systems in the building

would certainly not identify the root cause of the problem, and would be a waste of money to hire a refrigeration systems technician. Therefore, this signal is classified as false positive. Refer to the appendix for complete tallies of skill metrics for all faults.

8.5 Applying the Energy Signal Tool in Different Climate Zones

Testing the Energy Signal Tool with the same base model, and identical faults but located in other climate zones is a way to test the robustness of the tool for widespread, rapid deployment across a building portfolio. The case study building is located in climate zone 6a (cold). The two alternate testing locations were selected for diversity of climate to be in Portland (zone 4c, mixed marine) and Atlanta (zone 3a, warm, moist). In the figure below, the location of the case-study building is shown with a black star, and the two alternate testing locations are shown with red stars. For each new test location, a recent typical meteorological year (TMY3) file was used to simulate an annual weather cycle.

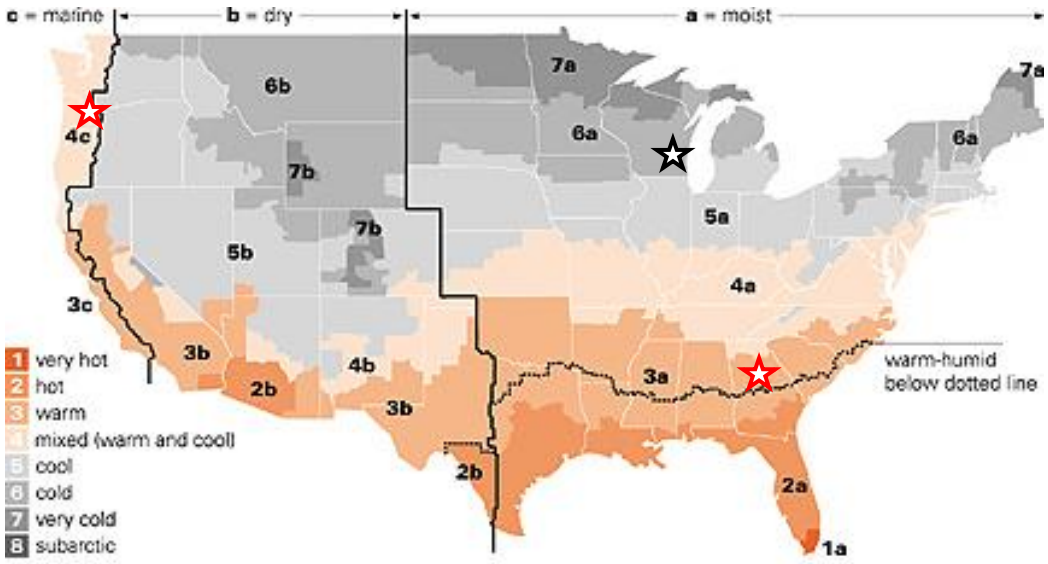


Figure 72: ASHRAE Climate Zones *and test locations*

(image source: http://www.achrnews.com/)

Other than using different weather files, the modeling process for the two new climate zones was identical to that in the case study building location. The same physical model and schedules were used. All HVAC systems including cooling coils, heating burners,

and fans were auto sized to meet loads. Identical rates of ventilation airflow were used. An NREL design support document made a similar analysis on big-box retail store design across these climate zones (Bonnema *et. al.*, 2013). The authors used auto sizing of HVAC equipment, and changed only the wall insulation level for best practice for different buildings. That insulation level required for climate zone 6a is slightly greater than that required for climate zones 3a and 4c; thus, it was left unchanged in this work to allow for more direct comparison of signal results. For other types of buildings, other parameters such as window type would change by climate zone as well, and could impact the parameter significance rankings. In this case however, fenestration is a very minimal component of the building and was left untouched. The table below shows the difference in end-use energy consumption profile for the same building in the three different climates.

Table 23: Summary of energy end use by climate zone

Annual Energy Use by Climate Zone (kBtu/sq ft-yr)			
End-Use:	***3a***	***4c***	***6a***
Heating (Elec)	0.02	0.03	0.53
Cooling	4.43	1.94	2.08
Interior Lighting	15.70	15.70	15.70
Exterior Lighting	1.46	1.61	2.24
Interior Equipment	8.49	8.49	8.49
Fans	2.42	2.36	2.42
Refrigeration	12.62	11.75	12.20
Nat Gas	6.82	12.35	23.10
Total:	**52.0**	**54.2**	**66.8**

To begin the testing process in the new climate zones, a new global sensitivity analysis was carried out to identify which parameters were most significant to each end use by climate zone. A summary of these findings, as compared to those of the original case study location is given below. Original parameter distribution assumptions (rather than the

updated distribution ranges generated from the calibration process[15]) were used to generate expected results for the new climate zones. Then, tuning of base case assumptions for parameter values was done produce no false positives for the "Base" case across the four representative months used in the study. Finally, the faulted scenarios from before were simulated in each climate zone.

8.5.1 Summary of Energy Signal Differences by Climate Zone

When in the context of different climates, there are several underpinnings to the Energy Signal Tool processor that change. Those include:

- Seasonal normalization factors due to seasonal variation in end-use magnitude
- Parameter significance for each end use, and the impact on end-use controllability
- Fault impact due to weather
- Answers to cost matrix input questions

As discussed earlier, seasonal normalization factors are used to dampen signals for periods when annualized consumption is much below the expected annual maximum. They have the greatest effect in climates with extremes, such as the upper-Midwest region and climate zone 6a. In climate zones 3a and 4c, there is less of a need for seasonal normalization of risk tolerance. The exception to this is the increased spread in monthly HVAC energy in climate zone 3a, due to very high summer cooling loads.

In each climate zone, a multivariate regression analysis revealed distinct results for the global parameter sensitivity analysis. A summary of these results is given in Figure 73 a,b below.

[15] There was no data against which to calibrate the models of the new climate zones. As illustrated in the previous section, the use of original vs. updated parameter uncertainty distributions did not alter the end-use consumption variance significantly. Thus it remains a fair comparison among the different climate zones.

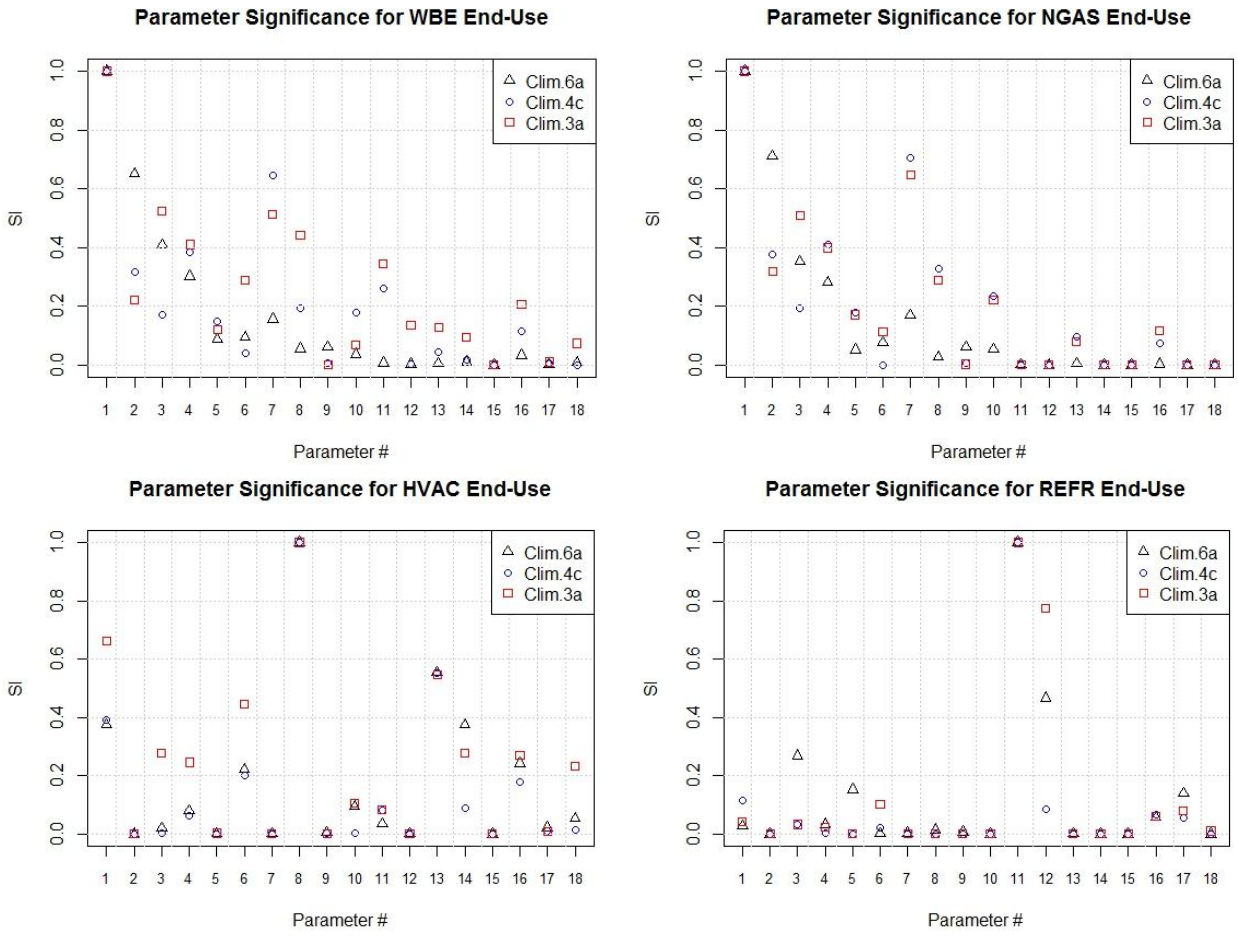


Figure 73 a,b: Comparison of parameter significance index by climate zone. Key given below.

Table 24: Key of parameter names in figure above

Par #:	*Description:*
1	Space infiltration reduction (%)
2	Gas Burner Eff (%)
3	Heating Setpoint (^{o}C)
4	Wall/Roof Conductivity multiplier
5	Heating supply air temp (^{o}C)
6	Heating sizing factor
7	HX Sensible Effectiveness (%)
8	Fan Pressure Rise (multiplier)
9	Minimum system outside air ratio
10	Rated roof absorptance
11	Refrig suction piping UA value
12	Condensing temp minimum, delta
13	Fan Efficiency (%)
14	DX Cool Coil COP
15	Anti-sweat energy multiplier
16	T-Stat float heat/cool (^{o}C)
17	Operating walkin ref temp (^{o}C)
18	Cooling Setpoint (^{o}C)

There were numerous parameters with minor differences and several with notable major differences among significance indices among the distinct climate zones. Gas burner efficiency (par #2) is of course more significant in colder climate zones. The sensible effectiveness of the air-to-air heat recovery device (par #7) was found to be much less significant in the cold climate. This is due to the fact that the plate heat exchanger operates with a frost prevention threshold of 1.7 ^{0}C. There are many more hours above this level, but below the heating balance point in climates 4c and 3a than occur in 6a. Fan pressure rise (par #8) was more significant in climate zones 4c and 3a mainly due to the larger percentage of total energy it comprises. Refrigeration minimum condensing temperature (par #12) was found to be more significant in the warmer climate zones in proportion to the number of annual hours with warm outdoor air temperatures.

The differences in parameter significance for each climate zone resulted in a new ratio of *controllable* to *uncontrollable* parameter composition for each end use (see previous section). Those results are summarized in the table below:

Table 25: Summary of end-use controllability ratios for each climate zone

	Whole Building Energy	Natural Gas	HVAC	Refrigeration
CLIMATE 3A				
Σ C:	1.974	1.987	1.179	0.960
Σ U:	2.615	1.881	2.981	1.170
U/(U+C)	**0.570**	***0.486***	**0.717**	**0.549**
CLIMATE 4C				
Σ C:	1.611	1.771	0.300	0.245
Σ U:	1.920	1.837	2.291	1.141
U/(U+C)	**0.544**	**0.509**	***0.884***	**0.824**
CLIMATE 6A				
Σ C:	1.298	1.235	0.817	1.096
Σ U:	1.623	1.562	2.270	1.077
U/(U+C)	**0.556**	**0.558**	**0.735**	**0.496**

In Table 25 above, the sum of scaled parameter influence factors controllability and uncontrollable parameters comprising each end-use is given. Once again, these values are obtained by normalizing the SI values of each parameter, obtained by the multivariate regression global sensitivity analysis, to the maximum in each end-use, and then summing values separately for parameters classified as controllable or uncontrollable to the facilities manager. The function *U/(U+C)* expresses the degree of *uncontrollability the facilities manager has* over affecting each end use. End-uses with a higher degree of uncontrollability are given more room for risk tolerance, since there are less likely to be effective actions that can be taken to address problems with unexpected energy consumption. The building in climate zone 4c has a notably higher degree of uncontrollability for HVAC and refrigeration end-uses. This is due to the mild outdoor temperatures in this climate that reduce the significance of controllable parameters such as cooling COP and refrigeration condensing temperature.

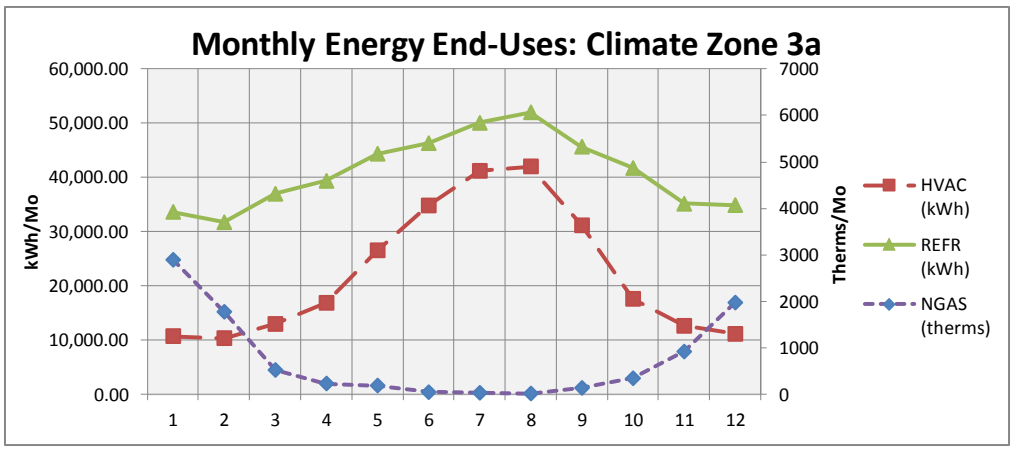
The cost matrices were slightly different for some of the end-uses across the three climate zones. For example, where cooling energy has a high ratio of maintenance to operational cost in climate zone 6a, the reverse is true in climate zone 3a. Conversely, the

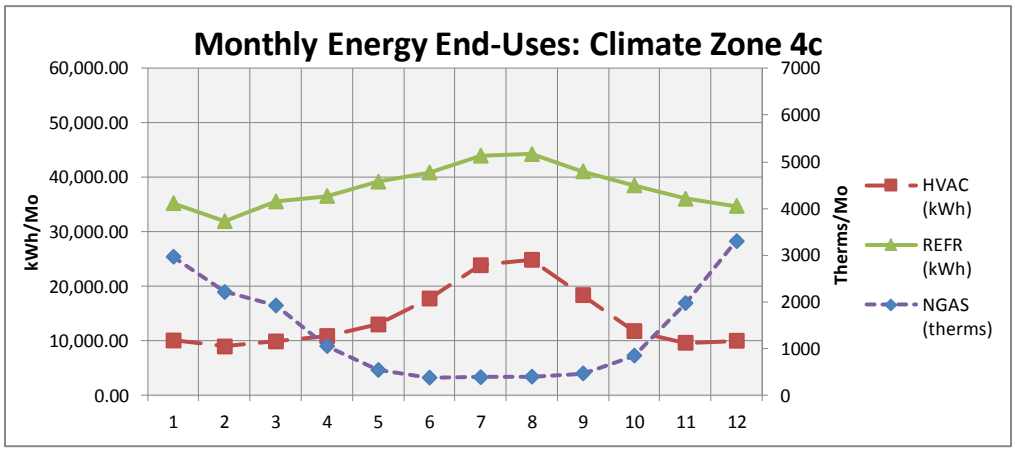
cost to maintain natural gas heating equipment is low compared to its use in the cold climate. Answers to question 2 are identical for all climate zones, representing consistent organizational goals at all stores. Table 26 below gives the expected answers provided by facilities management to the two cost matrix definition questions.

Table 26: Answers to cost matrix definition questions for three climate zones

Questions about End Uses:	*Climate*	Whole Bldg Energy	Natural Gas	HVAC	Refrigeration
1. Is the ratio of (cost to maintain) / (cost to operate) **high or low**?	Climate 3a	low	high	low	high
	Climate 4c	low	neutral	neutral	high
	Climate 6a	low	low	neutral	high
2. Does **high or low** energy in this end-use indicate a serious problem relating to operation, comfort, or product stability?	Climate 3a	neutral	neutral	low	low
	Climate 4c	neutral	neutral	low	low
	Climate 6a	neutral	neutral	low	low

The following figures illustrate the seasonal variation in the major end uses of the buildings in each climate zone. Note that climate zone 6a exhibits the greatest variation in monthly natural gas use. Climate zone 3a exhibits the greatest variation in monthly HVAC, as well as refrigeration energy use. Recall from the previous section that a normalization factor is applied to dampen the tendency for generating an action signal in off-peak end use consumption periods.





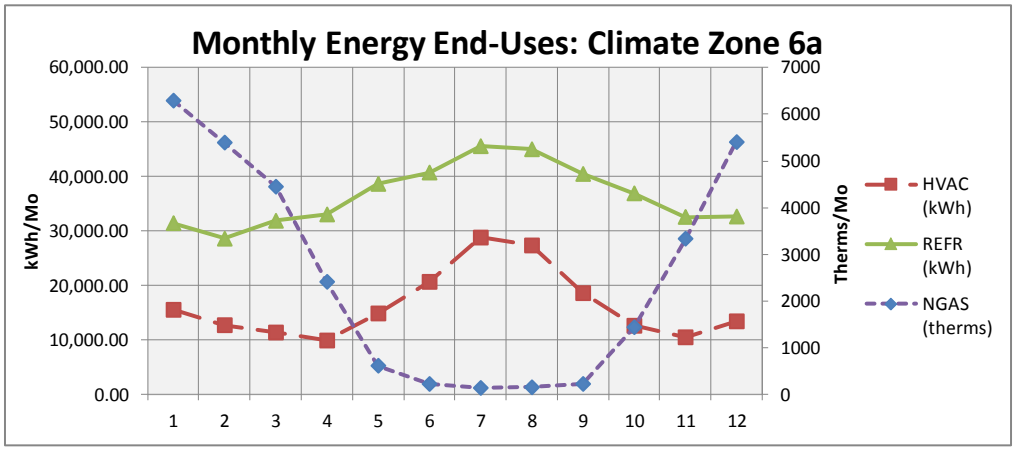
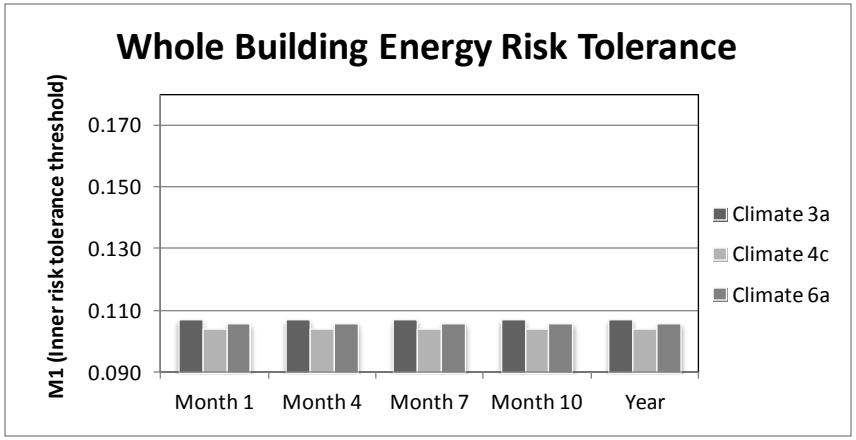
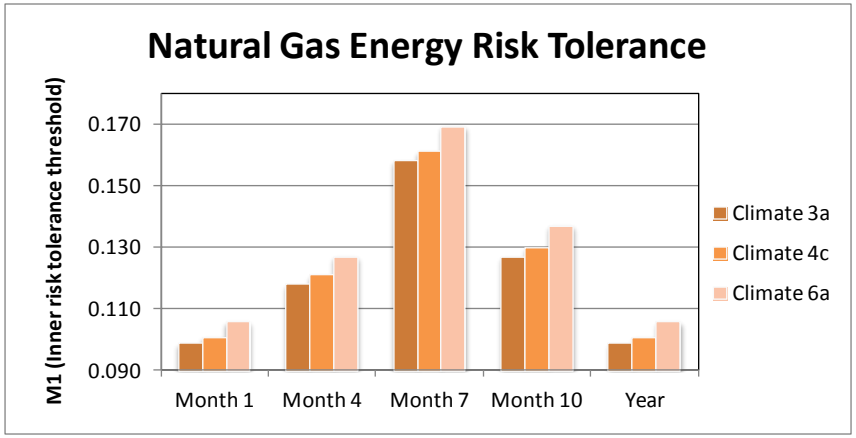
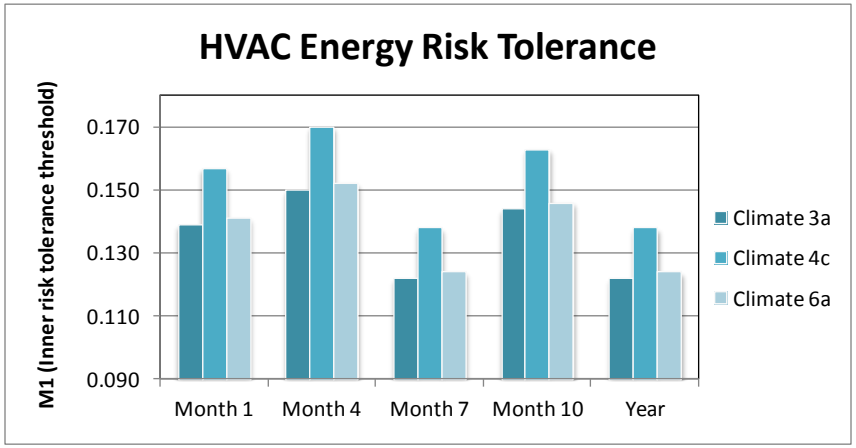


Figure 74 a, b, c: Monthly energy end-use for each climate zone

The combination of different ratios uncontrollability and seasonal end-use variation result in a new set of adjusted M_1 and M_2 risk tolerance threshold values for the end uses of each climate zone. Figure 75 below gives values of M_1 for each case (where M_2 is always equal to $M_1 * 2$).







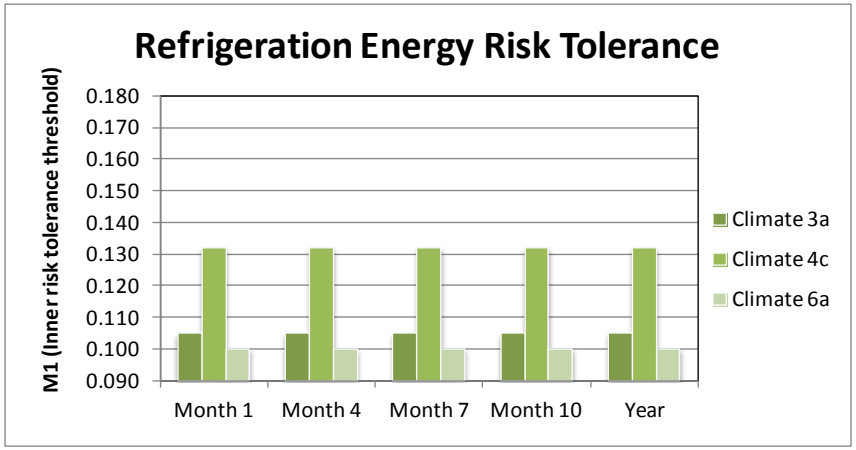


Figure 75 a, b, c, d: Comparison of inner risk tolerance threshold values (M_1) for end-uses in each climate zone

The figures above illustrate several differences between adjusted risk tolerance thresholds for each climate zone. First, the colder climates have greater annual variation in risk tolerance for natural gas use due to seasonal normalization. For the HVAC energy signals, the risk tolerance threshold values are generally lower in the more extreme climates (3a and 6a), while magnitude of seasonal variation is similar. Overall, it can be concluded that the Energy Signal Tool is expected to be more sensitive for climate zones 3a and 6a than for the mild climate zone 4c. In other words, it is likely that the FOD metric will be lowest for climate zone 4c.

The exceptions to allowing *false negative* signal classification noted above also apply to several periods for the new climate zones. That is, it is completely unreasonable to expect the Energy Signal Tool to detect certain faults with significant seasonality. In addition to the exceptions mentioned for the case study location, it is expected that heat exchanger failure with not generate a signal in expanded periods of mild weather. The table below gives a summary of these exceptions in reference to the fault testing periods and weather data plots in the appendix.

Table 27: Exceptions to possible signals for rule-based fault classification

Fault:	Climate 3a	Climate 4c	Climate 6a
Fault B – Economizer (HVAC)	-	Month 1, Wk 4, Wk 43	Month 1, Wk 4, Wk 43
Fault G – Cool T-stat (HVAC)	-	Month 1, Wk 4, Wk 43	Month 1, Wk 4, Wk 43
Fault H – HX Failure (HVAC or natural gas)	Wk 43	Wk 17, Month 10, Wk 43	-
Fault I – RH Control (HVAC)	Wk 4	Month 1, Wk 4, Wk 43	Month 1, Wk 4, Wk 43

8.5.2 Results of Energy Signal Tool Testing on Different Climate Zones

All ten faults were modeled for the two additional climate zones in the same manner as done for the case study location. This generated two additional sets of synthetic fault

data. Parameter values in the baseline model were adjusted to minimize baseline FAR across all monitoring periods. The Energy Signal Tool was then applied to each of these cases. Complete results of signal output can be found in the appendix. Skill metrics were assessed for the two additional climate zones, taking into account the signal exceptions noted in Table 27 above. The table below gives a summary of skill metrics.

Table 28: Summary of classification skill metrics for three climate zones

Climate Zone 3a	BASE	Fault A	*Fault B*	Fault C	Fault D	Fault E	Fault F	Fault G	Fault H	Fault I	Fault J
FOD	N/A	0.81	*0.11*	0.81	0.93	0.70	0.81	0.94	0.74	0.94	0.74
FAR	0.14	0.06	*0.20*	0.13	0.11	0.04	0.03	0.40	0.06	0.46	0.07
Climate Zone 4c	**BASE**	**Fault A**	***Fault B***	**Fault C**	**Fault D**	***Fault E***	***Fault F***	**Fault G**	**Fault H**	**Fault I**	**Fault J**
FOD	N/A	0.78	*0.08*	0.86	0.94	*0.56*	*0.60*	1.00	0.90	0.79	0.70
FAR	0.14	0.00	*0.16*	0.38	0.09	*0.08*	*0.11*	0.47	0.00	0.24	0.00
Climate Zone 6a	**BASE**	**Fault A**	***Fault B***	**Fault C**	**Fault D**	**Fault E**	**Fault F**	**Fault G**	**Fault H**	**Fault I**	**Fault J**
FOD	N/A	0.85	*0.11*	0.96	0.83	1.00	1.00	0.90	0.81	0.67	0.74
FAR	0.08	0.16	*0.11*	0.03	0.03	0.22	0.08	0.28	0.13	0.38	0.17

The signal tool was more sensitive to schedule-related faults in model representing the cold climate zone. This is due to a more constant need for indoor conditioning. As expected, the tool was more sensitive to detecting faults with the humidification system in the warm moist climate zones. The heat exchanger bypass fault was detected most often in the mild climate 4a, since the heat recovery device will often make the difference between using additional air conditioning or none at all. The signal tool was not sufficiently sensitive to detecting a fault in the economizer of one air-handling unit in the store. It also had a low frequency of detection for schedule-related faults in the mild climate 4c. This can be attributed to the reduced need for environmental conditioning during the nighttime hours of this climate. See the appendix for a complete print of fault testing signal generation results for each time period.

8.5.3 Sensitivity Ananlysis for Improving Skill

It was expected that the frequency of detection would be greater for some faults than others. Indeed, one of the main purposes of the tool is to set risk tolerance thresholds and prioritize actions based on organizational goals. However, to investigate the question of how sensitive the energy signal tool would have to be in order to identify each fault, a small sensitivity analysis was performed. For each scenario where FOD had been less than 66%, the values of M_1 and M_2 were reduced from their nominal values of 0.1 and 0.2, respectively, until the point where the FOD of the fault scenario increased to a passing score of 66% or greater.

Table 29: Sensitivity analysis results summary for improving skill

	Initial Skill: (M_1=0.1, M_2 = 0.2)		*Improved Skill*		*Final Risk Thresholds*	
Fault:	**FOD**	**FAR**	**FOD**	**FAR**	**M_1**	**M_2**
Fault B, Climate 6a	*0.11*	*0.11*	*0.22*	*0.28*	*0.03*	*0.06*
Fault E, Climate 4c	0.58	0.04	0.58	0.11	0.05	0.1
Fault F, Climate 4c	*0.60*	*0.11*	*0.63*	*0.14*	*0.05*	*0.1*

The skill metrics were not improved for any of the faults. In the case of fault B, a single economizer on a small air handling unit is not significant to any end-use. Faults E and F did not produce noticeable changes in the HVAC energy in any time period. This may be because fans on cycle on an infrequent basis at night to meet loads in an empty building. In addition, climate zone 4c is quite mild, meaning that little conditioning is needed. In all of these cases, reducing the tolerance threshold for risk was detrimental in that it increased the false alarm ratio. Higher quality sub-metering equipment, with lower standard measurement error, would be required to monitor for faults such as these.

8.6 Portfolio Level Analysis: Ranking Faults by Signal Priority Ratio

The ten distinct faults across three climate zones can collectively be thought of as 30 unique buildings in three climate zones. This amount of data is sufficient to explore the portfolio maintenance prioritization support that the Energy Signal Tool provides. Rather than displaying 30 sets of traffic lights, we can leverage the quantitative output of the ESTool for comparison of how the data processing algorithms prioritize actions in response to faults. The two quantitative outputs of the Energy Signal Tool are: Signal Priority Ratio (how strong the signal is based on both energy cost and subjective costs, Equation 27) and Mean Cost Deviation (mean of 95 percentile cost difference between expected and observed, Equation 28). When prioritizing in terms of the Signal Ratio, it is only valid to make comparisons among signals of the same end-use. The Signal Ratio is just a scalar, not a quantity whose magnitude has any direct meaning other than making rankings of actions to take for one end use; which can be made across multiple buildings. All cost deviations have been converted to absolute value for ease of comparison.

The following figures show the results of the top 15 faults by end use. They are ranked in ascending order of priority according to Signal Ratio (shown at right), while a comparison of deviation from mean expected cost is shown at left. Both scales have been transformed by the logarithmic function to encompass a better view of the entire range of signal strength and cost deviation values. Not surprisingly, the outdoor air damper fault (Fault A) produced a cost deviation in climate 6a that ranked higher than any other fault for the first three end-uses. Fan malfunction (Fault C) exhibited the most variability in terms of action urgency across the three climate zones. This can be explained by the fact that in the mild climate 4c, the main function of the fan is to provide ventilation air to the building – rather than heating or cooling. Meanwhile, the fans serve in large part to provide heating to climate 6a (highest priority in natural gas end-use of the three), and provide cooling to climate 3a.

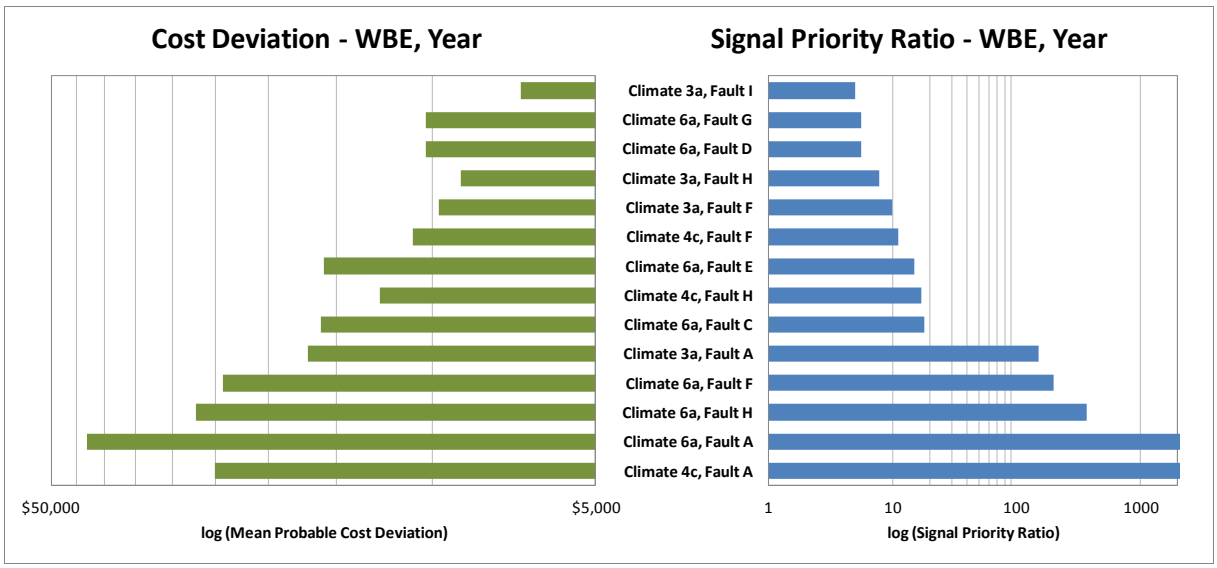


Figure 76: Top 15 faults in yearly whole-building energy end-use, in ascending signal priority ratio

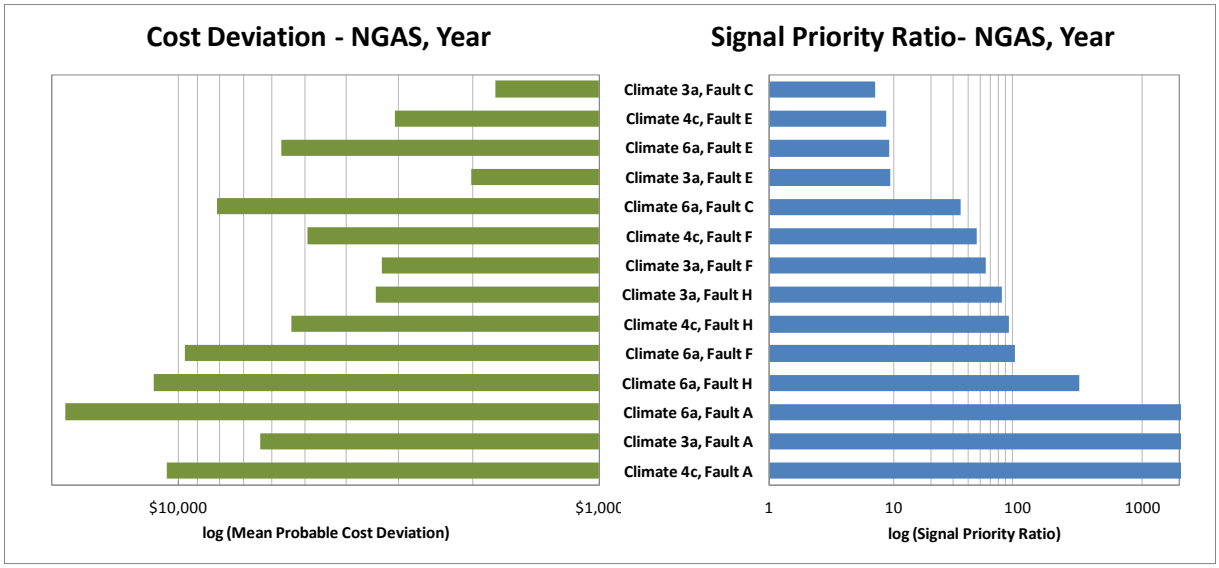


Figure 77: Top 15 faults in yearly natural gas energy end-use, in ascending signal priority ratio

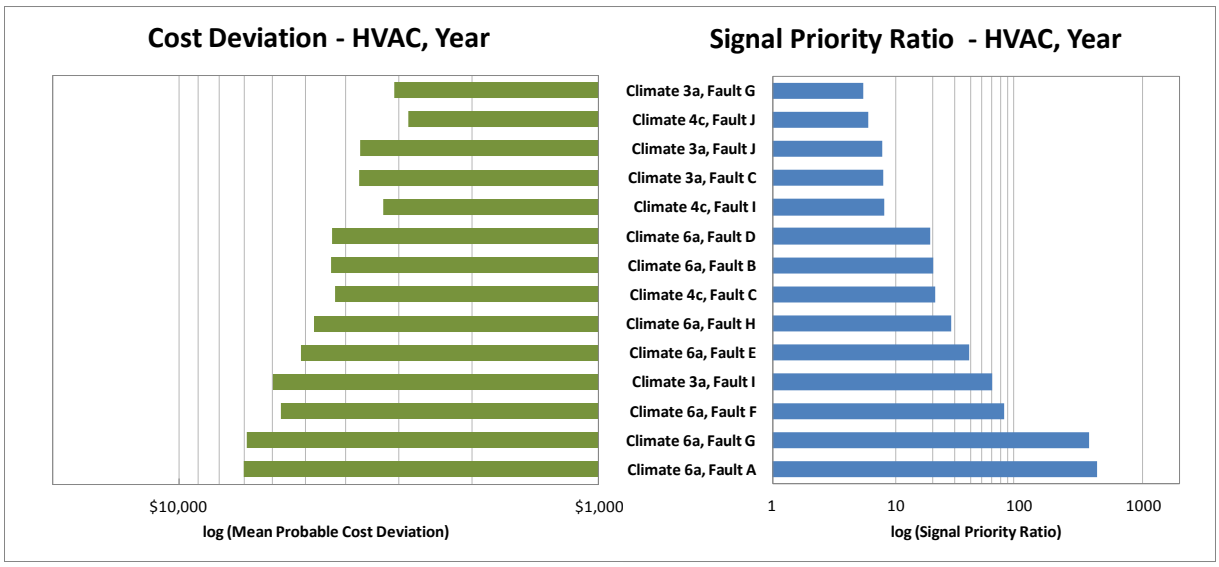


Figure 78: Top 15 faults in yearly HVAC energy end-use, in ascending signal priority ratio

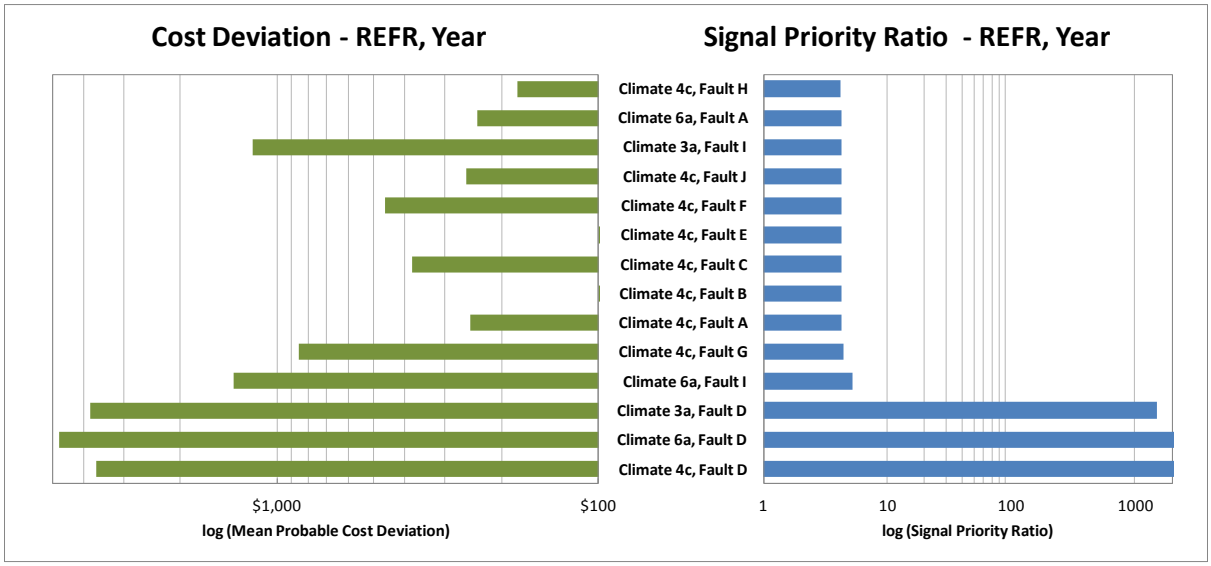


Figure 79: Top 15 faults in yearly refrigeration energy end-use, in ascending signal priority ratio

See the appendix for tables of all 30 faults as observed in months 1, 4, 7, and 10 for each end use. It should be noted that in these tables, all deviations in energy cost have been annualized (multiplied by 12 for periods of one month) and are expressed in absolute value.

The following figures compare the mean expected cost deviation to the signal priority ratio for all faults across all climate zones. In order to show what correlation

relationships exist between the two, both quantities have been transformed into respective ranks for each time period. Ranks are made in descending order, and are then re-defined as priorities in ascending order; *i.e.*, the lowest ranks are the top priorities. Thus, the points in the lower left corner of each graph represent the most severe faults. The ranks of the two quantities are not necessarily expected to match. This is due to the differences inherent to each of the fault prioritization metrics; the cost deviation value prioritizes based only on energy cost, while the signal ratio prioritizes based on energy cost, action cost, and the other organizational objectives built into the cost matrix. Across the portfolio of buildings, there are also different cost deviation magnitudes caused by any given fault in a given time period. For example, the cost deviation of fault "D" in climate zone 3a will be greater than that of fault "D" in climate zone 6a in the month of April.

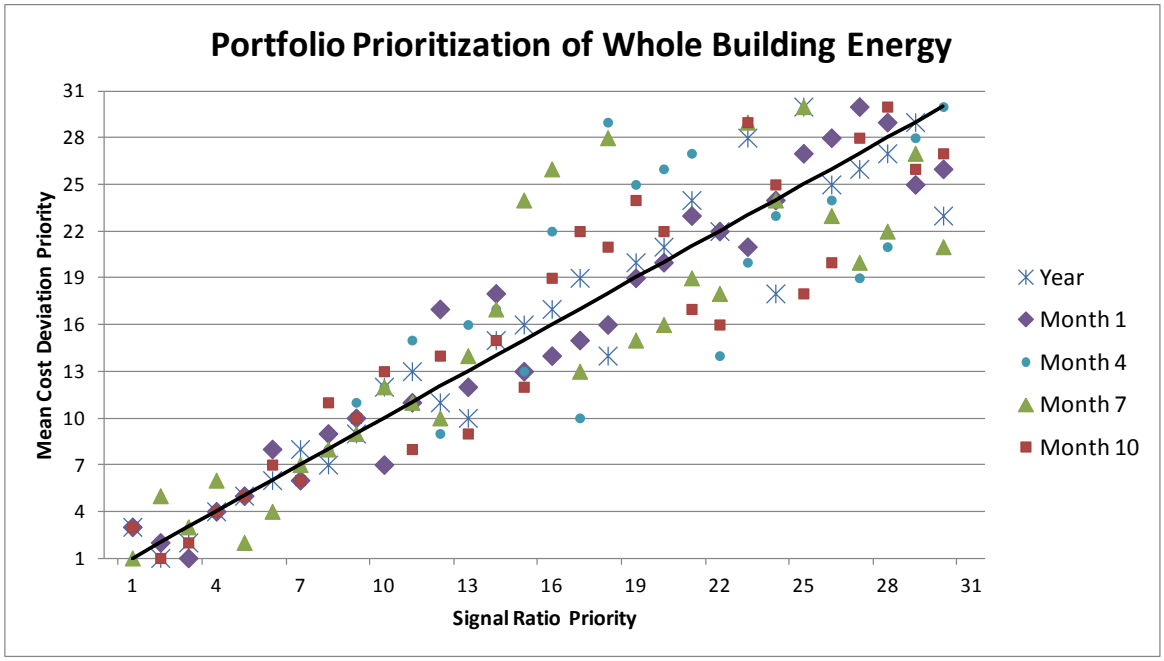


Figure 80: Comparison of whole building energy signal priority ratio to mean cost deviation for synthetic "portfolio" results

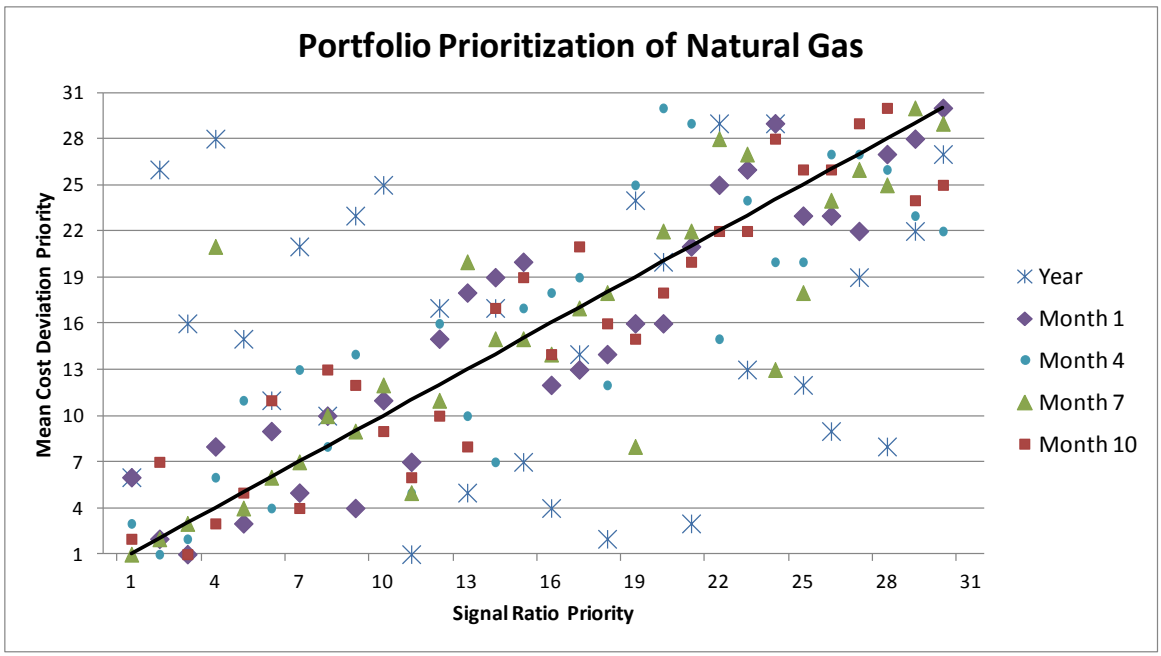


Figure 81: Comparison of natural gas energy signal priority ratio to mean cost deviation for synthetic "portfolio" results

Of all the end use signals, the rank of signal priority ratio is most closely correlated to the cost deviation rank for whole building energy use, as observed in Figure 80 above. The correlation of cost to signal priority for natural gas energy is more scattered – especially in "Month 7" and "Year". This can be attributed to the cost matrix definitions for each climate zone shown in Table 26 above. By saying that the signal should be sensitive to the fact that the ratio of cost of maintaining vs. the operating natural gas equipment is relatively high in climate zone 3a, any given cost deviation in natural gas use will have a lower signal priority ratio. Conversely, since the cost to maintain vs. the cost to operate is low for climate zone 6a, this signal will be relatively strong for any given cost deviation. Thus, the scattering of points in Figure 81 is a result of three distinct cost matrices for natural gas use.

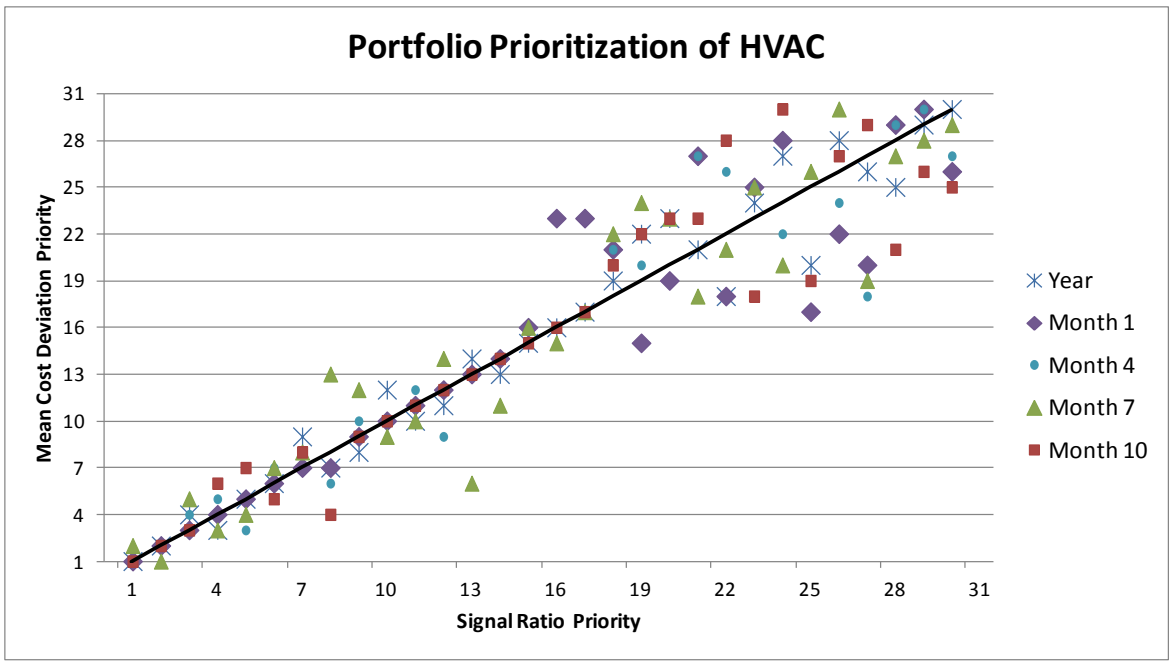


Figure 82: Comparison of HVAC energy signal priority ratio to mean cost deviation for synthetic "portfolio" results

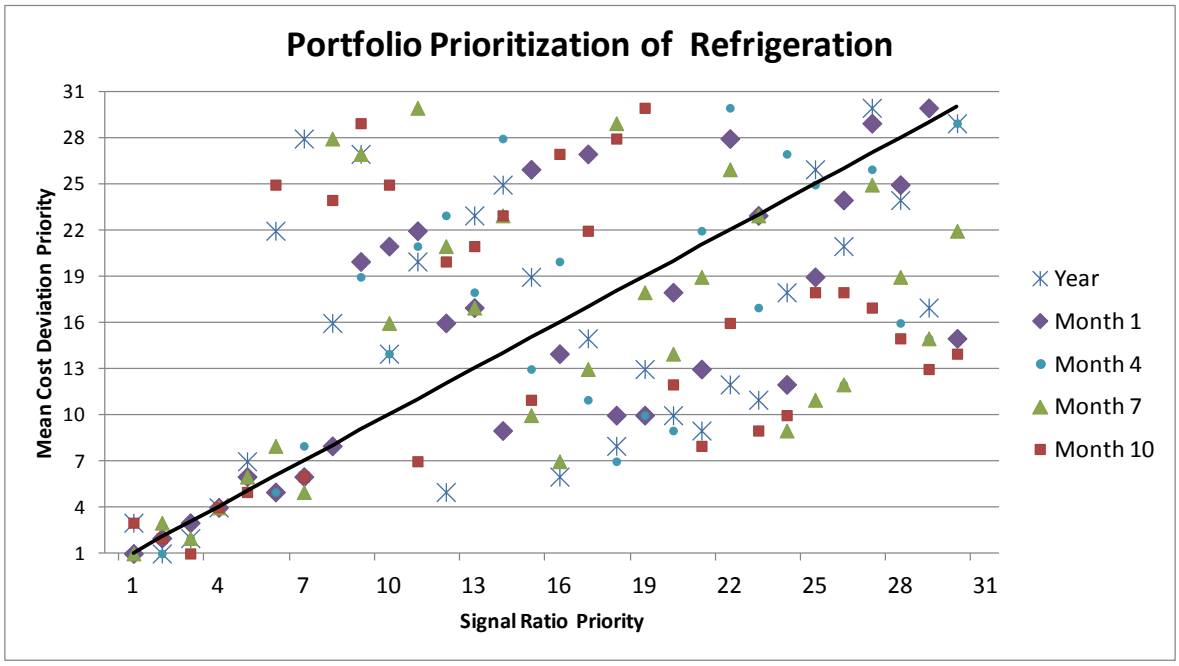


Figure 83: Comparison of refrigeration energy signal priority ratio to mean cost deviation for synthetic "portfolio" results

Figure 82 above shows that portfolio signal priority rankings are closely correlated to cost for the top ~50% of HVAC faults signals. Then, as priority decreases, the correlation

becomes less weakens due to an increasing ratio of potential maintenance cost to potential energy savings. For the refrigeration end-use this correlation falls apart even sooner, due to the generally low cost deviation in refrigeration energy end-use for any fault (refer to Figure 79 above).

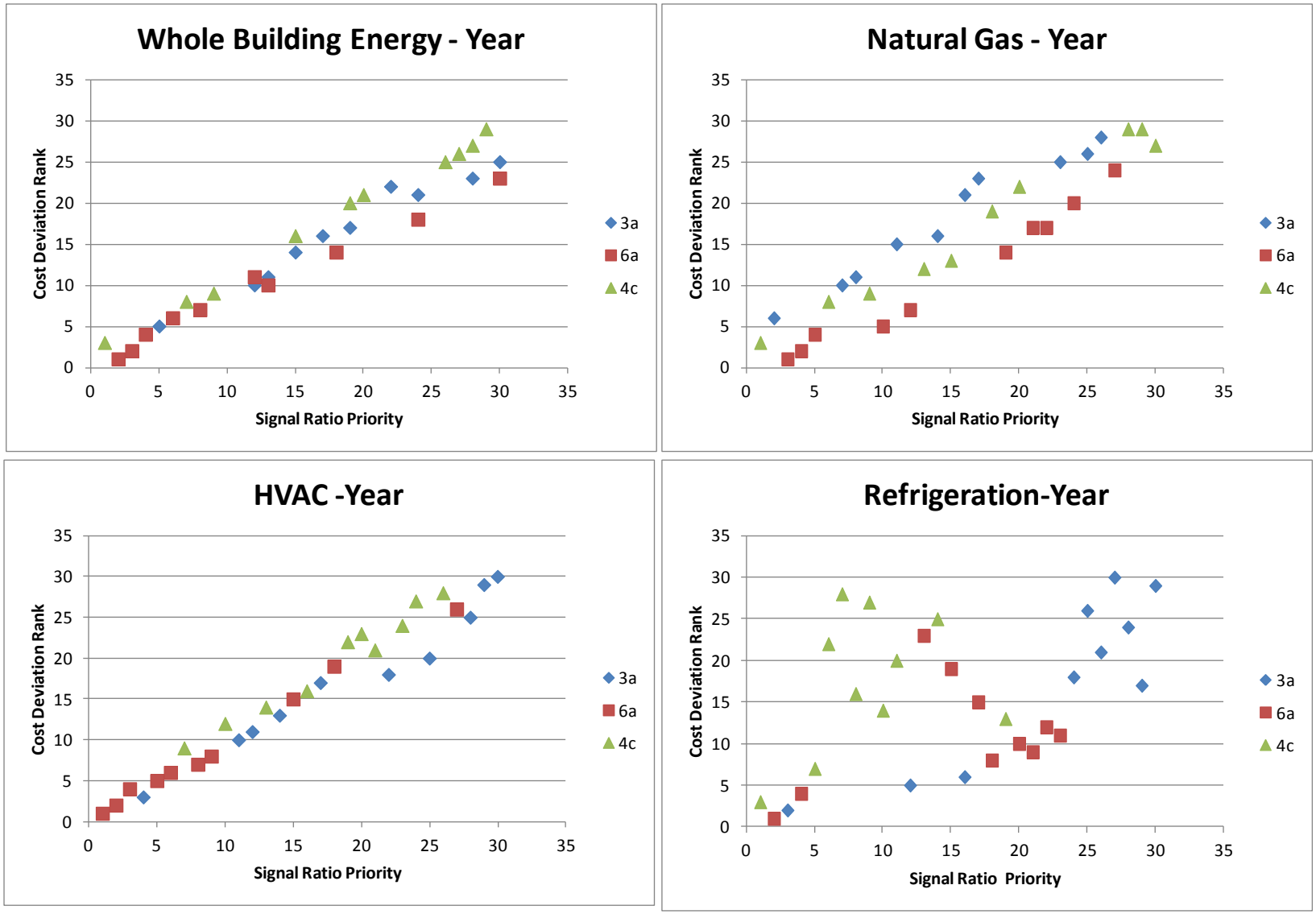


Figure 84 a, b, c, d: Signal priority ratio *rank* and cost deviation *rank* comparison decomposed by climate zone

Figure 84 above displays a decomposition of the cost/signal strength ratio by climate zone. This reveals more about how the risk tolerance threshold adjustment for controllability, as well as the cost matrix definition questions shape the signals. In the lower right corner, three distinct patterns of cost/signal ratio for refrigeration energy end use signals can be observed. Recall that in climate zone 4c, it was determined that energy use due to refrigeration was less controllable than other climate zones (refer to Table 25 above). This has caused the signals to be weaker for any given cost deviation compared to the other climate zones.

The clear divisions by climate zone in the rank comparisons in the upper-right graph of Figure 84 above illustrate the different weighting of natural gas energy cost in each climate zone. Indeed, this energy use varies most among climate zones. For a given cost deviation rank level, faults in zone 3a are most likely to get a higher signal priority ranking. This preference was built into the cost matrix by answering question #1 “high” for climate 3a, “N/A” for climate 4c, and “low” for climate 6a (refer to Table 26 above).

9 Conclusions and Extensions to Further Work

9.1 Summary and Conclusions

Through some initial work with a big-box retail industry partner, as well as additional literature review, the original concept (Henze *et al.*, 2015) of how to apply the Energy Signal Tool has been refined. The additional statistical processing theory introduced in this work has rendered the Energy Signal Tool capable of being deployed rapidly in a variety of locations. Expanding the possibilities for the cost matrix has promise to incorporate additional organizational goals into the Energy Signal Tool decision support. Rather than being best suited for single building fault detection, a case has been made for broadening the application of the ESTool to portfolio energy management and retrofit prioritization. The greatest power in this self-benchmarking tool lies in the ability to quickly assess whole-building and end-use performance to identify areas with potential for both energy savings *and* achieving broader organizational goals. The Energy Signal Tool may not be especially useful for identifying the root causes of faults with multiple building systems interactions. It was proved to be useful for quantifying fault severity with a sophisticated context of uncertainty analysis. The Energy Signal Tool is useful for directing facilities management attention to those *buildings in a portfolio* that merit further attention to diagnose specific faults.

Uncertainty analysis with Latin Hypercube Monte Carlo sampling was used to perturb significant parameters and generate expected ranges of consumption for each monitored end-use. Given that sufficient knowledge goes into creating the benchmark model and uncertainty is minimal, these expected ranges can represent the probable range in fault-free operation. It was shown that the concepts of the Energy Signal Tool can be

used with the OpenStudio detailed modeling platform to prioritize energy related faults in building portfolio. It was possible to illustrate this with synthetic data, but not with data from the case study building. Working with real building data and attempting to calibrate a model to it revealed just how difficult and complex these tasks can be. It can be concluded from the model calibration exercise that methods of calibrating to whole-building energy use data are not sufficient to develop a tool that monitors disaggregated end-uses within the building.

It may also be possible to improve the method of modeling the benchmark with an alternative approach to model calibration. One idea for such an approach was discovered late in the writing of this work. Subbarao *et al.* (2014) have proposed what they term an "*actuarial*" framework of assessing building retrofit energy savings. These authors show how data from "neighborhoods" of similar buildings within a portfolio can be used to generate PDFs of pre-retrofit energy consumption. When some pre and post-retrofit data is available from several buildings in the neighborhood, savings for a single building can be estimated in the form of a PDF through a convolution of pre and post retrofit savings differences. As discussed earlier, it is often necessary to expand the ranges of categorical variables to generate a sufficiently large set of buildings for peer comparison. The neighborhood methodology uses simulation to correct for differences in a loosely related peer group of buildings. The authors cite one example of correcting a peer group of similar buildings situated in slightly different locations within a weather neighborhood. Since the correlation of building EUI to weather conditions will be similar for peer buildings within this neighborhood, the expected EUI of a building in one location can be normalized to what would be expected in another nearby, but not identical, location. The authors note that it is much more difficult to correct for differences in building characteristics within an expanded peer group. This would involve identifying the important characteristics that determine energy use. Unless the dimensionality of expanded peer group differences is

small enough to be accounted for with available measurements, this would probably involve generating data with large simulation batches. Nevertheless, the approach of creating an expanded and normalized peer group for performance assessment may have promise.

Monte Carlo sampling of a model with 18 parameters for calibration requires at least 8000 (the number used in this work) samples, if not more. Generating and modeling this number of samples has a duration of a least 14 hours, at a cost of at least $40. This would currently be a challenge to implement in a real building application of the tool, when the benchmark would need to be recalibrated to account for any major change, and global SA would need to be updated for ESTool input. Other forms of global sensitivity analysis, such as the Morris method (as mentioned in section 5.2.1, would be more computationally efficient when possible to implement. Reducing computational time for model calibration efforts would require new methods such as meta-modeling (see Eisenhower *et al.*, 2012), and optimization to arrive at calibration solutions sooner.

The Energy Signal Tool is intended to be applied with facilities manager involvement. It was unreasonable to expect the uncertainty in a limited number of physical building parameters to generate a wide enough range of possibilities to bridge the gap between a model developed with little client stakeholder involvement and one that can be used for precise decision support. Thorough heuristic model calibration is crucial for the benchmark model be validated for use in the ESTool. Facilities manager involvement is paramount to this. The additional knowledge needed would have required resources that were not available to the researcher or the industry partner. While this involvement was not available from the retail partner for this work, it is still conceivable that an organization would see the value in sufficient stakeholder involvement to make implementing the ESTool a reality.

Thourough testing of the refined ESTool concepts was done with synthetic fault scenarios in place of applying it directly to the case study building. Synthetic fault testing helped to show how sensitive the tool is to detecting faults. It proved the Energy Signal Tool had advantages over other methods of performance benchmarking; which use whole-building energy consumption data alone. It also demonstrated the shortcomings of the tool in pointing to specific end-uses for root causes of faults. For many of the typical faults tested, (e.g., thermostat set point changes, scheduling errors, fan malfunction) signals were generated in many time periods across multiple end-uses. While this is valid, it does not help with deciding which building system(s) require corrective action to address the problem. With a more extensive and well-grouped sub-metering infrastructure, it would be possible to incorporate more monitored end-uses in the Energy Signal Tool, and then develop a set of rules based on consumption patterns to better identify root causes. It is concluded that it would be simpler to rely on traditional FDD techniques to identify specific root causes of faults.

9.2 Extensions to Future Work

The biggest challenge to overcome before the Energy Signal Tool cam be implemented remains calibrating the benchmark model. As mentioned, it is not so much the single representative model as is the importance of having minimal bias in the end-use distributions. A far greater parameter set may be needed to achieve a model that produces all “green” signals when compared to the observed data that went into calibrating it. The work that can be done to achieve this includes developing better measures to investigate all possible uncertain parameters, including full schedule manipulation, and developing algorithms to calibrate the model to each end-use desired to monitor. This will involve working extensively with OpenStudio server calibration measures and possibly optimization routines. Standards for a statistical process apart from following ASHRAE - 14 guidelines must be established to achieve proper model calibration for the ESTool.

For this tool to be broadly applied to an entire building portfolio, such as the commercial big-box retailer partner for this work, an automated process of generating and keeping up to date a large set of building models is needed. NREL has been developing a project called COFEE (Customer Optimization for Energy Efficiency; see Eber, 2014) to enable a major electric utility company to generate baseline models of thousands of customer buildings. The software behind COFEE leverages Google images along with the existing database of customer information to rapidly generate a baseline Open Studio model of each building. The goal of COFEE is to enable the utility to increase the efficiency of its demand side management program by targeting the customers with the highest model-predicted potential to benefit from DSM incentives. It is very conceivable that a program such as COFEE could be used by corporations who manage a large portfolio of buildings to generate models and keep then up to date. This could be done through an interface that pulls in information that the organization already keeps track of (building inventory, area, occupancy, installed systems). The challenge here is organizing this stream of information and creating a dynamic repository of models and model results to draw information from.

Portfolio managers have had access to other, more functional based, fault detection diagnostic (FDD) infrastructure for many years. Traditional FDD is based on the infrastructure of the building management system. It has come to include features that incorporate system measurements such as loop temperatures and rates of energy consumption. This can amount to giving the facilities manager an essentially binary signal of whether or not the building systems are operating to meet functional requirements such as safety and comfort. If a fault is detected in function, an alarm is sent to the facilities manager that alerts exactly what component needs to be addressed. If FDD were to be integrated with something such as the Energy Signal Tool the result would be a single tool more valuable than the two would be apart from one another. This way, signals could be

prioritized based on energy cost as well as their importance to the vital functions of the building by examining FDD information for correlation to energy end use consumption. Future work in binding the Energy Signal Tool to FDD capabilities would be quite valuable.

Other future work should attempt to integrate a larger set of, and more varied types of uncertain parameters into the pre-screening process. Unknowns such as controls scheduling, hours of use and hours of occupancy are additional parameters that can have an effect on the model but could not be incorporated in this work due to technical limitations of premature software used. With more advanced Ruby measures, it might be possible to perform model calibrations, including schedule parameters, on the basis of matching the model to *hourly* (rather than monthly) data. It is also conceivable that a list of potentially significant parameters based on building type, primary systems, envelope characteristics, and monitored end-uses could be automatically generated for each building type – given a sufficiently large library of measures to draw from. This would allow a user without much knowledge of which parameters are potentially significant to make a reasonable characterization of uncertainty in the model.

In practice, the benchmark model for an Energy Signal Tool would need to be kept dynamically in tune with any operational or equipment changes to the building. When the ESTool identifies a fault in the building, feedback from the facilities manager who has made an inspection of the fault is needed to inform the energy manager whether the fault was fixed, could be ignored, or if the fault was not observed at all. Depending where the responsibility to maintain a baseline model may fall (on the customer or on the energy performance assessment service provider), it may be advantageous to integrate the benchmark modeling process with the BMS. As described in Chapter 2, there is the potential for a synergistic existence of both traditional FDD and the ESTool; the former can pinpoint faults' locations, and the latter can quantify their severity. Currently, any link between the ESTool and BMS-based FDD capabilities would need to be made with direct

human involvement. Future work could be done to develop hardware and software links between the BMS, FDD, and the ESTool. Perhaps this dynamic interface could also allow for notification of parameter updates from the BMS.

Future work is needed to link the energy efficiency opportunities identified by the tool to broader organizational risks and opportunities. Energy is far from the largest expense that an organization has; that is usually compensation for their employees. Thus, the greatest opportunity for identifying areas where organizational performance could be improved is in employee productiveness and well-being. Some organizations, such as a big-box retailers, use their buildings for the majority of customer interactions. In these cases, an important cost consideration might be customer satisfaction and variables that affect sales volume. Eventually, this additional information could be part of the cost matrix input. Currently, the cost matrix is made up on scalar values that are highly subjective in nature. In reality, developing the cost matrix for each end use should take into consideration measured data for things like actual equipment maintenance costs, utility rates, value of inventory, and sales volume or worker productivity. Uncertainty would also be a component of the cost function, and could be derived from empirical data. Additional research is needed to uncover the quantitative correlations that exist between energy use and other organizational priorities.

It proved very difficult to produce an EnergyPlus model that described the energy use of an existing building to the level of accuracy needed for an Energy Signal Tool. Rather than using detailed simulation modeling, it might be possible instead to create baseline models using a data-based emulator. Using Gaussian processes, it is possible to build an emulator much akin to numerical regression techniques. The emulator would create a multivariate model which takes information about environmental conditions and predicts end-use energy consumption based on past history of consumption. This would reduce computational time required for generating expected end use distributions on the order of

99.99%. This would make it more realistic, given current computing power limitations, to use the tool for performance assessment. The process of calibrating such a model to operational data of disaggregated end-uses would be purely mathematical, and thus much easier. The disadvantage to using the numerical emulator is that it might not be valid for predicting monthly end use during seasonal shifts. Another drawback of using mathematical models would be the inability to test retrofit scenarios before are implemented, and to model controls specifically as intended. However, if several are implemented in a portfolio and performance data is available, the model would certainly be capable of learning from this data and projecting the impact onto the entire portfolio. Perhaps data-based emulators will be the future of detailed energy modeling.

10 References

Akaike, Hirotugu (December 1974). "A new look at the statistical model identification". IEEE Transactions on Automatic Control 19 (6): 716–723.

ASHRAE 14-2002. Guideline 14-2002: Measurement of Energy and Demand Savings. American Society of Heating, Refrigeration and Air-Conditioning Engineers, Atlanta

ASHRAE, Handbook of Fundamentals, 2013.

Ball, B., N. Long. “OpenStudio Analysis Spreadsheet”. Last update accessed October 27, 2014.

Bonnema, E., et al. "50% Advanced Energy Design Guides." (2012).

Bonnema, Eric, Matt Leach, and Shanti Pless. Technical Support Document: Development of the Advanced Energy Design Guide for Medium to Big Box Retail Buildings–50% Energy Savings. No. NREL/TP-5500-52589. National Renewable Energy Laboratory (NREL), Golden, CO., 2013.

Booth, A. T., and R. Choudhary. "Decision making under uncertainty in the retrofit analysis of the UK housing stock: Implications for the Green Deal."Energy and Buildings 64 (2013): 292-308.

Bucking, S., Zmeureanu, R., Athientitis, A. A methodology for identifying the influence of design variations on building energy performance. Journal of Building Performance and Simulation, 7:6, 2014 (pp. 411 – 426)

Burhenne, Sebastian. Monte Carlo Based Uncertainty and Sensitivity Analysis for Building Performance Simulation. PhD Thesis. Fraunhaufer Institute for Solar Energy Systems, Germany. November, 2013.

Chung, William. "Review of building energy-use performance benchmarking methodologies." Applied Energy 88.5 (2011): 1470-1479.

Coakley, D., Raftery, P., Keane, M. “A review of methods to match building energy simulation models to measured data.” Renewable and Sustainable Energy Reviews 37 (2014), pp. 123-141.

Commercial Buildings Energy Consumption Survey.
http://www.eia.gov/consumption/commercial/data/ [Accessed 10 February, 2015]

Corrado, Vincenzo, and Houcem Eddine Mechri. "Uncertainty and sensitivity analysis for building energy rating." Journal of building physics (2009).

Costa, Andrea, et al. "Building operation and energy performance: Monitoring, analysis and optimisation toolkit." Applied Energy 101 (2013): 310-316.

de Wit, S. Uncertainty in Building Simulation. <In: Malkawi, A. M., Augenbroe, G., editors, Advanced

Building Simulation, pp. 25-59
DOE. 2012. EnergyPlus documentation. US Department of Energy: http://apps1.eere.energy.gov/buildings/EnergyPlus/pdfs/inputoutputreference.pdf [Accessed 10 April 2014]

Domíngues-Muñoz, F. et al. Uncertainty in the thermal conductivity of insulation materials. Proceedings of the eleventh IBPSA Conference. Glasgow, Scotland; July 27-30, 2009.

Doukas, Haris, Christos Nychtis, and John Psarras. "Assessing energy-saving measures in buildings through an intelligent decision support model." Building and environment 44.2 (2009): 290-298.

Eber, Kevin. “Analyzing the Energy Use of Massive Numbers of Existing Buildings” NREL Continuum, March 2014. http://www.nrel.gov/continuum/homes_buildings/computing.cfm [Accessed: 16 September, 2014

Eisenhower, B., O’Neill, Z., Fonoberov, V., Mezic, I. Uncertainty and sensitivity decomposition for building energy models, Journal of Building Performance Simulation. Vol. 5, No. 3, May 2012b (pp. 171-184)

Eisenhower, B., O’Neill, Z., Narayanan, S., Fonoberov, V., Mezic, I. A methodology for meta-model based optimization in building energy models, Energy and Buildings 47, 2012a (pp. 292 – 301)

Energy Information Administration: http://www.eia.gov/ [Accessed 4 February 2015]

Energy Star Portfolio Manager. www.energystar.gov/buildings [Accessed 10 February, 2015]

Fleming, Katherine, Nicholas Long, and Alex Swindler. "The Building Component Library: an Online Repository to Facilitate Building Energy Model Creation." Proceedings of the 2012 ACEEE Summer Study on Energy Efficient Buildings (2012).

Frankel, M., C Turner. Energy Performance of LEED® for New Construction Buildings. New Buildings Institute. White Salmon Washington, 2008. Final Report to the US Green Building Council

Gao, Xuefeng, and Ali Malkawi. "A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm." Energy and Buildings 84 (2014): 607-616.

Guo, Mingjin, et al. "Comparative Analysis on Energy Consumption of Commercial Buildings Based on Sub-metered Data." (2014).

Halkidi, Maria, Yannis Batistakis, and Michalis Vazirgiannis. "On clustering validation techniques." Journal of Intelligent Information Systems 17.2-3 (2001): 107-145.

Hedrick, Roger, Vernon Smith, and Kristin Field. "Restaurant Energy Use Benchmarking Guideline." Contract 303 (2011): 275-3000.

Henze, Gregor P., et al. "An energy signal tool for decision support in building energy

systems." Applied Energy 138 (2015): 51-70.

Henze, Gregor P., et al. Control Limits for Building Energy End Use Based on Engineering Judgment, Frequency Analysis, and Quantile Regression. National Renewable Energy Laboratory, 2014.

Hinge, Adam, et al. "Back to school on energy benchmarking." Proceedings of the 2002 ACEEE Summer Study. Monterey, Calif (2002).

Hirsch, Adam, et al. Results and Lessons Learned From the DOE Commercial Building Partnerships. National Renewable Energy Laboratory (NREL), Golden, CO., 2014.

Hopfe, Christina J., and Jan LM Hensen. "Uncertainty analysis in building performance simulation for design support." Energy and Buildings 43.10 (2011), pp. 2798-2805.

Loonen, R., Hensen, J. Dynamic sensitivity analysis for performance-based building design and operation. Proceedings of IBPSA 2013, Chambéry, France.

MacDonald, I.A. “Quantifying the effects of uncertainty in building simulation.” PhD thesis. University of Strathclyde, Scotland. 2002

Macdonald, Iain A. "Comparison of sampling techniques on the performance of Monte-Carlo based sensitivity analysis." Eleventh International IBPSA Conference. 2009.

Macumber, Daniel L., Brian L. Ball, and Nicholas L. Long. "A GRAPHICAL TOOL FOR CLOUD-BASED BUILDING ENERGY SIMULATION." Submitted EMC Simbuild (2014).

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K.(2014). cluster: Cluster Analysis Basics and Extensions. R package version 1.15.3.

Mills, E. Building Commissioning: A Golden Opportunity for Reducing Energy Costs and Greenhouse Gas Emissions. Lawrence Berkeley National Laboratory. Berkeley, CA. 2009.

Montgomery, Douglas C. Applied Statistics and Probability for Engineers 6th edition. Wiley, 2013.
NREL. Target Improves Efficiency in New Construction. US Department of Energy Buildings Technology Office, Energy Efficiency and Renewable Energy. October, 2013(a).

NSTC. “Submetering of Building Energy and Water Usage: Analysis and Recommendations of the Subcomittee on buildings technology research and development”. National Science and Technology Council Committee on Technology (NSTC). Report; October, 2011.

O’Donnel, J. et al., “Scenario Modelling: A holistic environmental and energy management method for building operation optimization.” Energy and Buildings (2012): 146-157

Parker J., Cropper P., and Shao L., A calibrated whole building simulation approach to assessing retrofit options for Birmingham airport, First Building Simulation and Optimization Conference Loughborough, UK 10-11 September 2012, 49-56

Pati, Debajyoti, Cheol-Soo Park, and Godfried Augenbroe. "Roles of building performance assessment in stakeholder dialogue in AEC." Automation in construction 15.4 (2006): 415-427.

Pavlak, Gregory S., et al. "Comparison of Traditional and Bayesian Calibration Techniques for Gray-Box Modeling." Journal of Architectural Engineering 20.2 (2013).

Reddy, T. Agami, Itzhak Maor, and Chanin Panjapornpon. "Calibrating detailed building energy simulation programs with measured data—Part I: General methodology (RP-1051)." Hvac&R Research 13.2 (2007a): 221-241.

Reddy, T. Agami, Itzhak Maor, and Chanin Panjapornpon. "Calibrating detailed building energy simulation programs with measured data—Part II: application to three case study office buildings (RP-1051)." Hvac&r Research 13.2 (2007b): 243-265.

Reddy, T. Agami. Applied data analysis and modeling for energy engineers and scientists. Springer, 2011.

Robertson, Joseph, Ben Polly, and Jon Collis. "Evaluation of Automated Model Calibration Techniques for Residential Building Energy Simulation." (2013).

Saltelli A, Chan K, Scott E M; Sensitivity Analysis, J Wiley and Sons, 2000.

Sharp, T.R. Benchmarking Energy Use in Schools. Paper presented at proceedings of the 1998 ACEE Summer Study on Energy Efficiency in Buildings (1998).

Stein, J. Retail Building Guide for Entrance Energy Efficiency Measures. NREL, 2012

Struck, Christian, and Jan Hensen. "On supporting design decisions in conceptual design addressing specification uncertainties using performance simulation." Proceedings of the 10th IBPSA Building Simulation Conference. 2007.

Subbarao, Krishnappa, Pavel V. Etingov, and T. Agami Reddy. An actuarial approach to retrofit savings in buildings. No. PNNL-SA-91271. Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2014.

Taasevigen, Danny J., and William Koran. User's Guide to the Energy Charting and Metrics Tool (ECAM). No. PNNL-21160. Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2012.

The Energy Index for Commercial Buildings. US Department of Energy. http://buildingsdatabook.eren.doe.gov/CBECS.aspx [Accessed 10 February, 2015

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0 R package: http://cran.r-project.org/web/packages/MASS/MASS.pdf

Wang, Liping, Paul Mathew, and Xiufeng Pang. "Uncertainties in energy consumption introduced by

building operations and weather for a medium-size office building." Energy and Buildings 53 (2012): 152-158.

11 Appendices

11.1 Appendix A: Sub-metered points and additional data

List of sub-metered points at case-study store

POINT	CHANNEL NAME	UNITS	Phase	Metered Load
1	TOTAL LOAD	KW	a	MDS
2	TOTAL LOAD	KW	b	MDS
3	TOTAL LOAD	KW	c	MDS
4	HP-1	KW	a	HP-1
5	MDP	KW	a	MDP-1
6	LD-1	KW	a	MDP-2
7	HL-1	KW	a	HL-1
8	HL-1	KW	b	HL-1
9	HL-1	KW	c	HL-1
10	HL-2	KW	a	HL-2
11	HL-2	KW	b	HL-2
12	HL-2	KW	c	HL-2
13	Generator	KW	a,a	HLE-1,LPE-1
14	Generator	KW	b,b	HLE-1,LPE-1
15	Generator	KW	c,c	HLE-1,LPE-1
16	LPE-2	KW	a	LPE-2
17	LP-1	KW	a	LP-1
18	LP-4	KW	a	LP-4
19	RTCR-1	KW	a	RTCR-1
20	RTU-10	KW	a	RTU-10
21	RTU-7	KW	a	RTU-7
22	RTU-8	KW	a	RTU-8
23	RTU-12	KW	a	RTU-12
24	RTU-12 Supply Fan	KW	a	RTU-12 Supply Fan
25	DHU-13	KW	a	DHU-13
26	DHU-13	KW	b	DHU-13
27	DHU-13	KW	c	DHU-13
28	DHU-13 - Reactivation Fan	KW	a	DHU-13 - Reactivation Fan
29	DHU-13 - Process Air Blower	KW	a	DHU-13 - Process Air Blower
30	DHU-13 - Compressor 1	KW	a	DHU-13 - Compressor 1
31	DHU-13 - Compressor 2	KW	a	DHU-13 - Compressor 2
32	DHU-13 - Compressor 3	KW	a	DHU-13 - Compressor 3
33	DHU-13 - Compressor 4	KW	a	DHU-13 - Compressor 4
34	DHU-13 - Condensor Fan 1	KW	a	DHU-13 - Condensor Fan 1
35	DHU-13 - Condensor Fan 2	KW	a	DHU-13 - Condensor Fan 2
36	DHU-13 - Condensor Fan 3	KW	a	DHU-13 - Condensor Fan 3

37	LD-2	KW	a	LD-2
38	HL-3	KW	a	HL-3
39	HL-3	KW	b	HL-3
40	HL-3	KW	c	HL-3
41	UPS-1	KW	a	UPS-1
42	Parking Lights	KW	a	HL-4
43	Parking Lights	KW	b	HL-4
44	Parking Lights	KW	c	HL-4
45	RTU 1	KW	a	RTU 1
46	RTU 2	KW	a	RTU 2
47	RTU 3	KW	a	RTU 3
48	RTU 4	KW	a	RTU 4
49	RTU 5	KW	a	RTU 5
50	RTU 6	KW	a	RTU 6
51	RTU 9	KW	a	RTU 9
52	RTU 11	KW	a	RTU 11
53	RTU-ERV-1	KW	a	RTU-ERV-1
54	LP-3	KW	a	LP-3
55	LP-5	KW	a	LP-5
56	LP-6	KW	a	LP-6
57	LP-7	KW	a	LP-7
58	LP-8	KW	a	LP-8
59	LP-9	KW	a	LP-9
60	LP-10A	KW	a	LP-10A
61	LP-10B	KW	a	LP-10B
62	LP-11	KW	a	LP-11
63	LP-12A	KW	a	LP-12A
64	LP-12B	KW	a	LP-12B
65	LPE-3	KW	a	LPE-3
66	Door Htrs	KW	a	LP-12A/cct 2,8
67	Door Htrs	KW	b	LP-12A/cct 4
68	Door Htrs	KW	c	LP-12A/cct 6
69	CU-1	KW	a	CU-1
70	CU-2	KW	a	CU-2
71	CU-3	KW	a	CU-3
72	CU-4	KW	a	CU-4
73	VAV-5-1	KW		VAV-5-1
74	VAV-5-2	KW		VAV-5-2
75	VAV-5-3	KW		VAV-5-3
76	VAV-5-5	KW	a	VAV-5-5
77	VAV-5-6	KW	a	VAV-5-6
78	VAV-5-7	KW		VAV-5-7
79	VAV-5-9	KW		VAV-5-9

80	VAV-11-1	KW		VAV-11-1
81	WH-1	KW	a	WH-1
82	WH-2	KW		WH-2
83	WH-3	KW		WH-3
84	WH-4	KW		WH-4
85	RTU-5 Supply Fan	KW	a	RTU-5 Supply Fan
86	RTU-9 Supply Fan	KW	a	RTU-9 Supply Fan
87	RTU-9 - Compressor 1	KW	a	RTU-9 - Compressor 1
88	RTU-9 - Compressor 2	KW	a	RTU-9 - Compressor 2
89	RTU-9 - Compressor 3	KW	a	RTU-9 - Compressor 3
90	RTU-9 - Compressor 4	KW	a	RTU-9 - Compressor 4
91	RTU-11 Supply Fan	KW	a	RTU-11 Supply Fan
92	RTU-11 - Compressor 1	KW	a	RTU-11 - Compressor 1
93	RTU-11 - Compressor 2	KW	a	RTU-11 - Compressor 2
94	RTU-11 - Compressor 3	KW	a	RTU-11 - Compressor 3
95	RTU-11 - Compressor 4	KW	a	RTU-11 - Compressor 4
96	RTU-11 Power Exhaust Fan	KW	a	RTU-11 Power Exhaust Fan
97	RTU-11 - Condensor Fan 1	KW	a	RTU-11 - Condensor Fan 1
98	RTU-11 - Condensor Fan 2	KW	a	RTU-11 - Condensor Fan 2
99	RTU-11 - Condensor Fan 3	KW	a	RTU-11 - Condensor Fan 3
100	RTU-11 - Condensor Fan 4	KW	a	RTU-11 - Condensor Fan 4
101	PRV-1	KW	a	PRV-1
102	LP-2	KW	a	LP-2
103	UH-1	KW	a	UH-1
104	Checklane Coolers	KW	a	Checklane Coolers
105	Checklane Coolers	KW	b	Checklane Coolers
106	Checklane Coolers	KW	c	Checklane Coolers
107	BG-1	KW	a	COMPACTOR
108	BG-2	KW	a	BALER
Virtual Points				
109	Generator	KW	a,a	HLE-1,LPE-1,2
110	Generator	KW	b,b	HLE-1,LPE-1,2
111	Generator	KW	c,c	HLE-1,LPE-1,2
112	Sales Floor Lights	KW	a,a	HL-1, HL-2
113	Sales Floor Lights	KW	b,b	HL-1, HL-2
114	Sales Floor Lights	KW	c,c	HL-1, HL-2

Additional energy sub-meter data

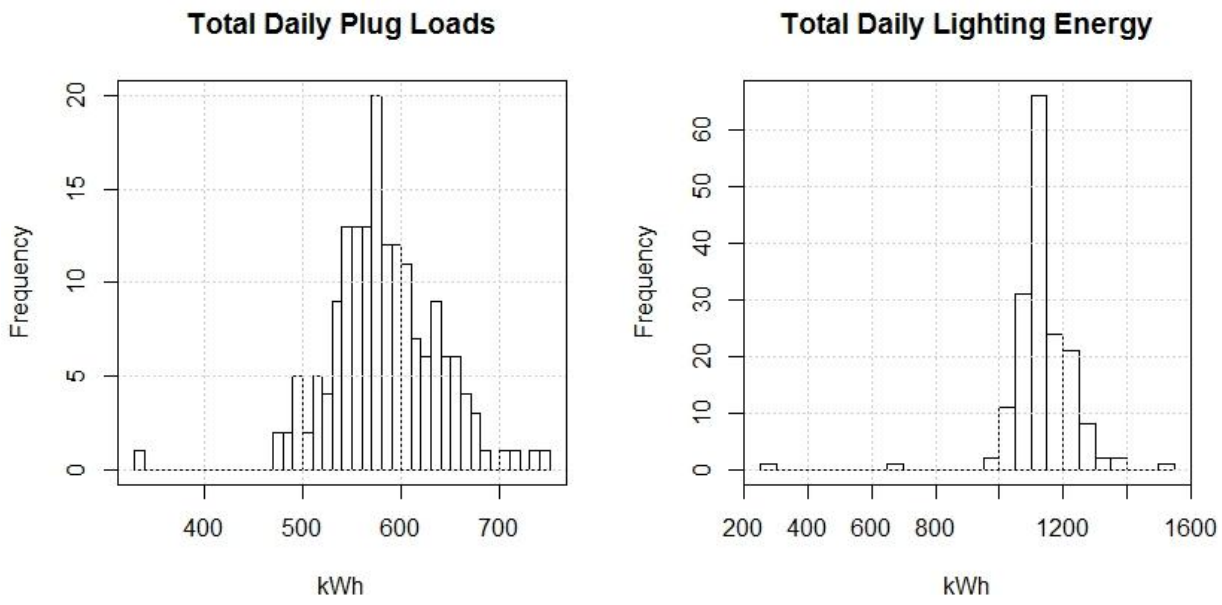


Figure 85: Lighting and Plug Load Daily Energy Distributions across the 170-day data collection period

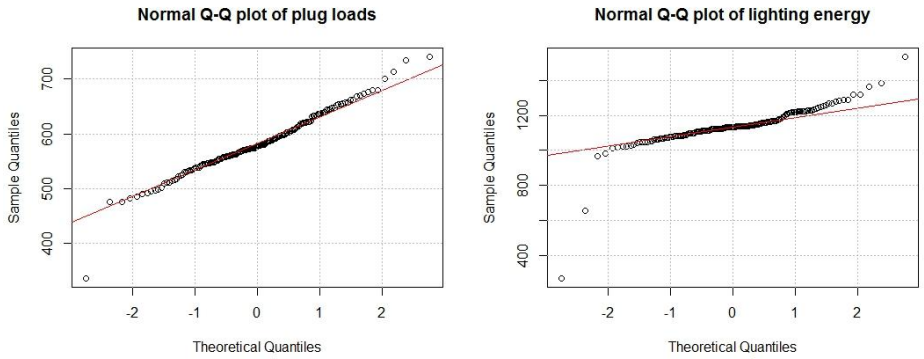


Figure 86: Q-Q plots for quantifying the normalcy of the two load distributions

Parameter *(weather independent)*	**Observed median value of Daily kWh**	**Stdev. Of daily kWh**	**95% Tolerance Interval about Median* for Daily kWh**
Plug Loads – non holiday	562.5	41.1	*484.7 < X < 645.6 kWh/day (+/- %)*
Plug Loads – holiday season	591.9	58.2	*477.74 < X < 705.97 kWh/day (+/-)*

Lighting Energy	1132.23	107.58	*921.4 < X < 1343.1 kWh/day (+/- 18.6%)*

**Assuming the data follows a normal distribution, centered about the median*

This data was not used for the model calibration since it was recorded accurately by the sub-metering equipment and is a benchmark to be calibrated against.

11.2 Appendix B: Parameter Characterization and Pre-Screening

Par. No.	Description:	Justification	Source
	Physical Properties		
10	Fan Pressure Rise (multiplier)	actual construction will not match rated value precisely	
12	Rated roof absorptance	actual value will decrease with level of dirt	conversations with Target
13	Wall Specific Heat	actual construction will not match rated value precisely	MacDonald, 2002
14	Wall/Roof Conductivity	actual construction will not match rated value precisely	Dominguez, 2008
30	Space Infiltration reduction (%)	varies with occupancy patterns, weather, etc.	
33	Space light load multiplier	varies slightly within known bounds	empirical data from submetering
34	Space thermal mass multiplier	varies with seasonal display and store configuration	
35	Space occupancy multiplier	varies greatly	conversations with Target
36	Space plug load multiplier	varies with seasonal displays within known bounds	empirical data from submetering
	HVAC Equipment		
8	Fan Efficiency (%)	can vary with equipment degredation	
9	Fan motor Efficiency (%)	can vary with equipment degredation	
11	HX Sensible Effectiveness (%)	can vary with equipment degredation	
15	minimum system airflow ratio	can vary with control settings and damper functionality	
17	Rated Cool Coil COP	can vary with equipment degredation	
18	Rated Gas Burner Eff (%)	can vary with equipment degredation	
31	Cooling Sizing Factor	design calculations not documented	
32	Heating Sizing Factor	design calculations not documented	
	Refrigeration		
37	Sum of UA for suction piping	actual installation details may vary from design	
20	Ratio Evap rate and Latent Cap	latent loads may vary with use	
23	Fraction of anti-sweat energy to case load	some cases often feel warm on the outside	conversations with Target
24	anti-sweat energy multiplier	this depends on space and dewpoint	
25	Delta in minimum ref. condensing temp (oC)	condenser temp sensor and modifications	
26	Condenser fans min flow (%)	where does the condenser fan VFD actually operate?	
27	Delta in operating Case Temp (oC)	possible to be set higher or lower	
28	Case Rated Runtime Fraction	actual runtime varies with application	conversations with Target
29	Delta in operating Walkin Temp (oC)	more likely to be set lower as safety precaution	
	Controls / Set Points		
1	Economizer mode OA temp low limit (oC)	can be adjusted easily at the BMS, depends on sensor	
2	Economizer mode OA temp high limit (oC)	can be adjusted easily at the BMS, depends on sensor	
3	Suply fan minimum flow (flow ratio)	where does the variable speed fan VFD actually operate?	
4	Supply fan medium flow (flow ratio)	where does the variable speed fan VFD actually operate?	
5	Max heat/cool point - tstat float (oC)	depends on sensor accuracy and setting at the BMS	
6	Cooling Setpoint (oC)	depends on sensor accuracy and setting at the BMS	
7	Heating Setpoint (oC)	depends on sensor accuracy and setting at the BMS	
16	Design heating supply air temp (oC)	depends on sensor accuracy and setting at the BMS	
19	Design heating supply air temp (oC)	depends on sensor accuracy and setting at the BMS	
21	Cycling run time (secs)	can be adjusted easily at the BMS	
22	Zone Mixing flow rate (m3/sec)	highly uncertain and dependent on zone configurations	

Table 30: Justification for parameter uncertainty choices

Table 31: List of screened model parameters and results of OAT Pre-Screening

Parameter Description (units):	μ / Initial value	min	max	SI Elec	SI NG	SI HVAC	SI Ref	SI WBE	Rank Elec	Rank NG	Rank EUI
Space infiltration reduction (%)	0	-40%	+40%	2.6%	60.4%	25.3%	0.2%	**17.5%**	4	1	**1**
Rated Gas Burner Eff (%)	80	60	81	0.0%	28.6%	0.0%	0.0%	**8.9%**	32	2	**2**
Heating Setpoint (oC)	15-20	-2	2	1.1%	15.7%	7.0%	1.2%	**6.0%**	7	3	**3**
Space light load multiplier	*various*	*-18.6%*	*+18.6%*	*13.2%*	*13.2%*	*78.5%*	*0.2%*	***5.7%***	*1*	*5*	***4***
Wall/Roof Conductivity multiplier	various	-20%	+20%	0.7%	13.5%	7.7%	0.5%	**4.9%**	10	4	**5**
Heating supply air temp (oC)	16.7	14.7	18.7	0.3%	10.6%	0.3%	1.4%	**3.5%**	13	6	**6**
Space plug load multiplier	*various*	*-18.2%*	*+18.2%*	*6.7%*	*6.0%*	*47.7%*	*0.2%*	***2.9%***	*2*	*9*	***7***
Heating sizing factor	1.2	1.1	1.4	0.3%	7.0%	2.1%	0.4%	**2.5%**	14	7	**8**
HX Sensible Effectiveness (%)	76	60	76	0.1%	6.4%	1.5%	0.2%	**2.1%**	19	8	**9**
Fan Pressure Rise (multiplier)	various	-30%	+30%	4.4%	4.0%	34.3%	0.1%	**1.9%**	3	11	**10**
Minimum system outside air ratio	0.3	0.3	0.5	0.1%	4.8%	1.6%	0.2%	**1.5%**	18	10	**11**
Rated roof absorptance	0.3	0.3	0.7	0.2%	3.5%	2.2%	0.0%	**0.9%**	15	12	**12**
Refrig suction piping UA value	various	-25%	+25%	1.2%	0.0%	17.5%	10.4%	**0.8%**	6	31	**13**
Delta in minimum ref. condensing temp (oC	21-24	-3	+3	1.1%	0.0%	0.3%	4.0%	**0.7%**	8	30	**14**
Fan Efficiency (%)	45	35	48	1.5%	1.2%	14.0%	0.0%	**0.7%**	5	13	**15**
Rated Cool Coil COP	4	3.5	4.1	0.9%	0.0%	7.9%	0.0%	**0.6%**	9	31	**16**
Anti-sweat energy multiplier	various	0.5	1	0.2%	0.6%	0.2%	0.5%	**0.3%**	17	15	**17**
Max heat/cool point - tstat float (oC)	0.278	0.2	0.8	0.4%	0.1%	4.9%	0.5%	**0.3%**	12	22	**18**
Operating walkin ref temp (oC)	0	-3	+1	0.4%	0.1%	1.7%	0.9%	**0.2%**	11	21	**19**
Cooling Setpoint (oC)	23-27	-2	2	0.2%	0.3%	1.8%	0.0%	**0.2%**	16	18	**20**
Space thermal mass multiplier	various	-30%	+30%	0.1%	0.8%	0.4%	0.1%	**0.2%**	21	14	21
System cycling run time (secs)	1800	1200	2400	0.0%	0.5%	0.1%	0.0%	**0.2%**	27	16	22
Wall Specific Heat multiplier	various	-29%	+29%	0.0%	0.4%	0.3%	0.1%	**0.1%**	26	17	23
Operating ref case temp (oC)	various	-2	+3	0.0%	0.2%	0.4%	0.3%	**0.1%**	25	19	24
Cooling sizing factor	1.2	1.1	1.4	0.1%	0.2%	1.5%	0.4%	**0.1%**	24	20	25
Supply fan medium flow (flow ratio)	0.8	0.7	0.9	0.1%	0.1%	0.7%	0.0%	**0.1%**	20	25	26
Suply fan minimum flow (flow ratio)	0.4	0.3	0.5	0.1%	0.0%	0.6%	0.0%	**0.1%**	22	27	27
Economizer mode OA temp high limit (oC)	18.3	16.3	20.3	0.1%	0.0%	0.3%	0.1%	**0.0%**	23	26	28
Building Occupancy (total people)	160	48	480	0.0%	0.1%	0.6%	0.2%	**0.0%**	29	23	29
Economizer mode OA temp low limit (oC)	10	8	12	0.0%	0.1%	0.4%	0.2%	**0.0%**	28	24	30
Cooling supply air temp (oC)	12.8	10.8	14.8	0.0%	0.0%	0.0%	0.0%	**0.0%**	30	28	31
Ratio Evap rate and Latent Cap	1.5	1	2	0.0%	0.0%	0.0%	0.0%	**0.0%**	31	29	32
Fan motor Efficiency (%)	90	75	93	0.0%	0.0%	0.0%	0.0%	**0.0%**	32	31	33
Zone Mixing flow rate (m3/sec)	2	0.5	3.5	0.0%	0.0%	0.0%	0.0%	**0.0%**	32	31	33
Fraction of anti-sweat energy to case load	0.85	0.7	1	0.0%	0.0%	0.0%	0.0%	**0.0%**	32	31	33
Condenser fans min flow (%)	20	20	50	0.0%	0.0%	0.0%	0.0%	**0.0%**	32	31	33
Ref case rated runtime fraction	95	70	95	0.0%	0.0%	0.0%	0.0%	**0.0%**	32	31	33

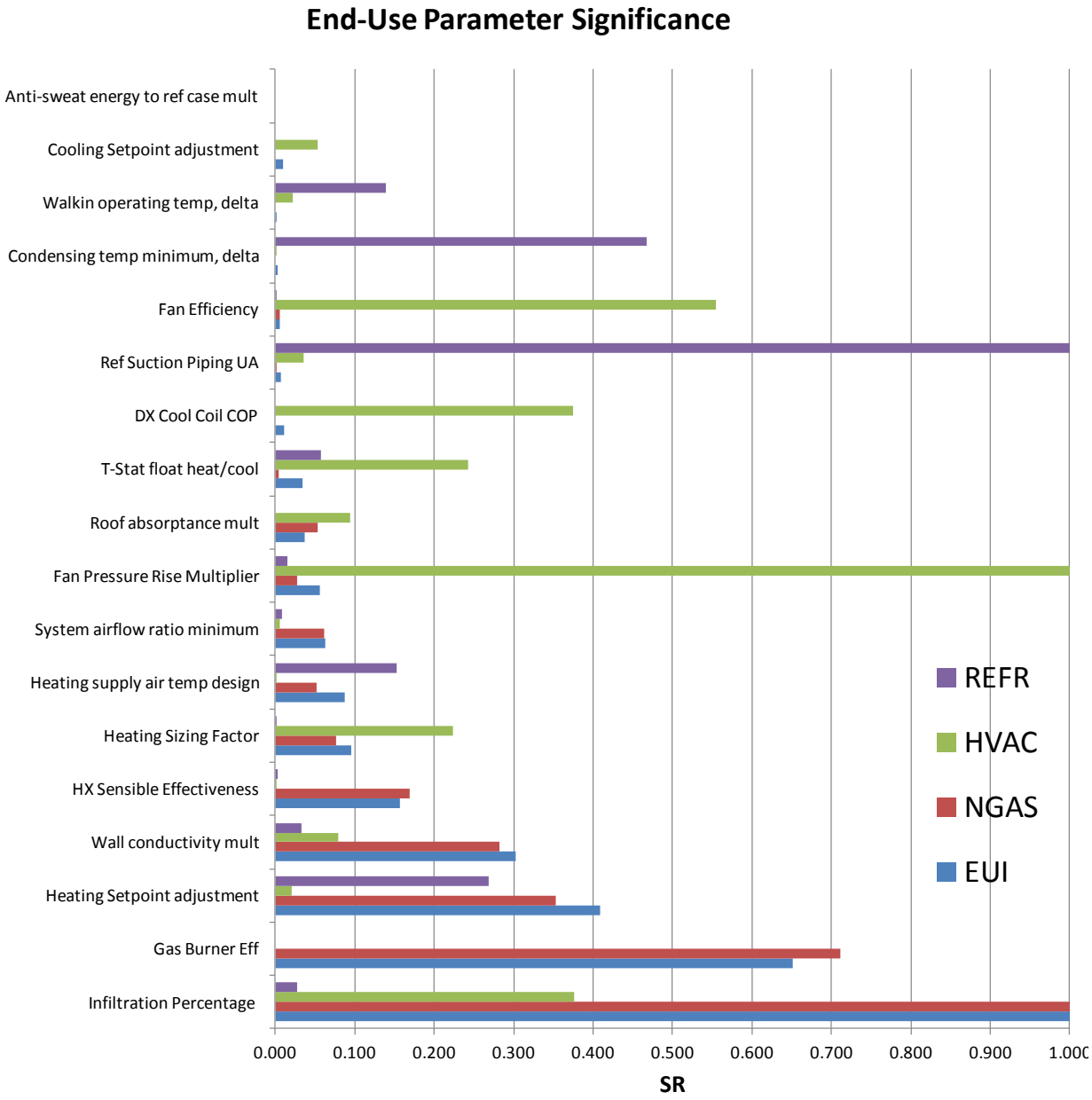


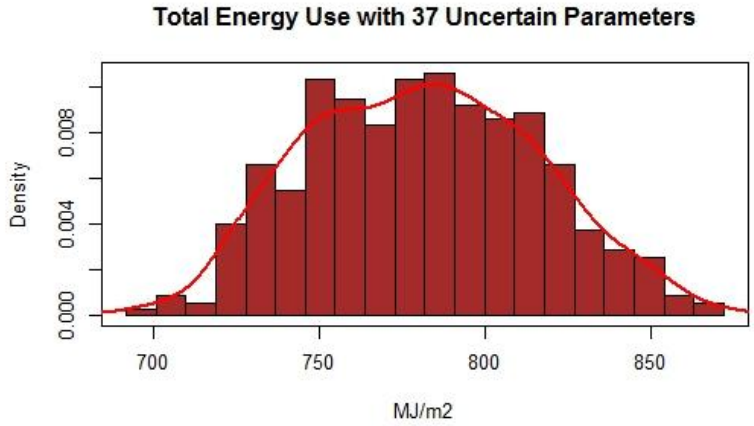
Figure 87: Parameter significance ratios obtained by using the regression method

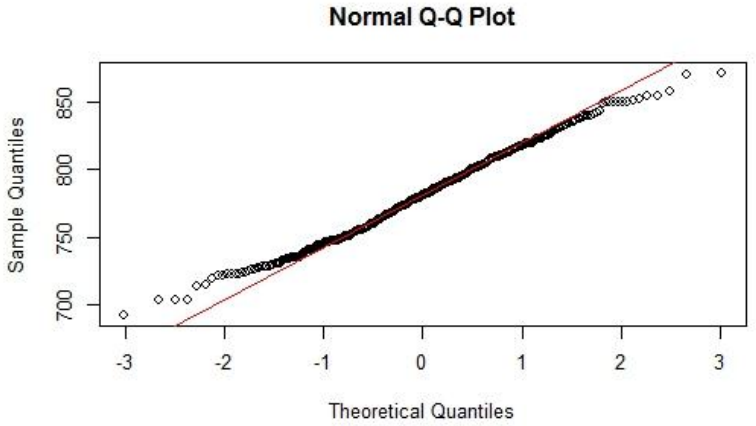
11.3 Appendix C: Full Results of Global Sensitivity Analysis

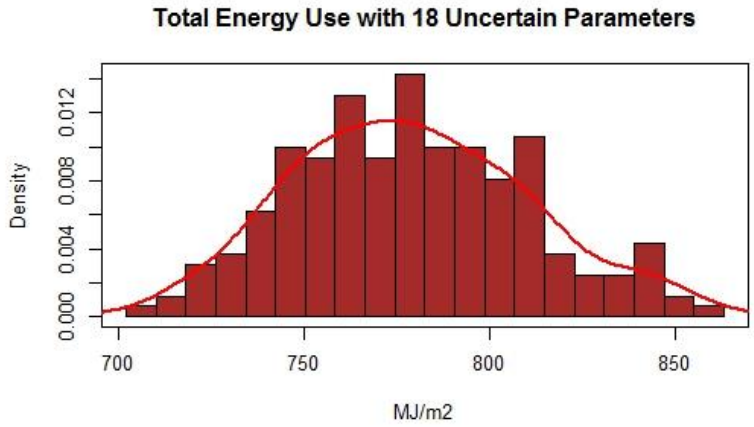
Table 32: Comparison of OAT and Multivariate Regression Scaled Parameter Significance Ratios

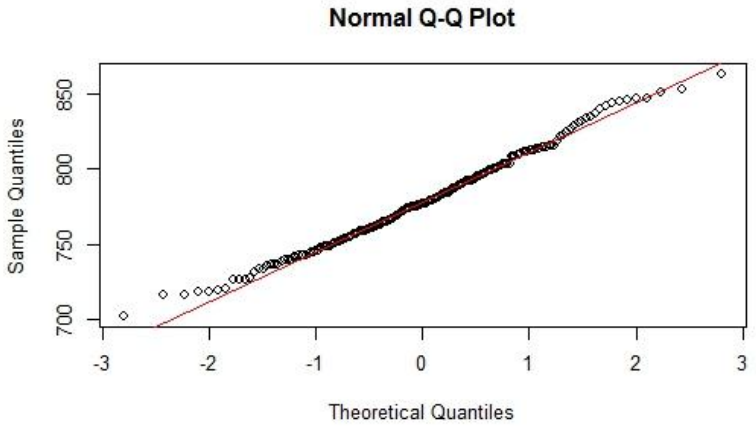
Parameter	WBE		NGAS		HVAC		REFR	
	MV	OAT	MV	OAT	MV	OAT	MV	OAT
Space infiltration reduction (%)	1.000	1.000	1.000	1.000	0.376	0.739	0.027	0.018
Gas Burner Eff (%)	0.652	0.509	0.711	0.473	0.000	0.000	0.000	0.000
Heating Setpoint (oC)	0.408	0.341	0.353	0.260	0.021	0.205	0.269	0.115
Wall/Roof Conductivity multiplier	0.302	0.281	0.282	0.223	0.079	0.226	0.033	0.052
Heating supply air temp (oC)	0.087	0.202	0.052	0.175	0.001	0.009	0.153	0.132
Heating sizing factor	0.095	0.141	0.076	0.115	0.223	0.062	0.002	0.034
HX Sensible Effectiveness (%)	0.157	0.120	0.169	0.107	0.001	0.043	0.003	0.016
Fan Pressure Rise (multiplier)	0.056	0.107	0.027	0.067	1.000	1.000	0.015	0.011
Minimum system outside air ratio	0.062	0.084	0.061	0.079	0.006	0.046	0.008	0.018
Rated roof absorptance	0.037	0.051	0.053	0.058	0.094	0.064	0.000	0.001
Refrig suction piping UA value	0.007	0.047	0.002	0.000	0.036	0.510	1.000	1.000
Condensing temp minimum, delta	0.003	0.042	0.000	0.000	0.001	0.008	0.467	0.383
Fan Efficiency (%)	0.005	0.037	0.006	0.019	0.555	0.408	0.001	0.000
DX Cool Coil COP	0.011	0.034	0.000	0.000	0.375	0.232	0.000	0.000
Anti-sweat energy multiplier	0.000	0.018	0.000	0.010	0.000	0.007	0.000	0.048
T-Stat float heat/cool (oC)	0.034	0.016	0.004	0.002	0.243	0.142	0.057	0.053
Operating walkin ref temp (oC)	0.002	0.014	0.000	0.002	0.022	0.051	0.139	0.082
Cooling Setpoint (oC)	0.010	0.012	0.000	0.005	0.053	0.053	0.000	0.004

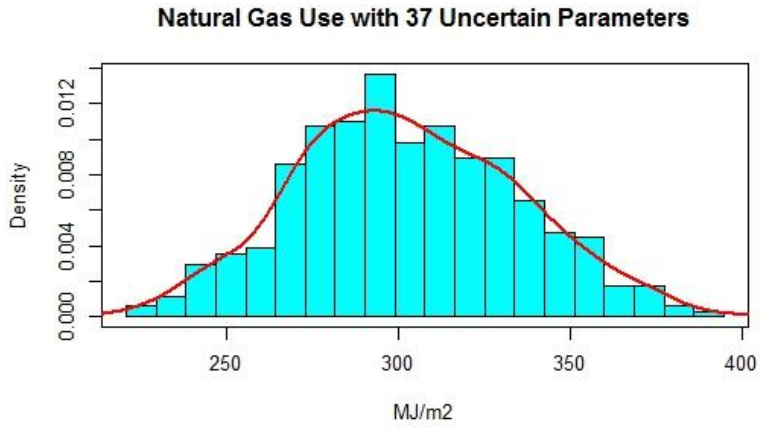
11.4 Appendix D: Calibration Validation Results by End Use

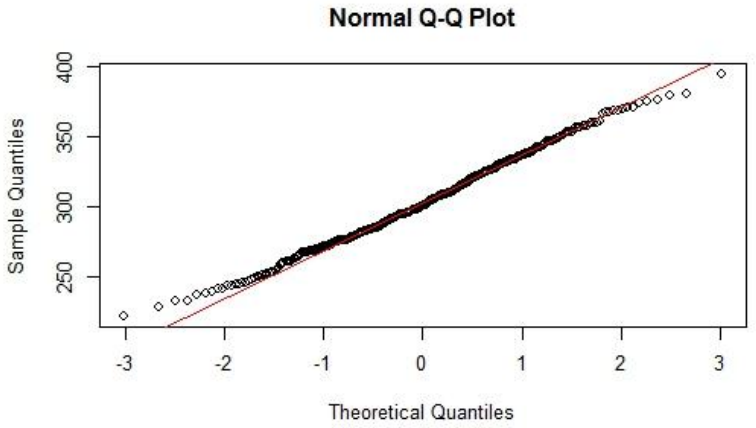


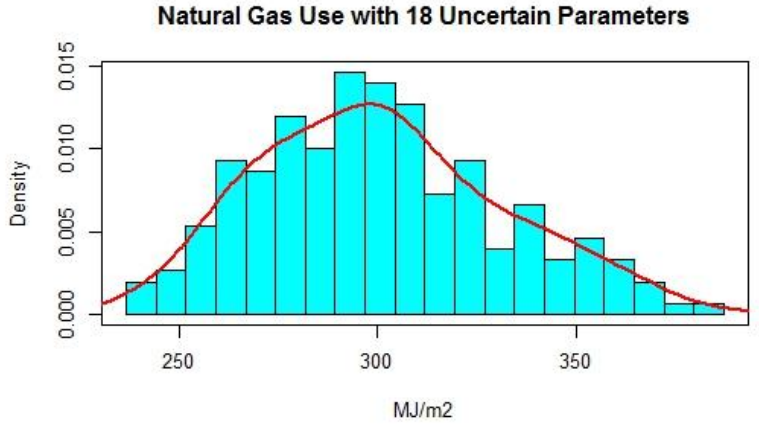


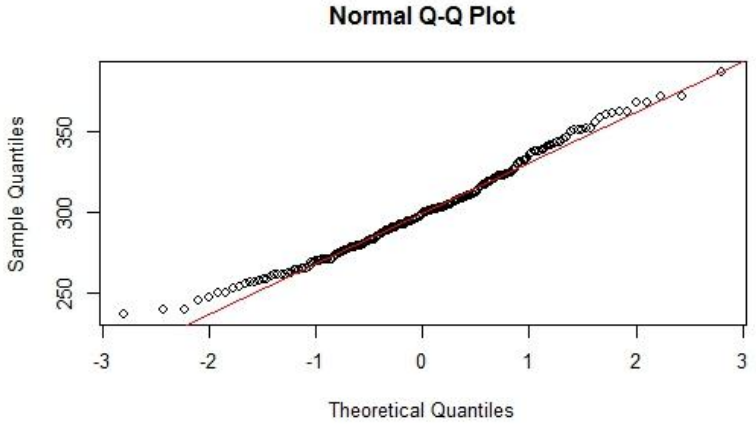


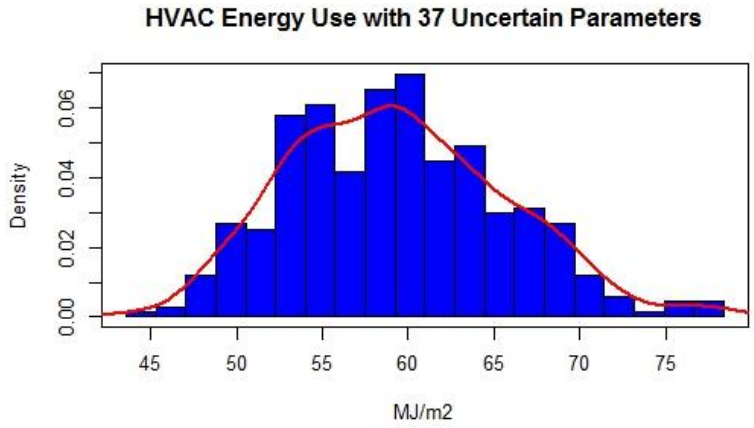


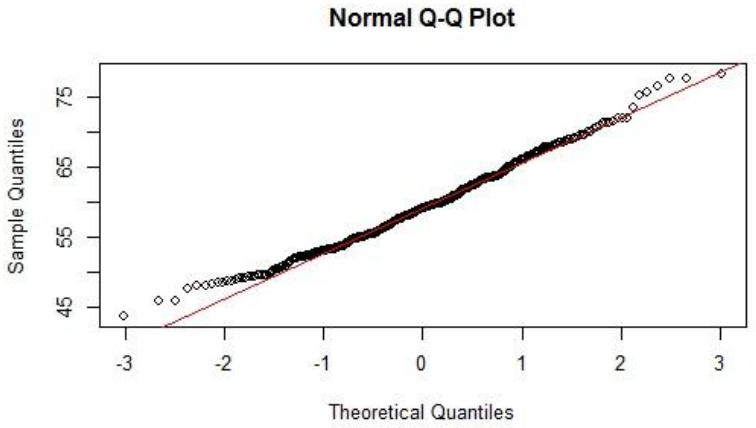

Natural Gas Use with 37 Uncertain Parameters
Density
0.000
0.004
0.008
0.012
250
300
350
400
MJ/m2

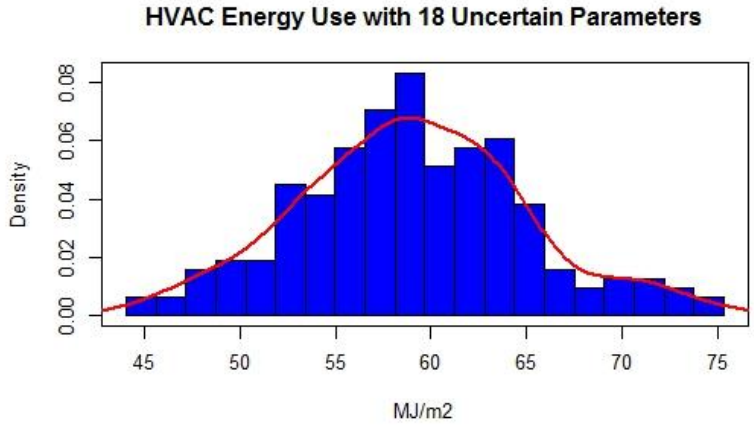

Normal Q-Q Plot
Sample Quantiles
250
300
350
400
-3
-2
-1
0
1
2
3
Theoretical Quantiles

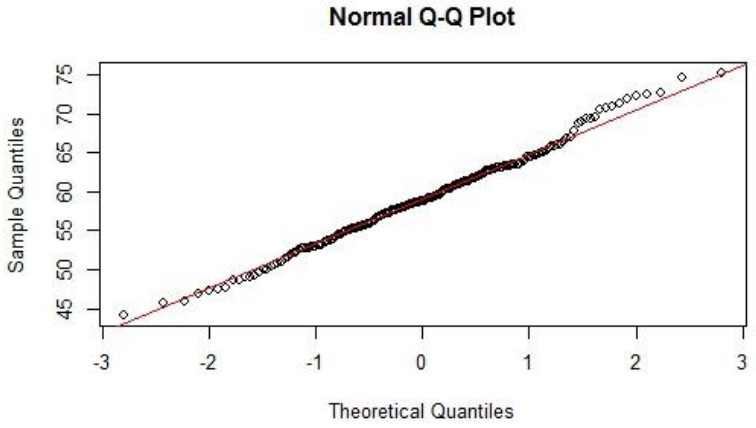

Natural Gas Use with 18 Uncertain Parameters
Density
0.000
0.005
0.010
0.015
250
300
350
MJ/m2

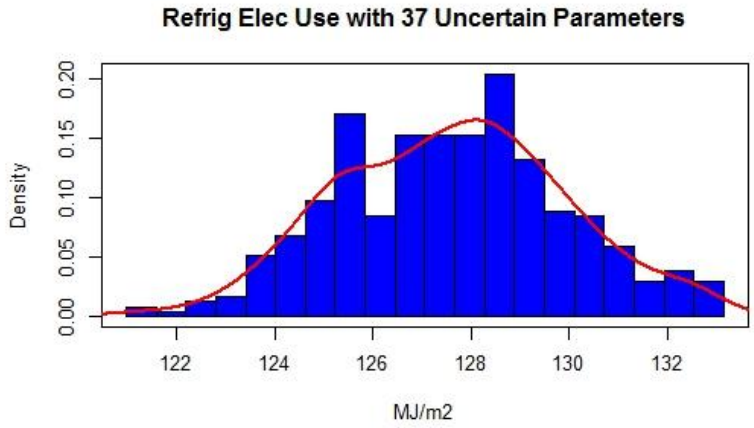

Normal Q-Q Plot
Sample Quantiles
250
300
350
-3
-2
-1
0
1
2
3
Theoretical Quantiles

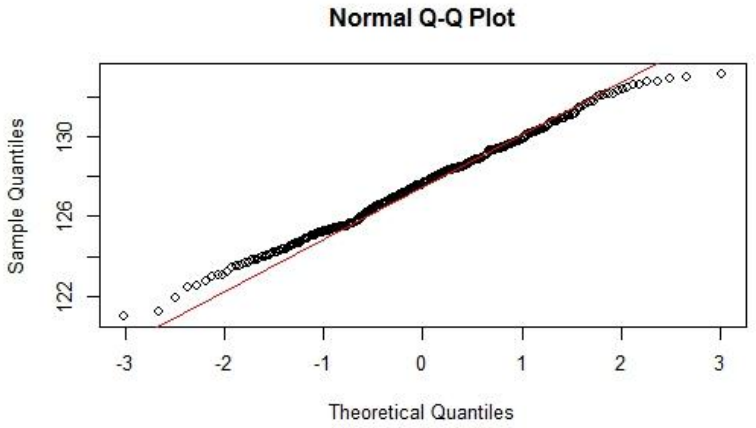

HVAC Energy Use with 37 Uncertain Parameters
Density
0.00
0.02
0.04
0.06
45
50
55
60
65
70
75
MJ/m2

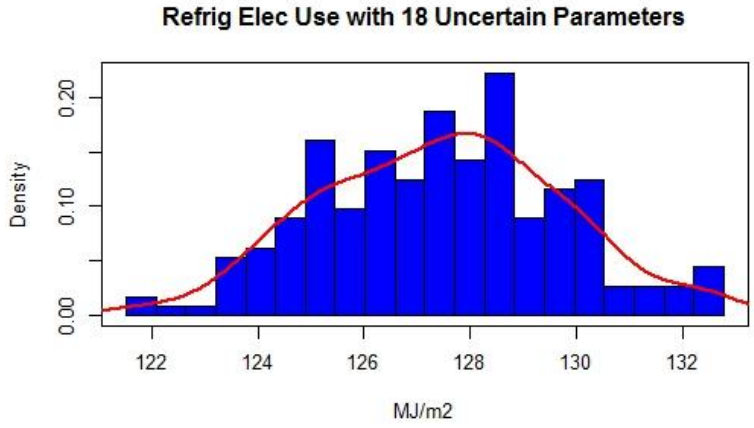

Normal Q-Q Plot
Sample Quantiles
45
55
65
75
-3
-2
-1
0
1
2
3
Theoretical Quantiles

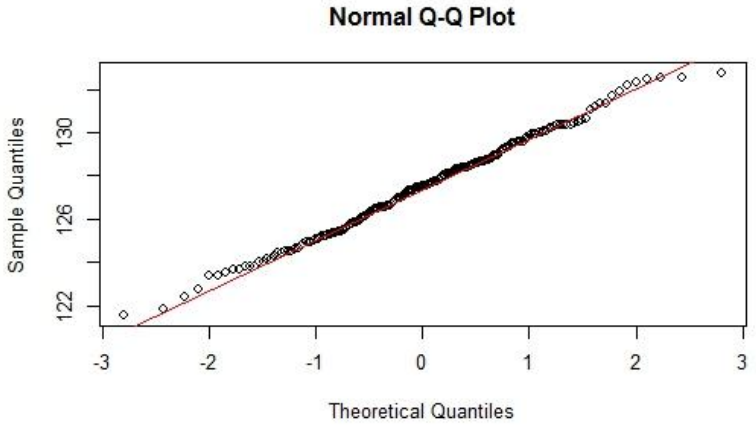

HVAC Energy Use with 18 Uncertain Parameters
Density
0.00
0.02
0.04
0.06
0.08
45
50
55
60
65
70
75
MJ/m2


Normal Q-Q Plot
Sample Quantiles
45
50
55
60
65
70
75
-3
-2
-1
0
1
2
3
Theoretical Quantiles


Refrig Elec Use with 37 Uncertain Parameters
Density
0.00
0.05
0.10
0.15
0.20
122
124
126
128
130
132
MJ/m2


Normal Q-Q Plot
Sample Quantiles
122
126
130
-3
-2
-1
0
1
2
3
Theoretical Quantiles


Refrig Elec Use with 18 Uncertain Parameters
Density
0.00
0.10
0.20
122
124
126
128
130
132
MJ/m2


Normal Q-Q Plot
Sample Quantiles
122
126
130
-3
-2
-1
0
1
2
3
Theoretical Quantiles

11.5 Appendix E: Calibration Statistics

The following plots show the quantity and uniqueness of the calibrated solutions generated by performing LHS sampling in various batch sizes. Due to the quasi-random nature of the LHS sampling algorithm, runs made with identical batch sizes produced identical or nearly identical results. Due to this, and to the practical limits on sample size as discussed above, additional unique solutions were achieved in the final large sample set by varying the sample size to avoid common multiples,

and concatenating these results into one large data set.

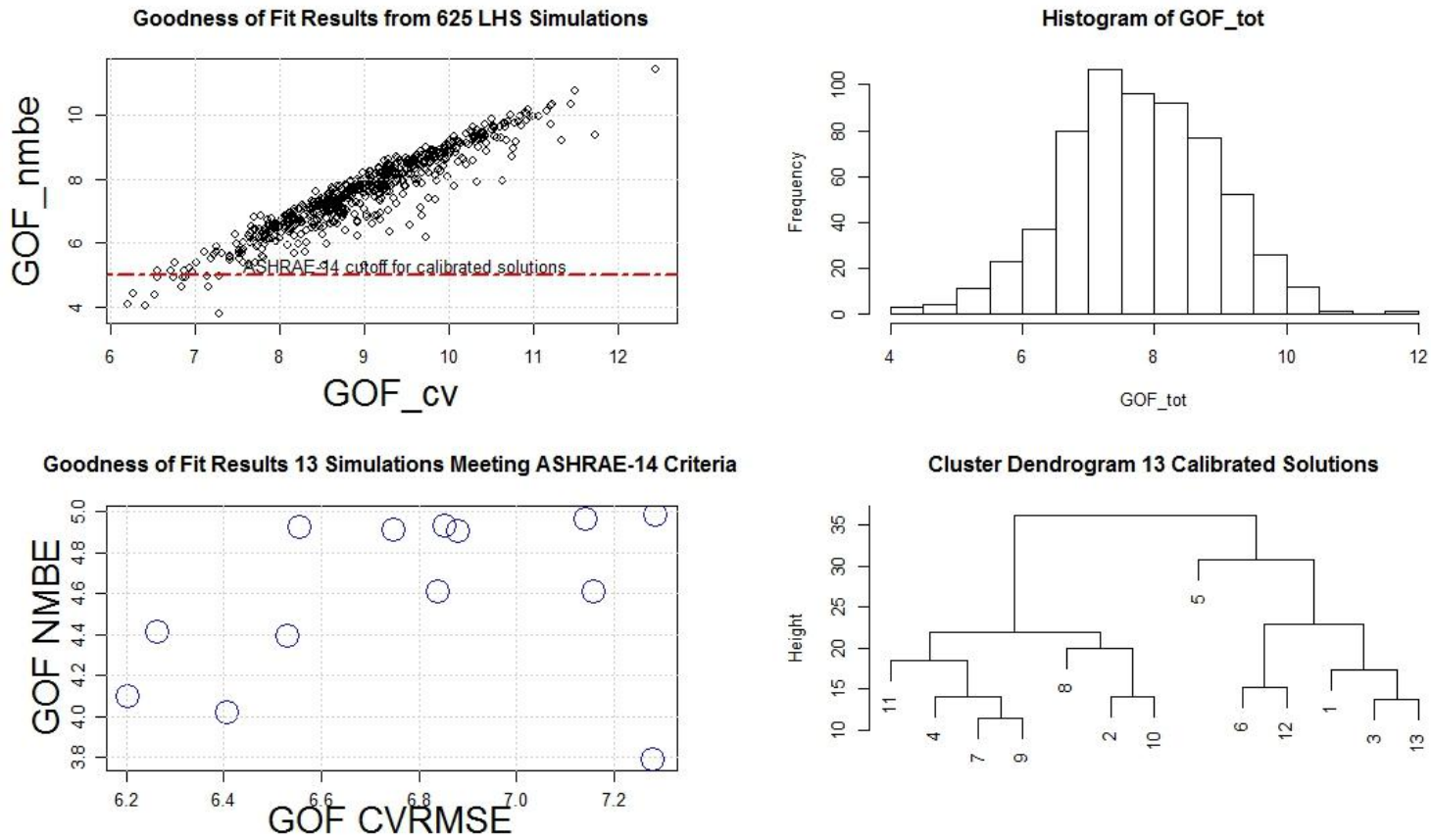


Figure 88: Calibration statistics for 625 monte carlo sample size

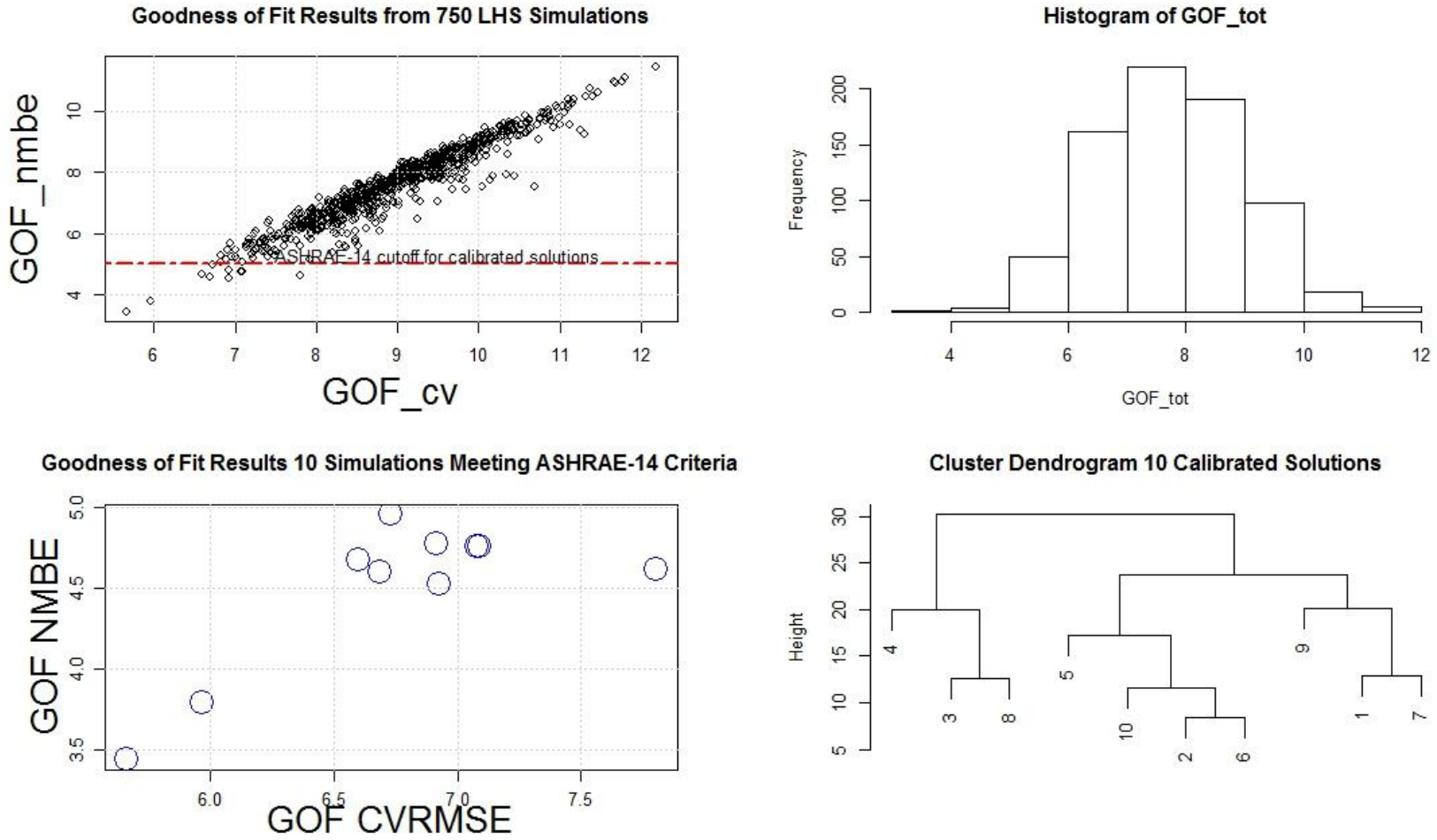


Figure 89: Calibration statistics for 750 monte carlo sample size

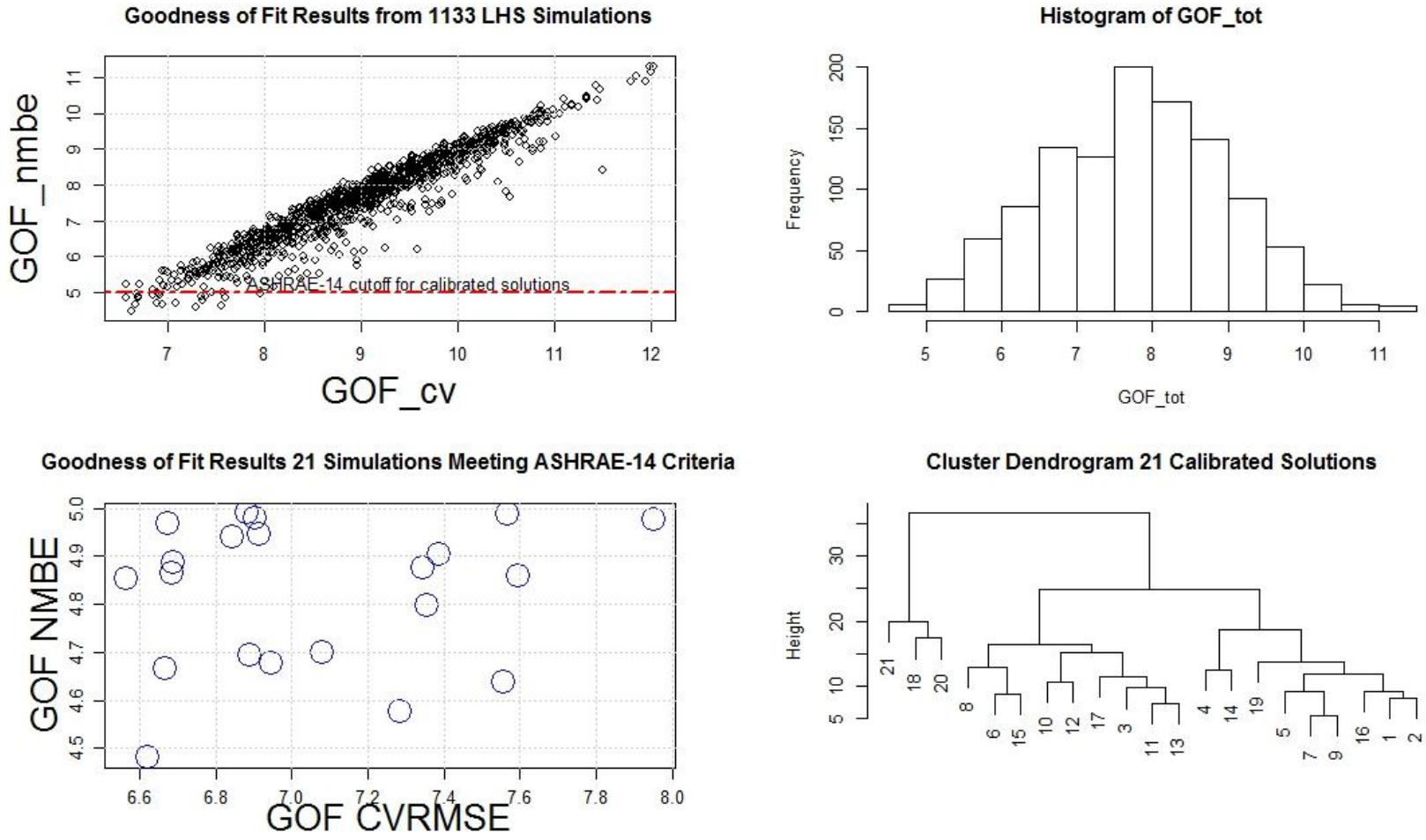


Figure 90: : Calibration statistics for 1133 monte carlo sample size

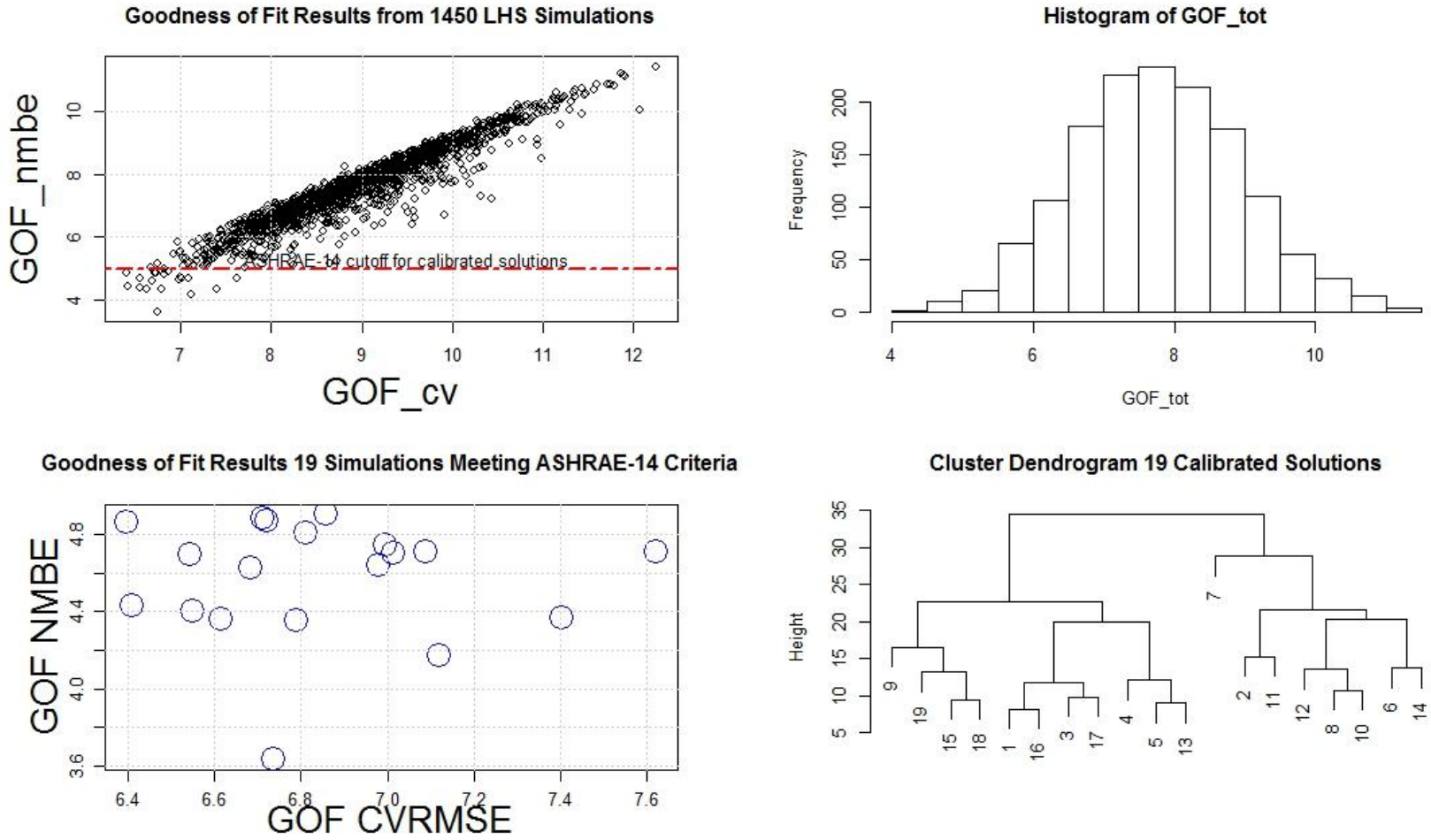


Figure 91: Calibration statistics for 1450 monte carlo sample size

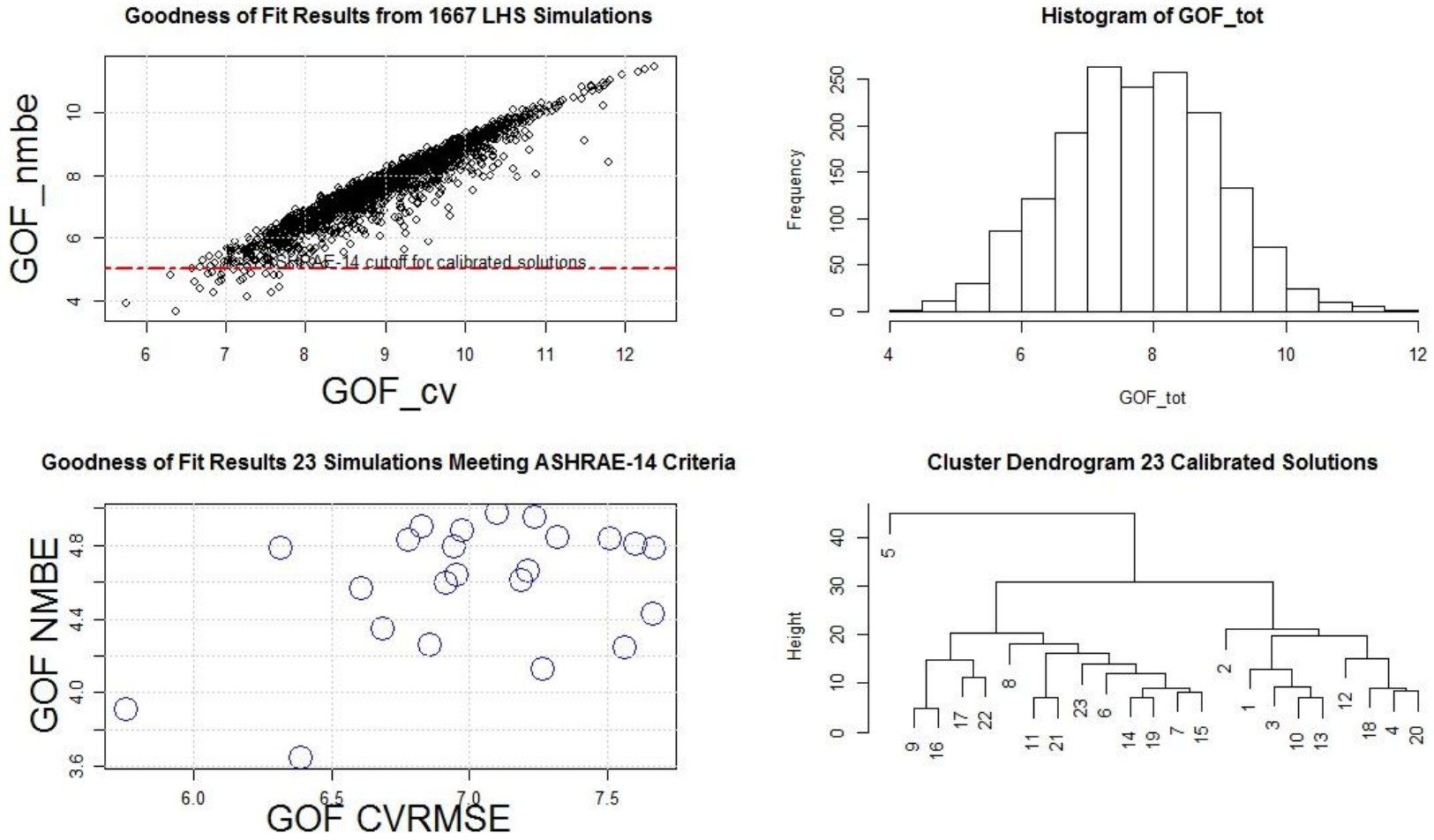


Figure 92: Calibration statistics for 1667 monte carlo sample size

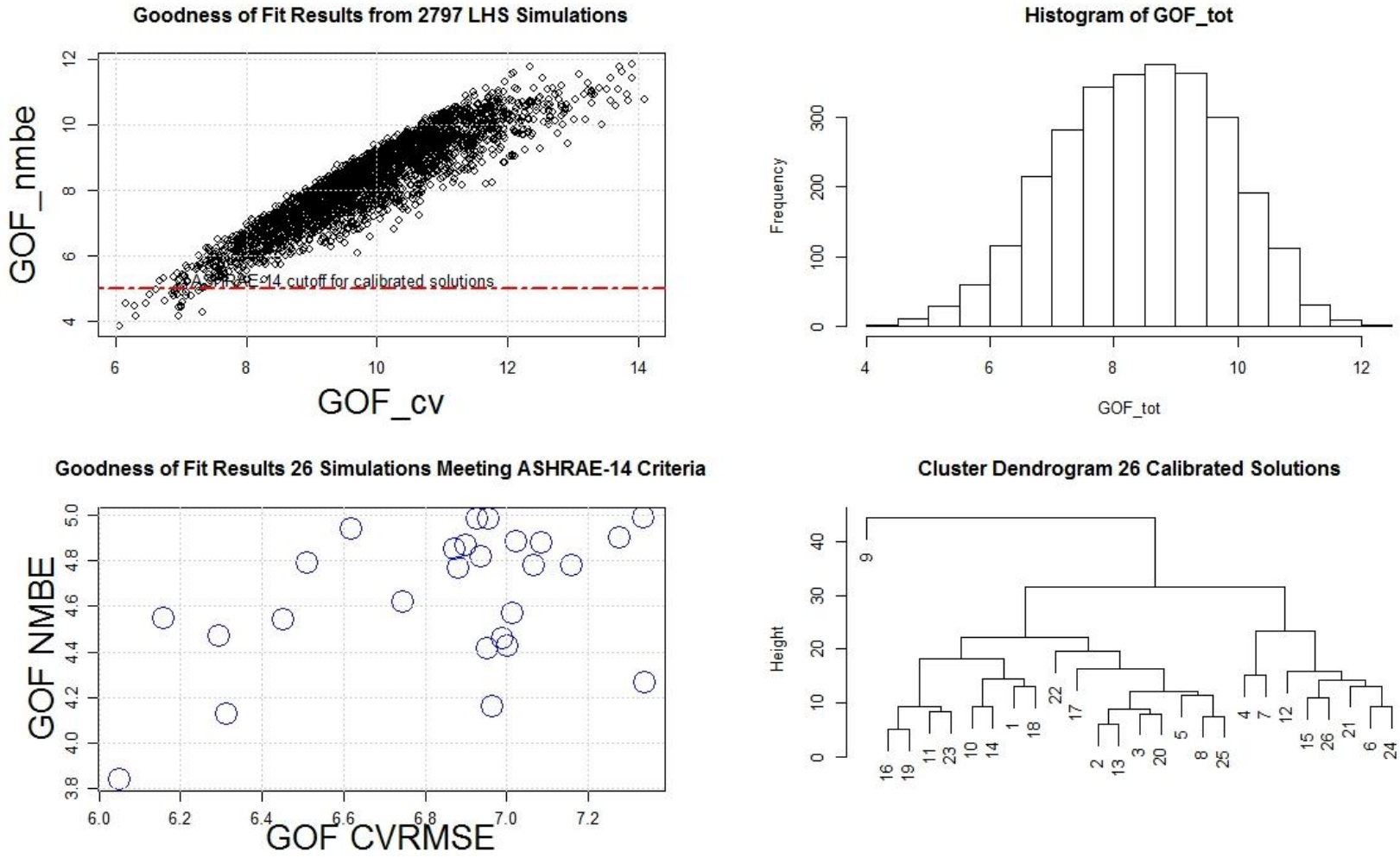


Figure 93: Calibration statistics for 2797monte carlo sample size

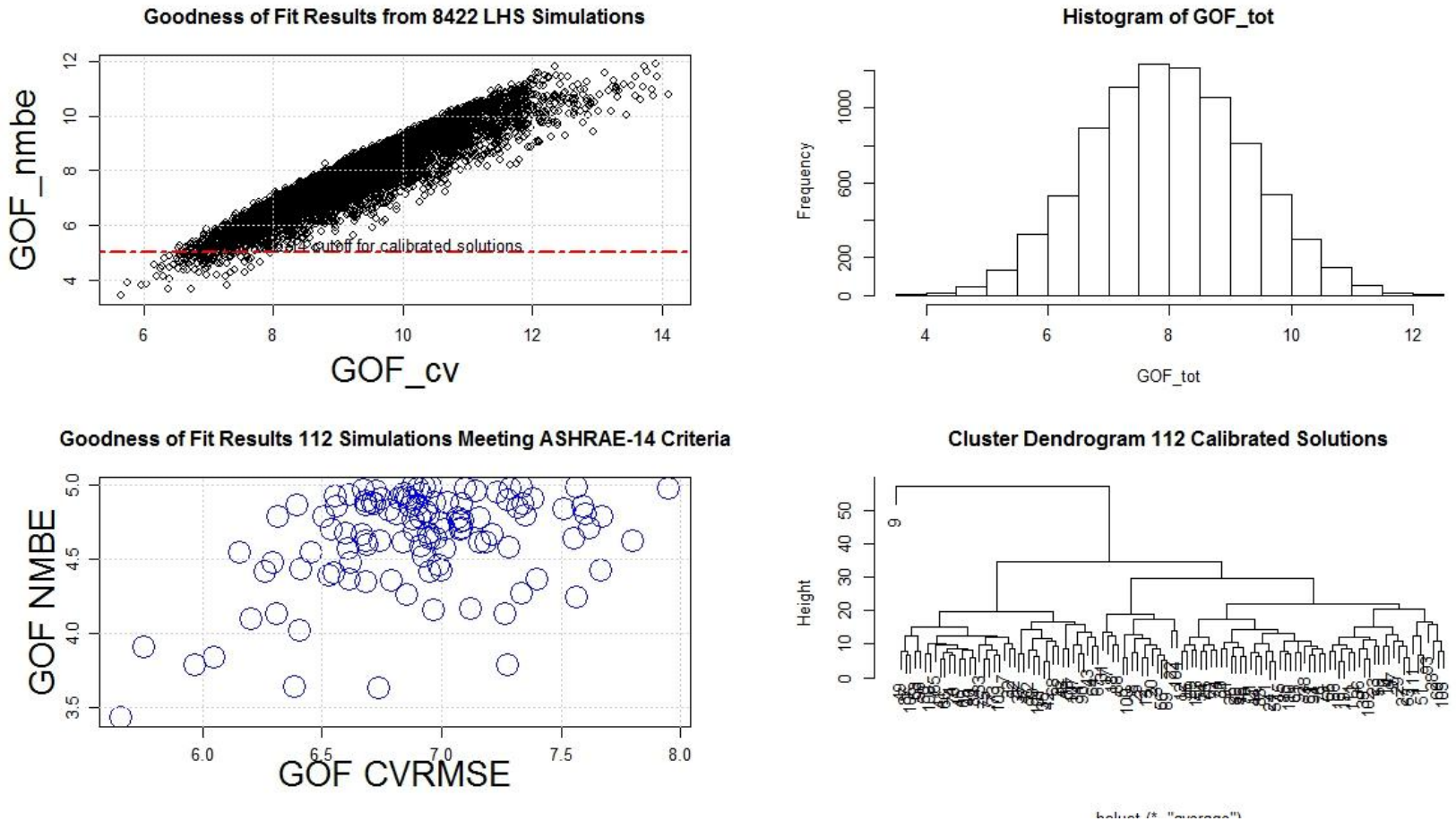


Figure 94: Calibration statistics for agglomerated batches

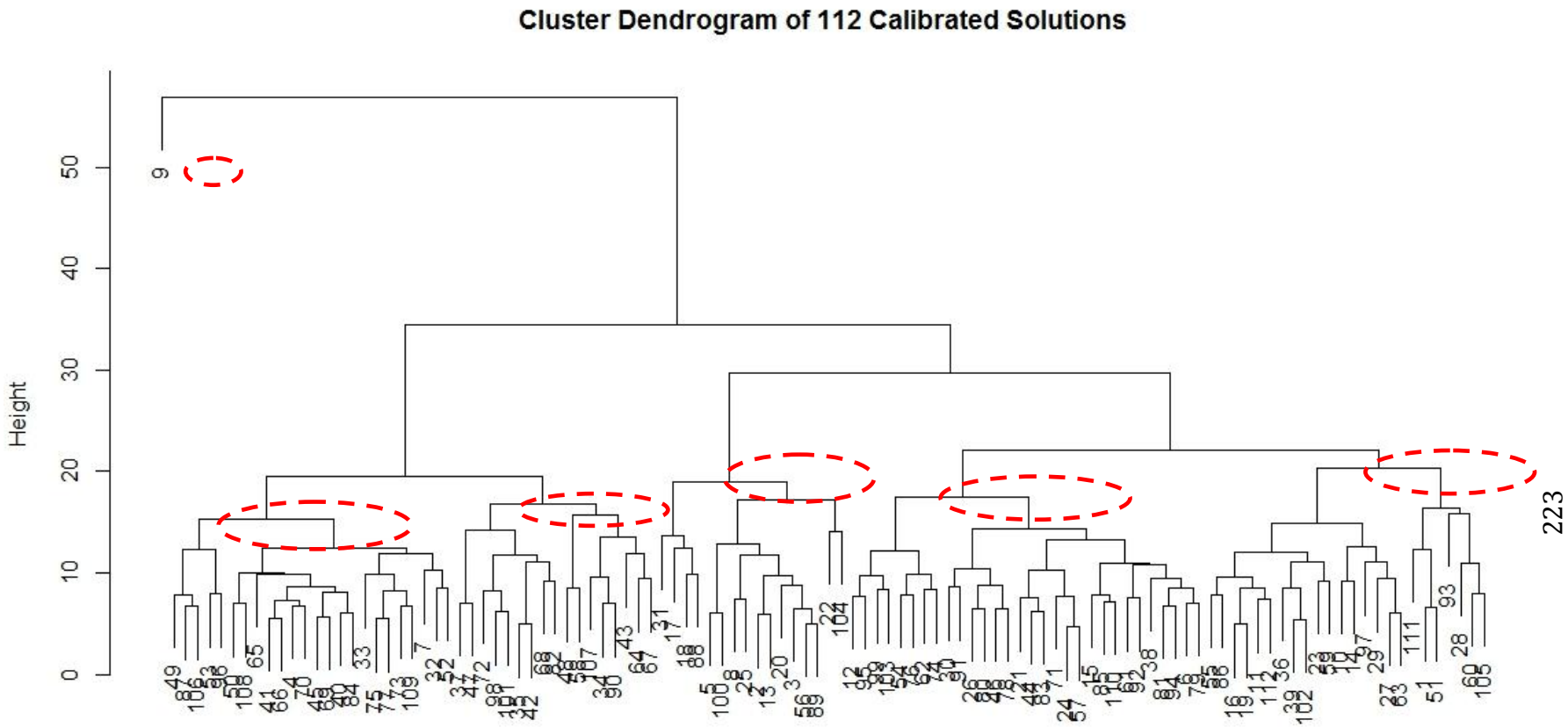


Figure 95: Final results of clustering analysis for the set of model calibration solutions, showing six natural crisp divisions.

11.6 Appendix F: Regression Statistics

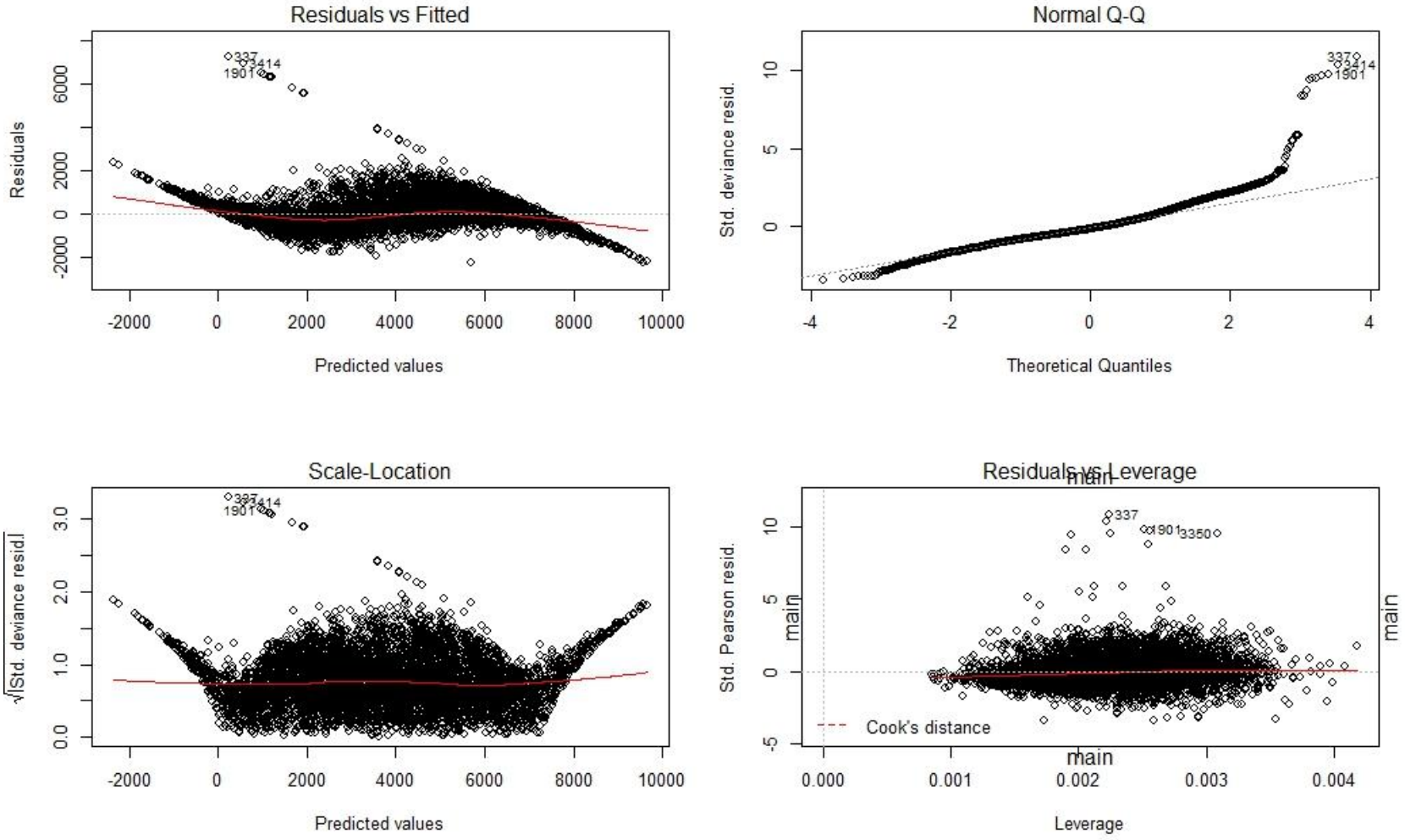


Figure 96: EUI, ranked

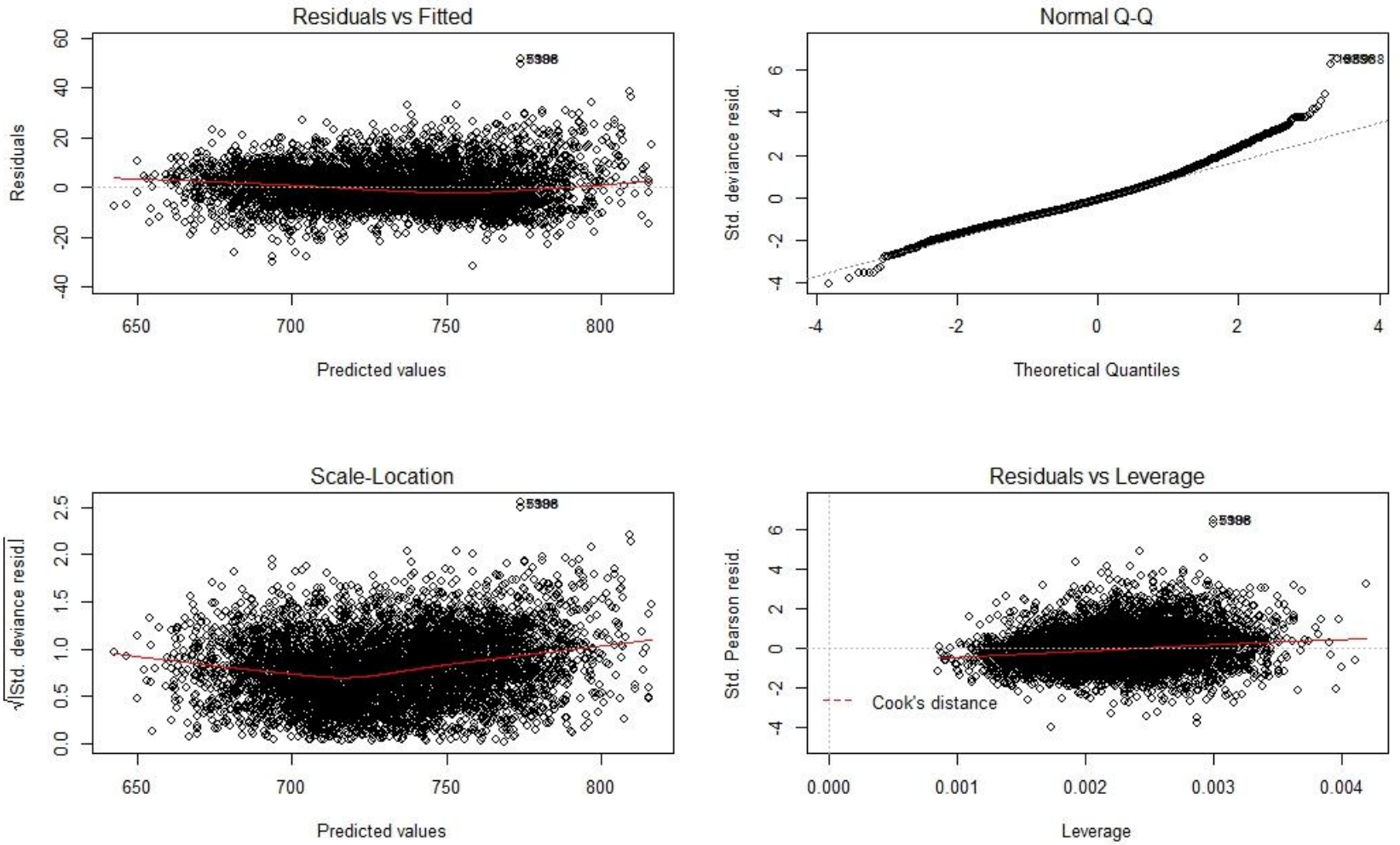


Figure 97: EUI, Un-ranked

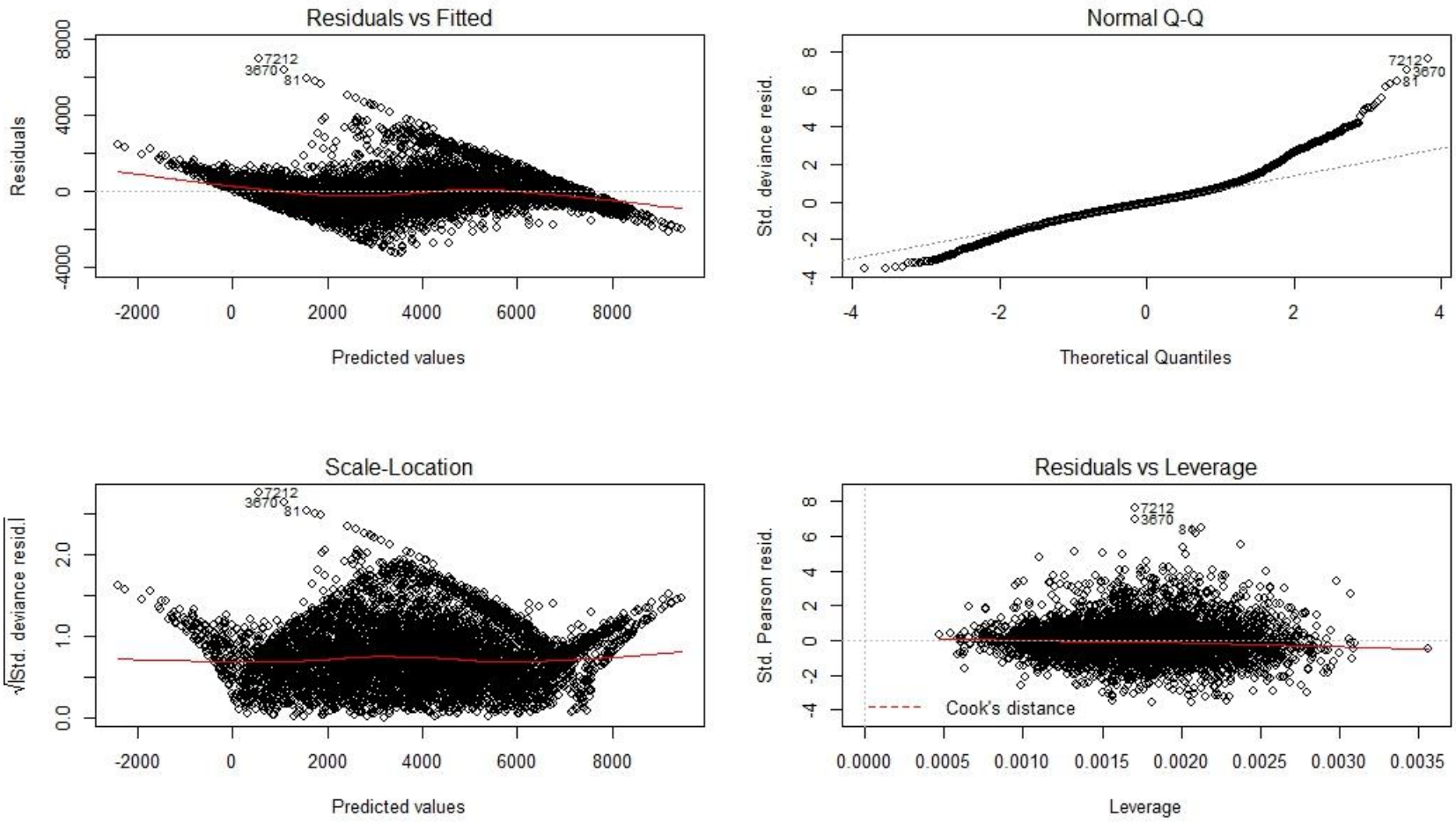


Figure 98:Refrigeration, Ranked

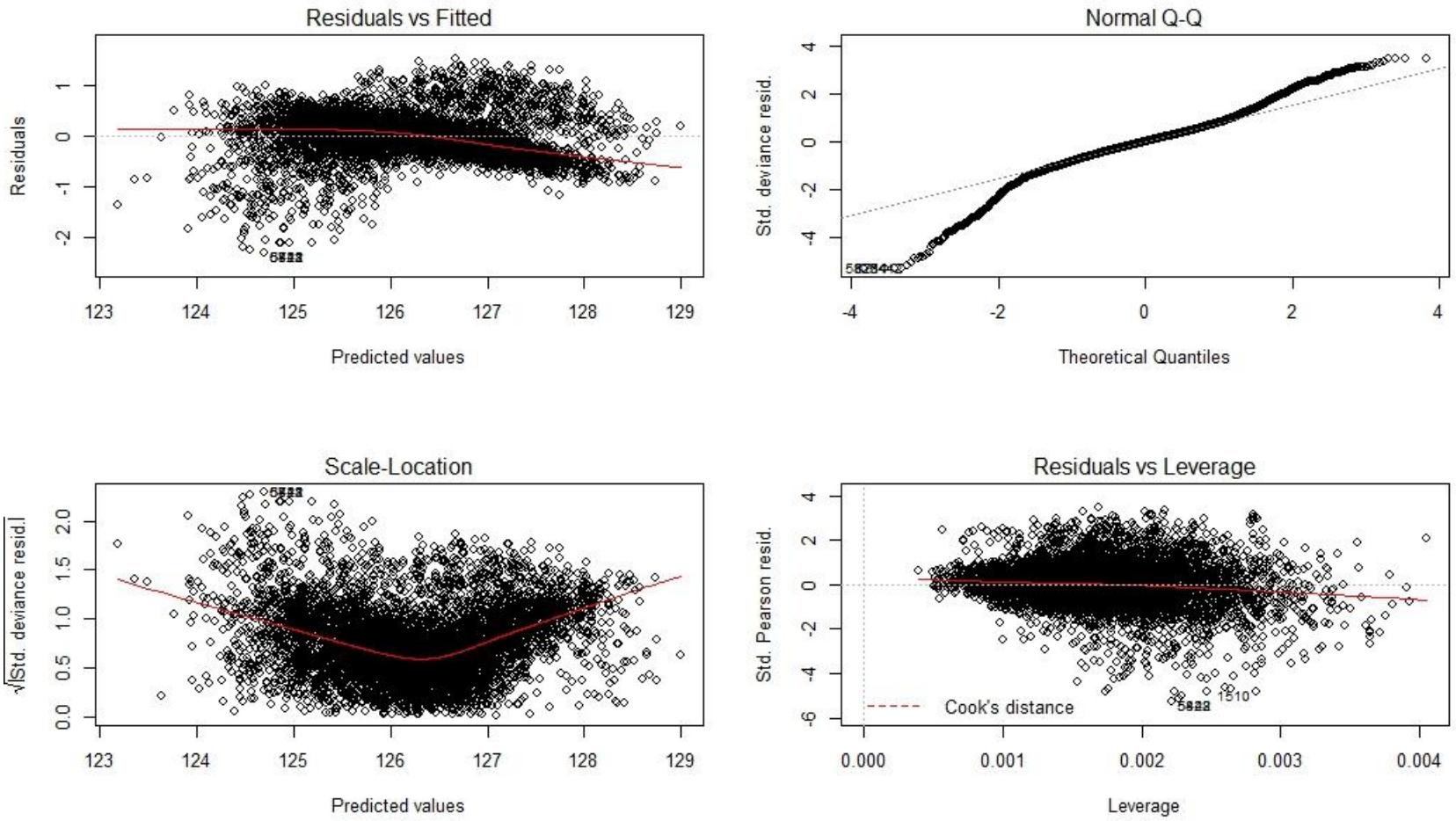
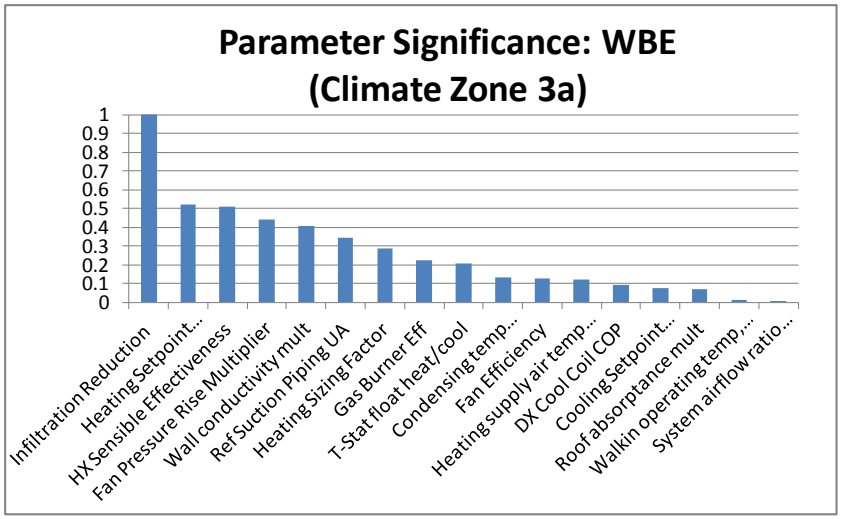
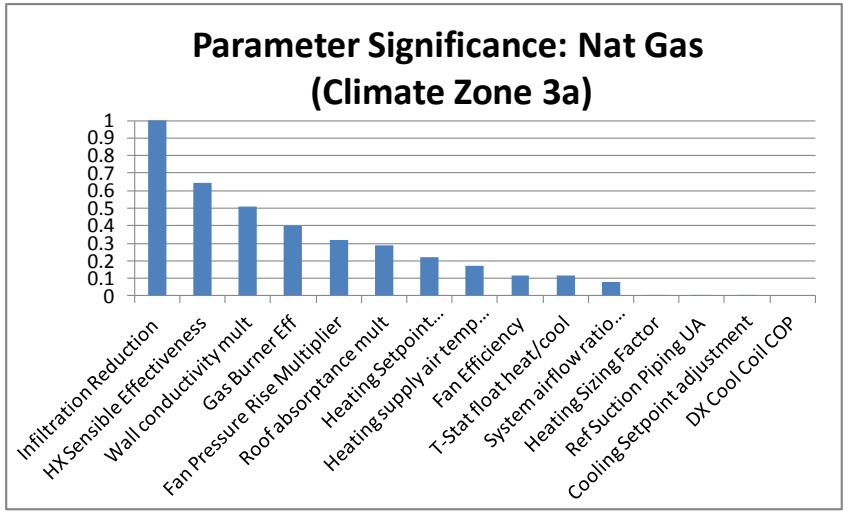
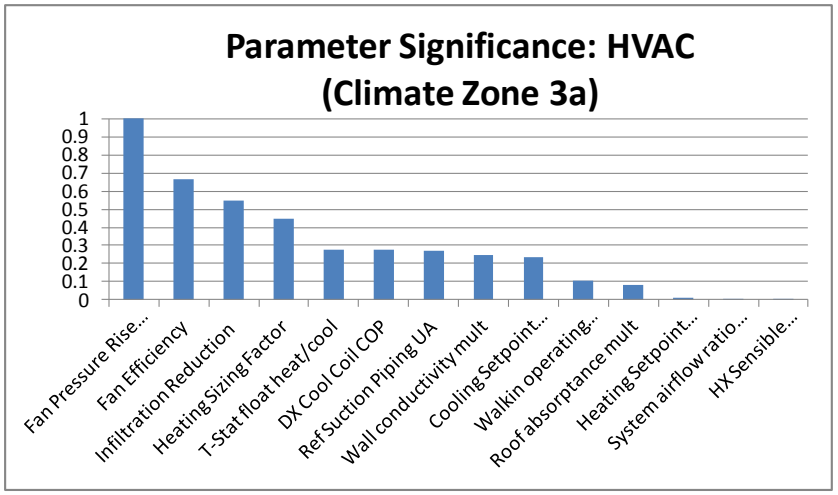


Figure 99: Refrigeration, Un-ranked







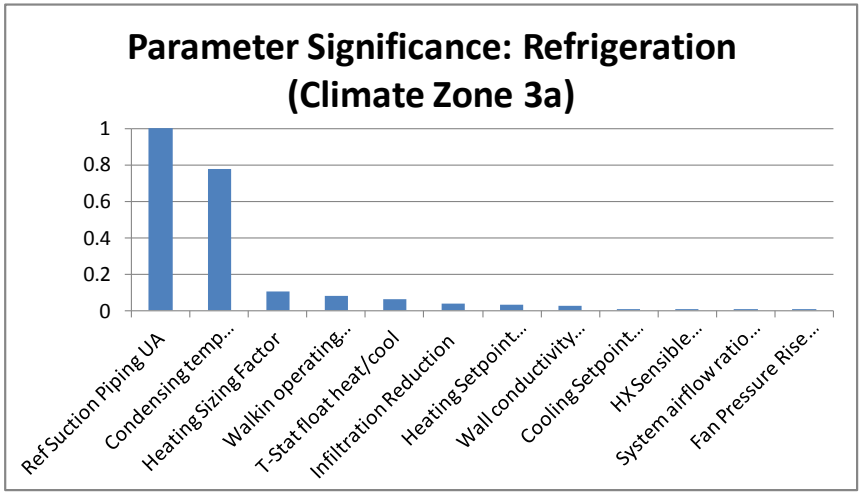


Figure 100: Significance index values for all end-uses. Case Study building

Table 33: Parameter significance values for climate zones 6a, 4c, and 3a

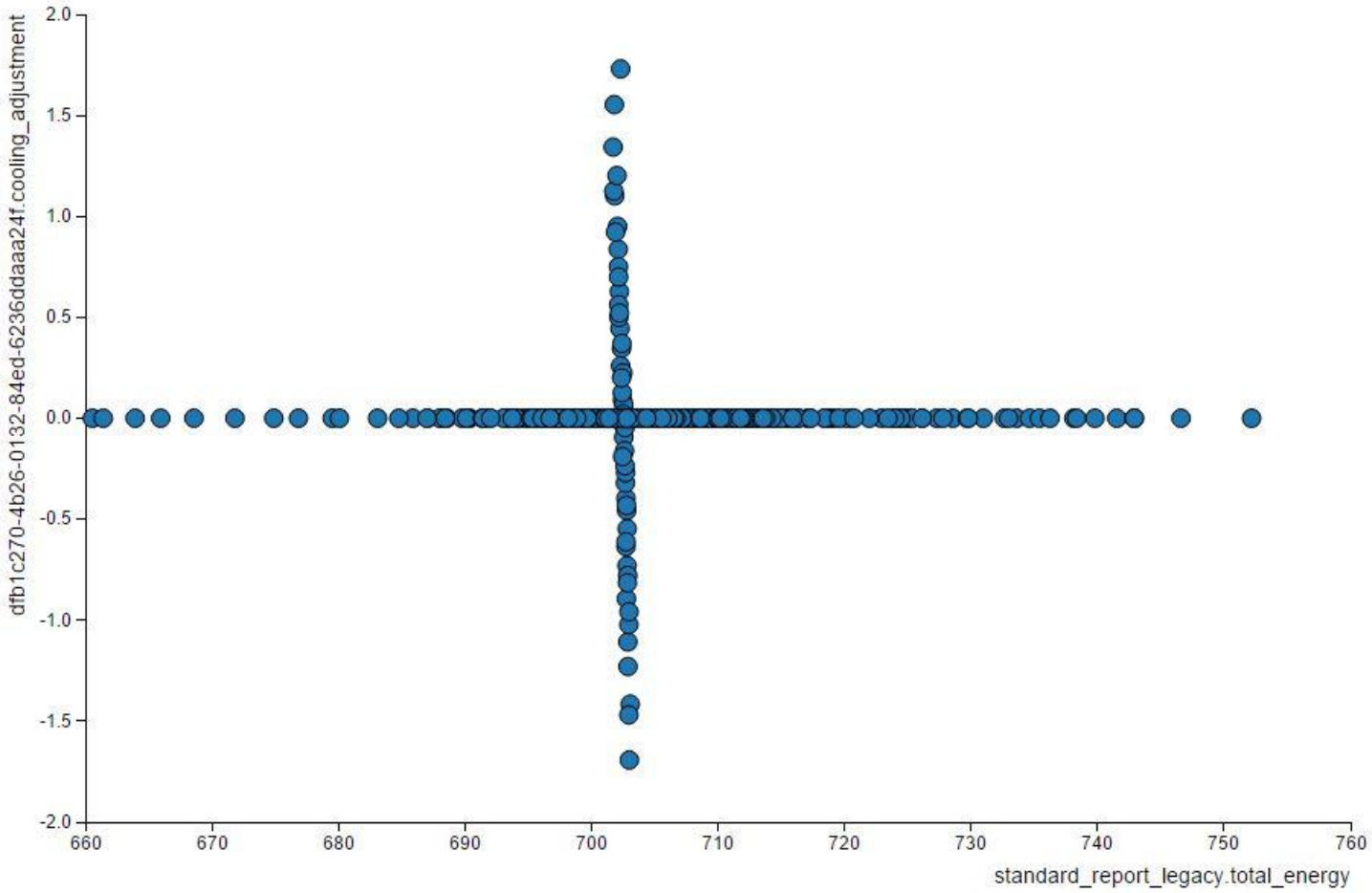
Parameter	WBE			NGAS			HVAC			REFR		
	6a	4c	3a	6a	4c	3a	6a	4c	3a	6a	4c	3a
Space infiltration reduction (%)	1.000	1.000	1.000	1.000	1.000	1.000	0.376	0.393	0.662	0.027	0.116	0.042
Gas Burner Eff (%)	0.652	0.315	0.223	0.711	0.375	0.319	NA	NA	NA	NA	NA	NA
Heating Setpoint (oC)	0.408	0.170	0.525	0.353	0.195	0.509	0.021	0.002	0.277	0.269	0.033	0.033
Wall/Roof Conductivity multiplier	0.302	0.385	0.410	0.282	0.410	0.399	0.079	0.062	0.246	0.033	0.004	0.025
Heating supply air temp (oC)	0.087	0.150	0.121	0.052	0.179	0.171	0.001	0.000	0.003	0.153	NA	NA
Heating sizing factor	0.095	0.040	0.290	0.076	0.001	0.114	0.223	0.202	0.446	0.002	0.021	0.102
HX Sensible Effectiveness (%)	0.157	0.646	0.513	0.169	0.704	0.646	0.001	0.000	0.000	0.003	0.001	0.001
Fan Pressure Rise (multiplier)	0.056	0.194	0.442	0.027	0.329	0.288	1.000	1.000	1.000	0.015	NA	NA
Minimum system outside air ratio	0.062	0.006	0.001	0.061	0.006	0.003	0.006	0.002	0.000	0.008	NA	NA
Rated roof absorptance	0.037	0.178	0.068	0.053	0.236	0.221	0.094	0.003	0.106	NA	NA	NA
Refrig suction piping UA value	0.007	0.259	0.344	0.002	0.001	0.000	0.036	0.082	0.082	1.000	1.000	1.000
Condensing temp minimum, delta	0.003	0.002	0.135	NA	NA	NA	0.001	0.000	0.000	0.467	0.087	0.774
Fan Efficiency (%)	0.005	0.043	0.129	0.006	0.096	0.080	0.555	0.552	0.545	0.001	0.000	0.000
DX Cool Coil COP	0.011	0.019	0.096	NA	NA	NA	0.375	0.089	0.278	0.000	NA	0.000
Anti-sweat energy multiplier	0.000	NA	NA	NA	NA	NA	0.000	0.000	0.000	0.000	NA	0.000
T-Stat float heat/cool (oC)	0.034	0.117	0.208	0.004	0.076	0.118	0.243	0.180	0.270	0.057	0.067	0.062
Operating walkin ref temp (oC)	0.002	0.007	0.011	NA	NA	NA	0.022	0.009	0.010	0.139	0.056	0.079
Cooling Setpoint (oC)	0.010	NA	0.074	0.000	0.000	0.001	0.053	0.015	0.234	0.000	0.000	0.011

11.7 Appendix G: LHS Sampling Correlation Plots

The following plots were generated from the case-study building model annual simulation sampling. Total energy is shown, on the X-axis, as correlated to various input parameters, described on the Y-axis of each plot.

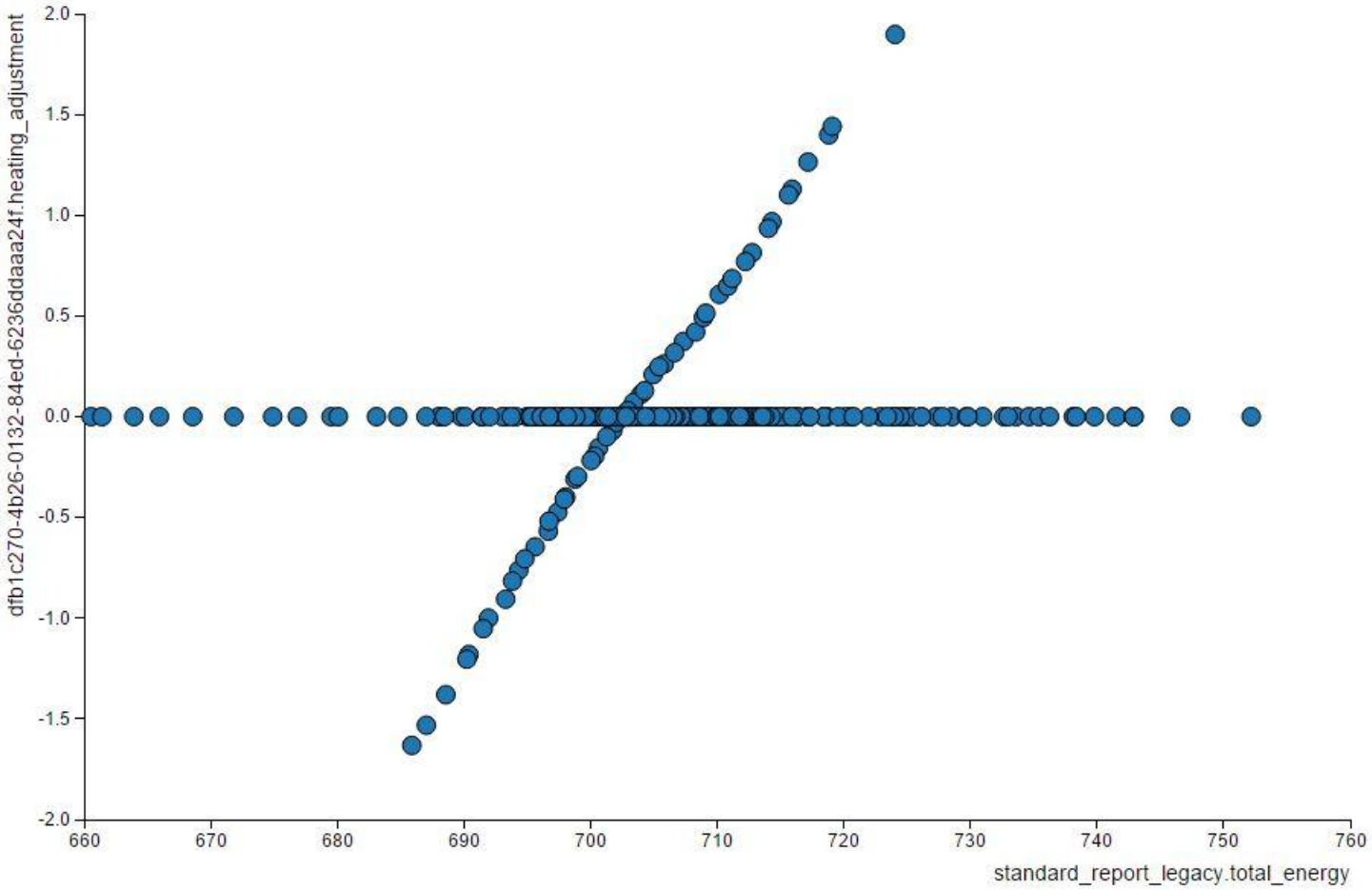
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



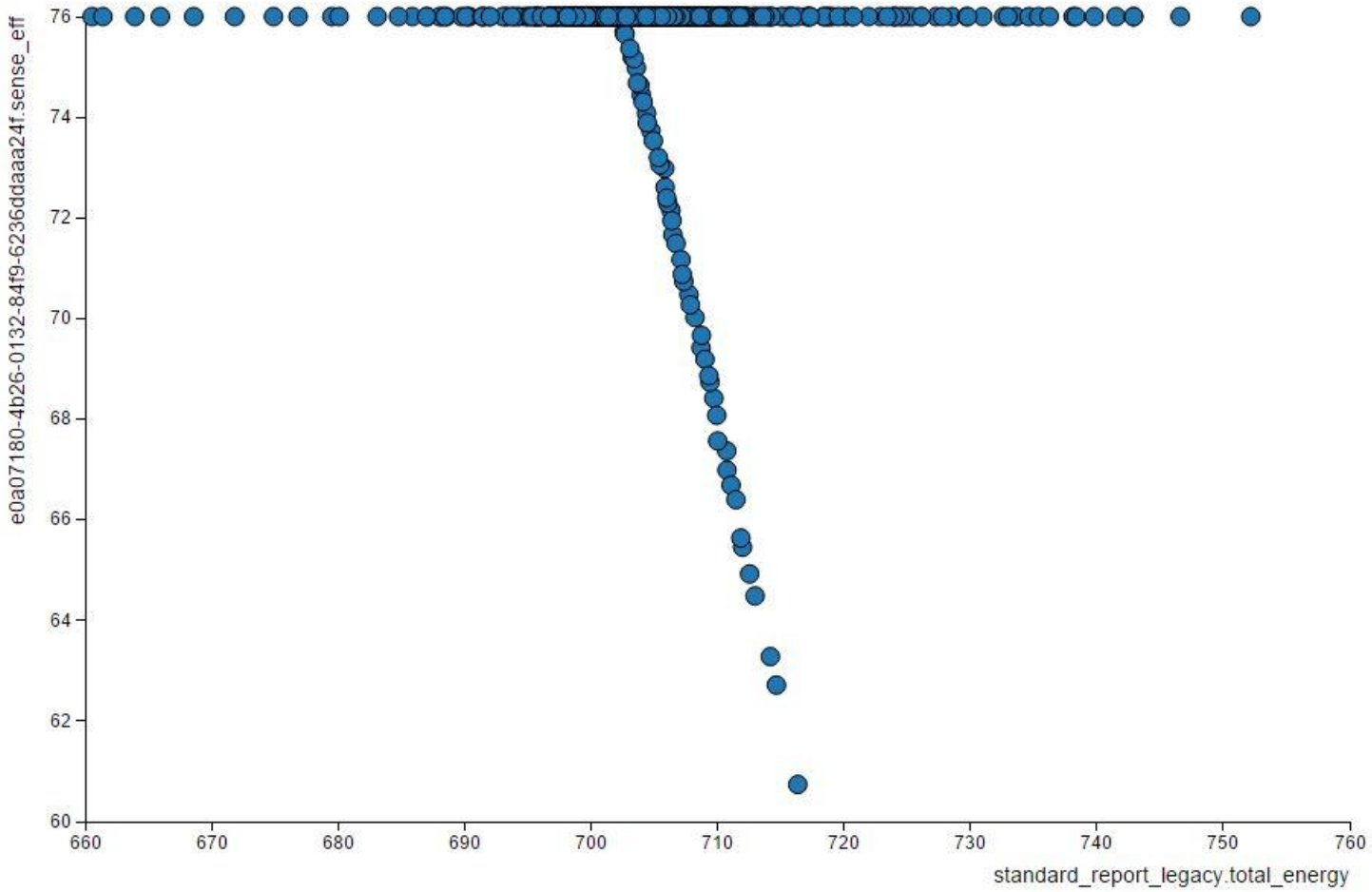
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



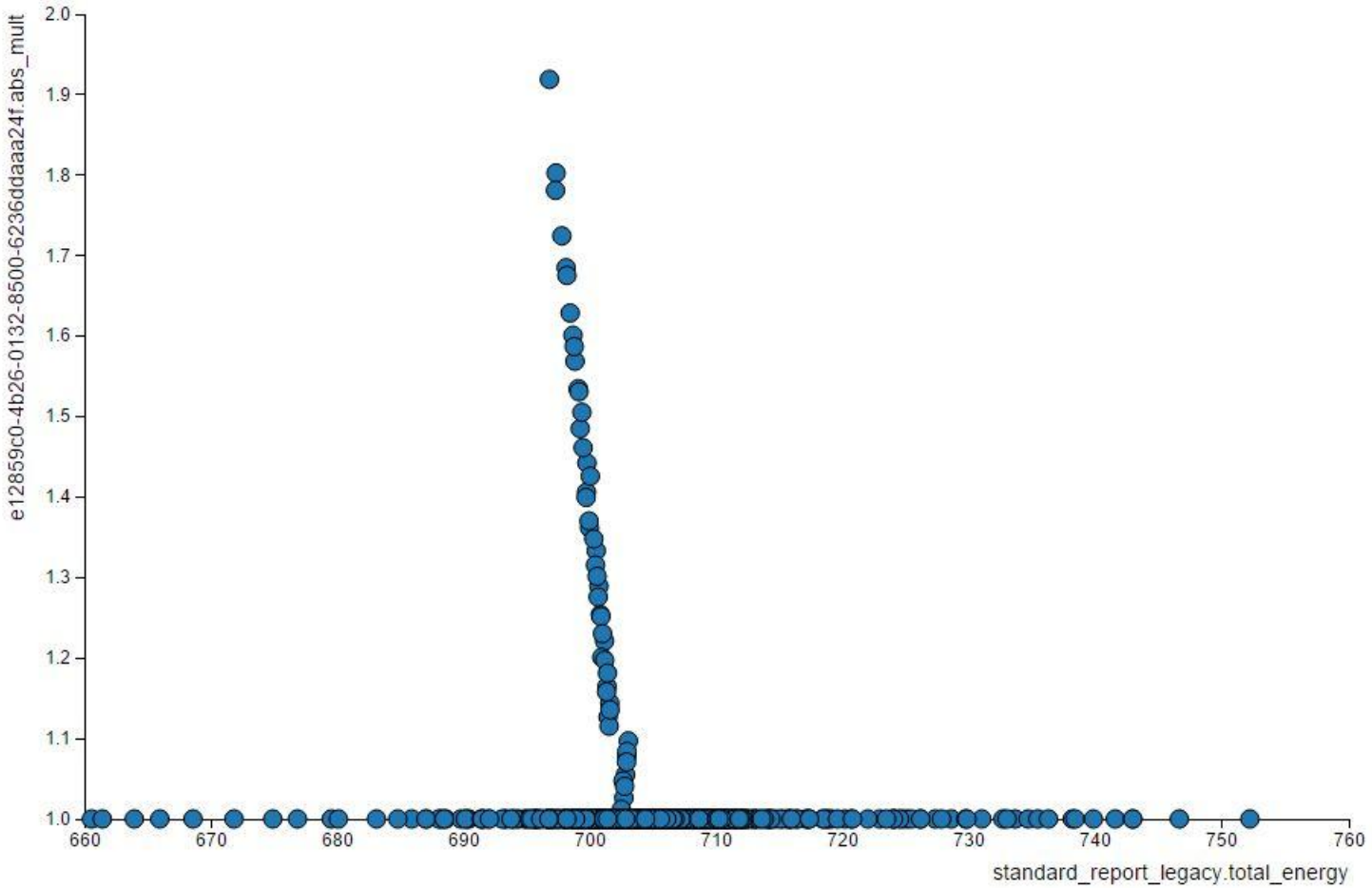
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



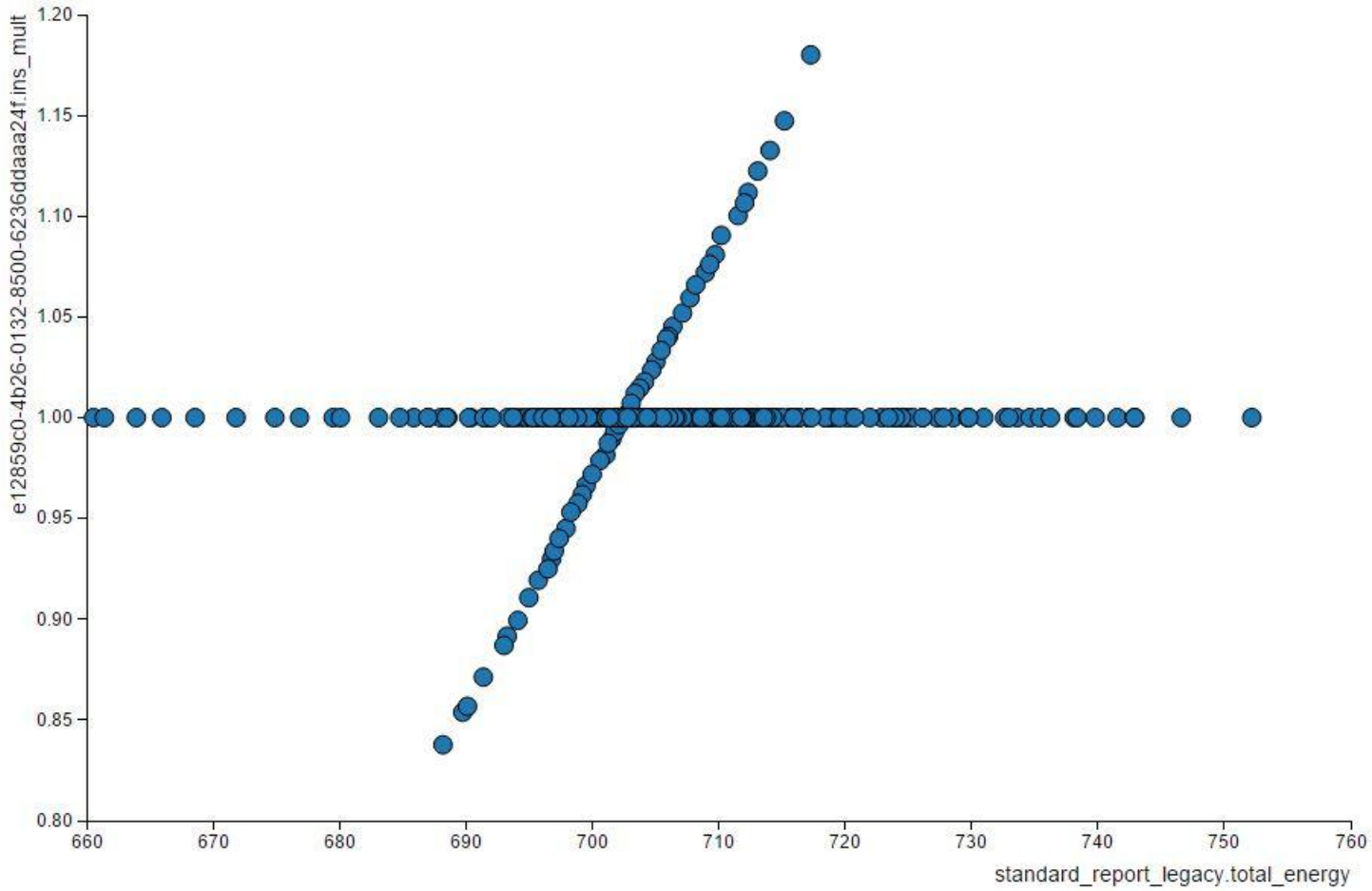
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



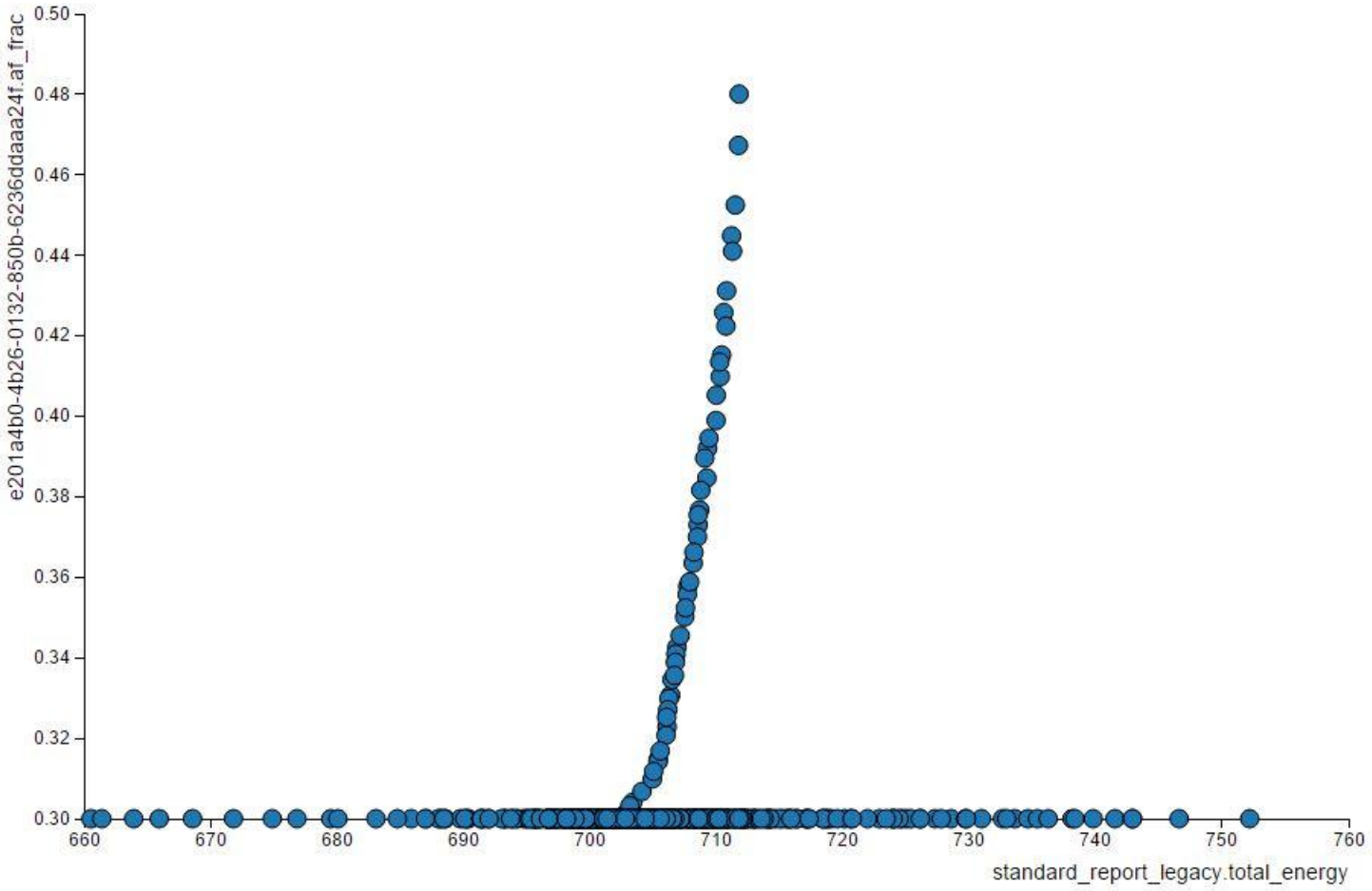
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



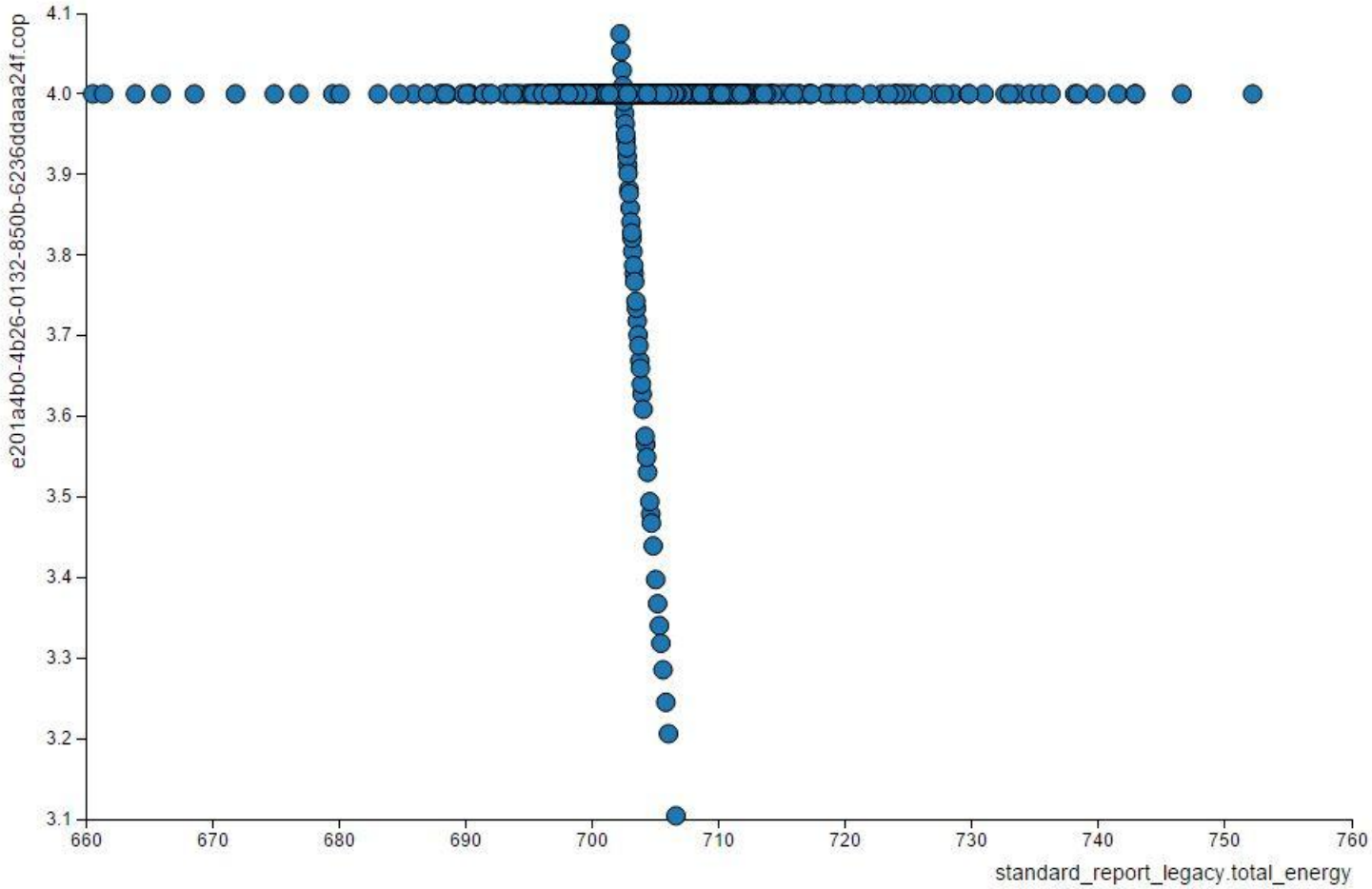
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



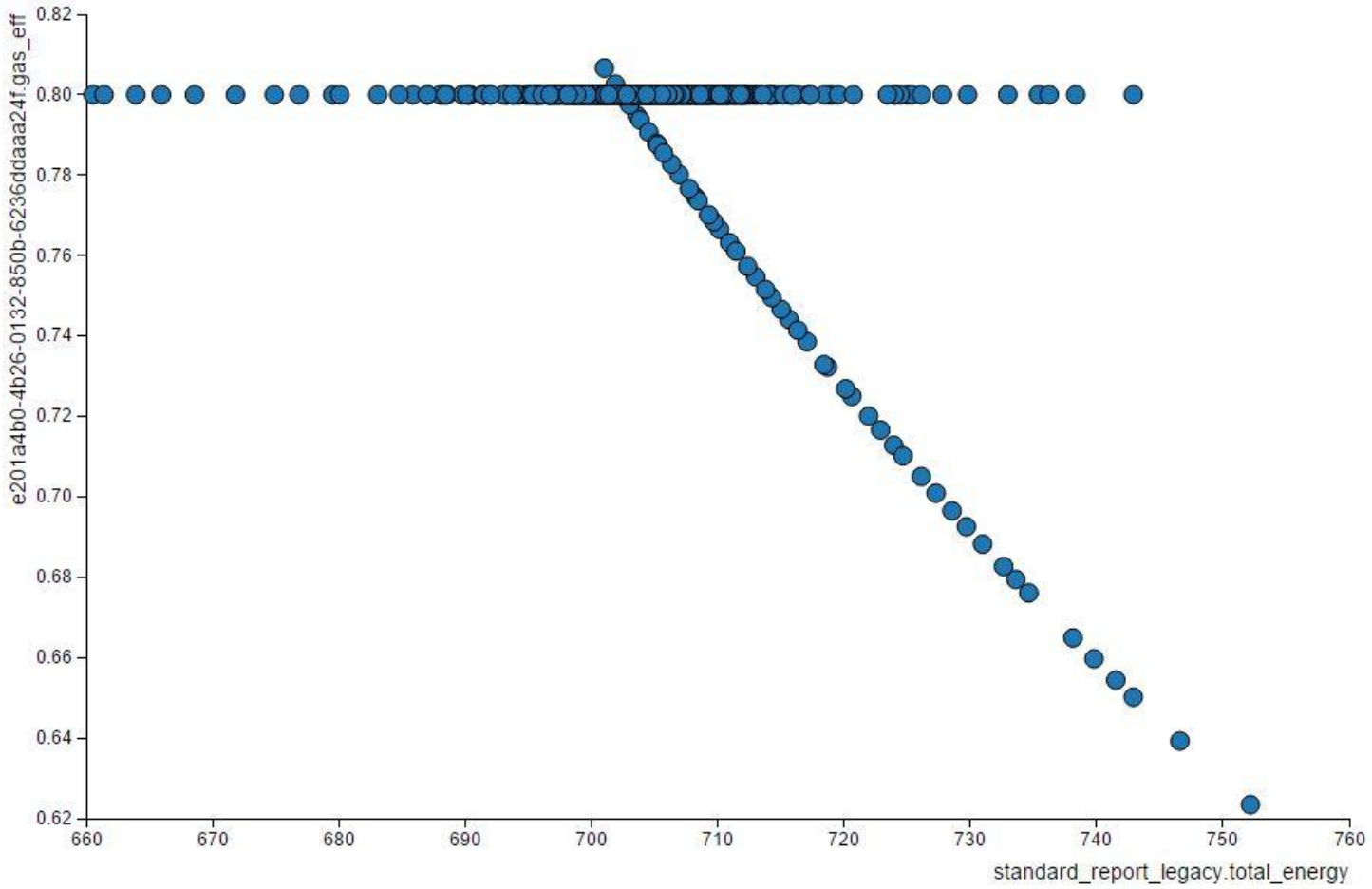
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



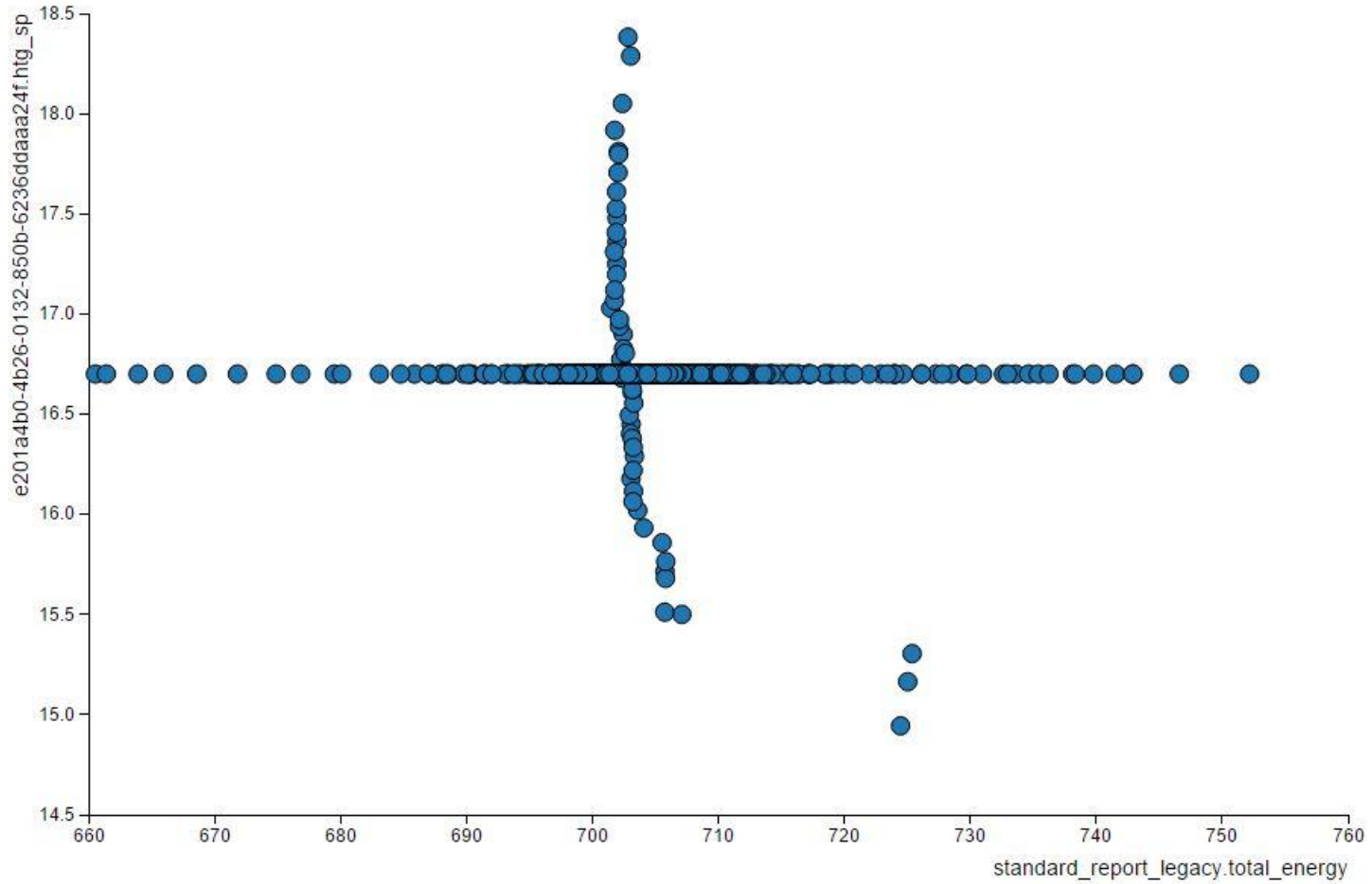
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



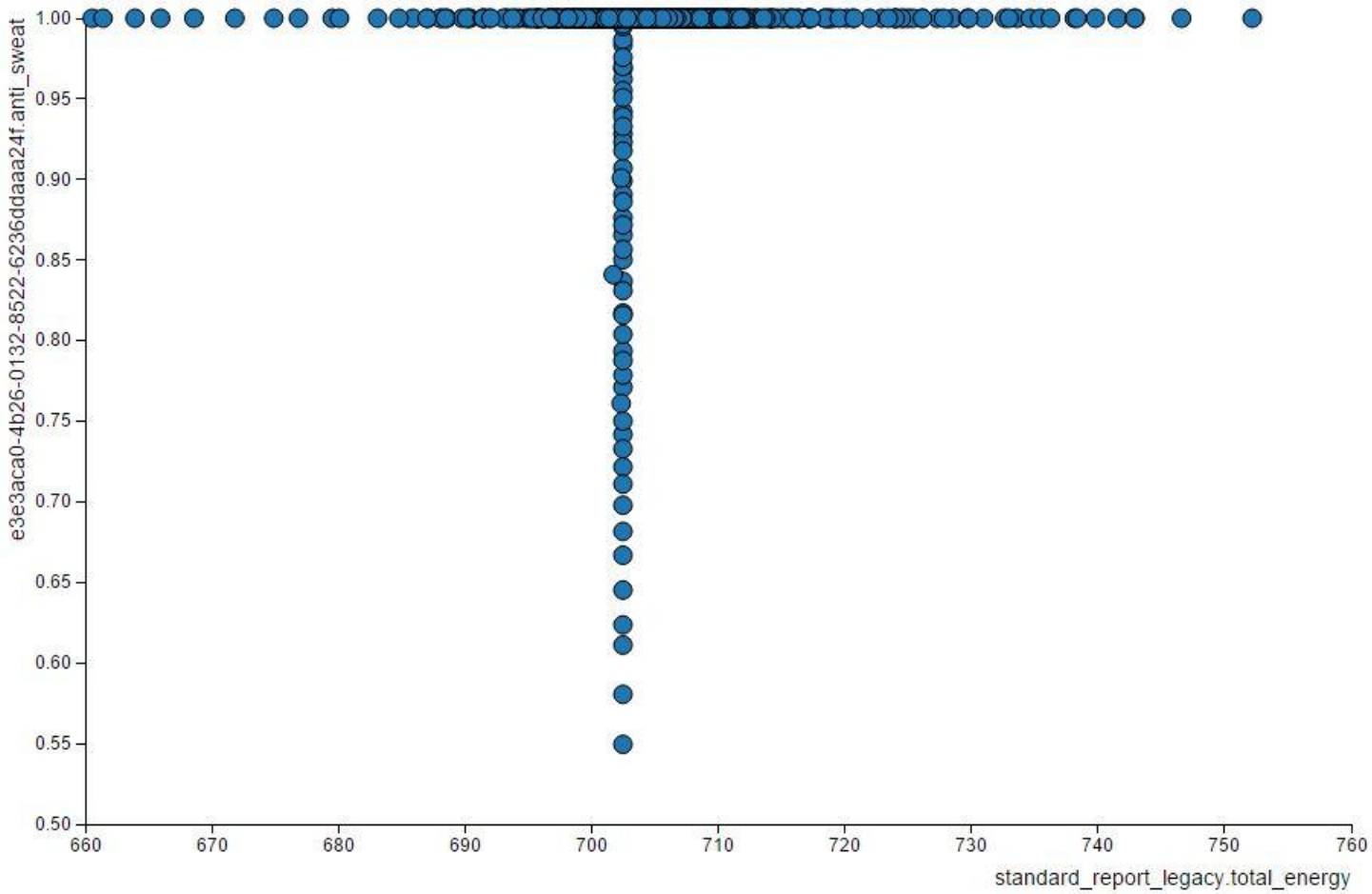
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



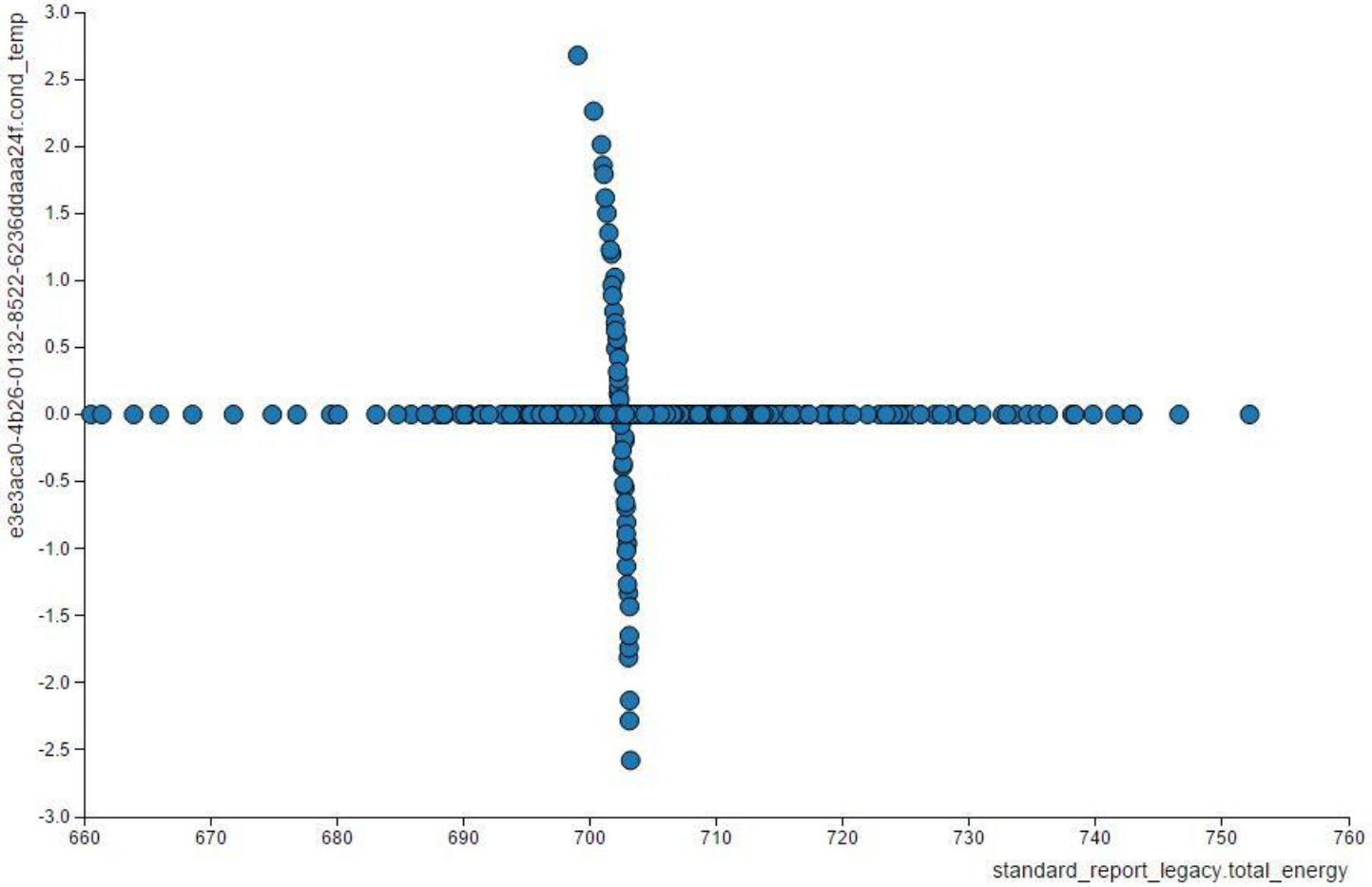
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



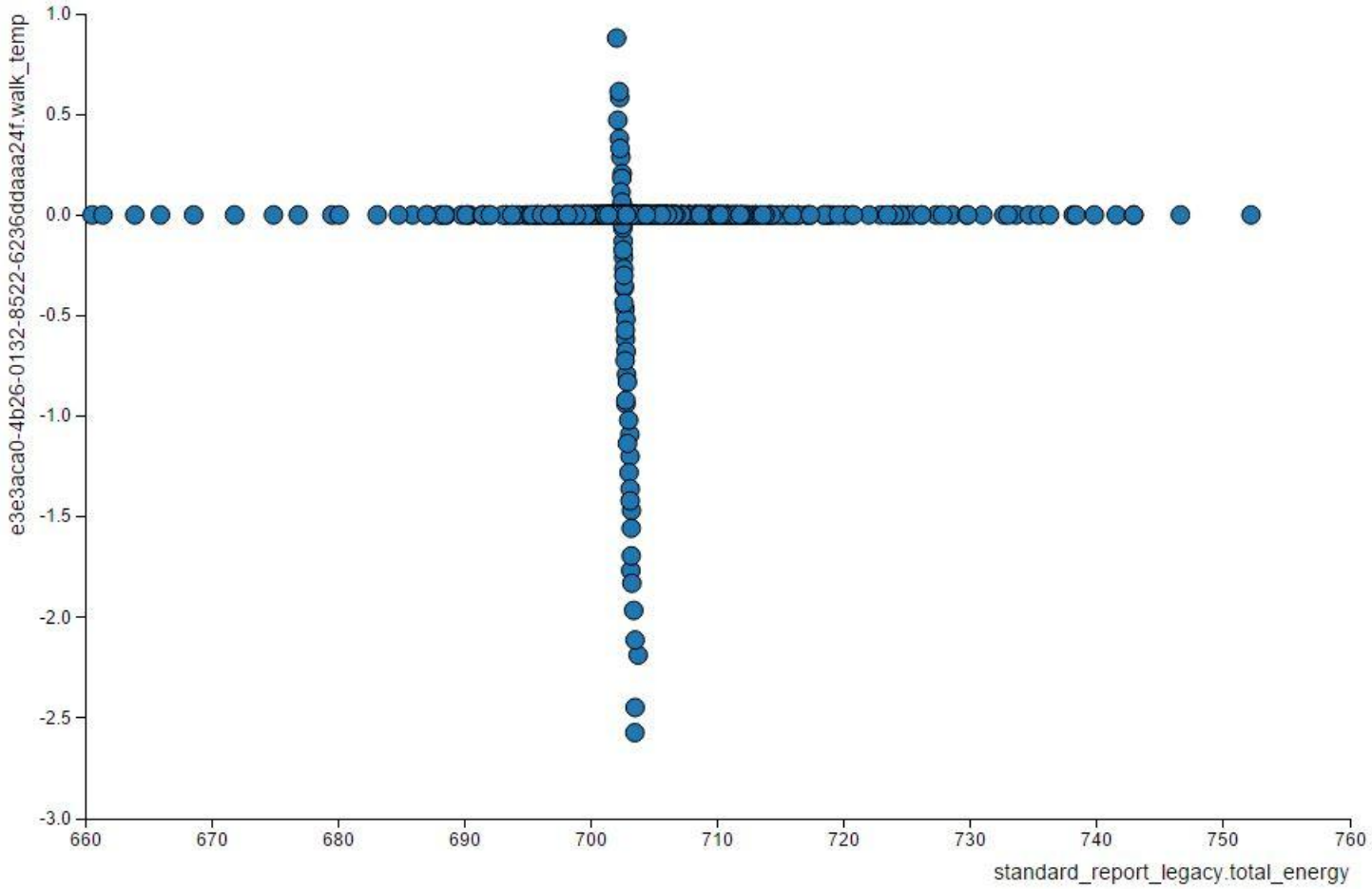
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



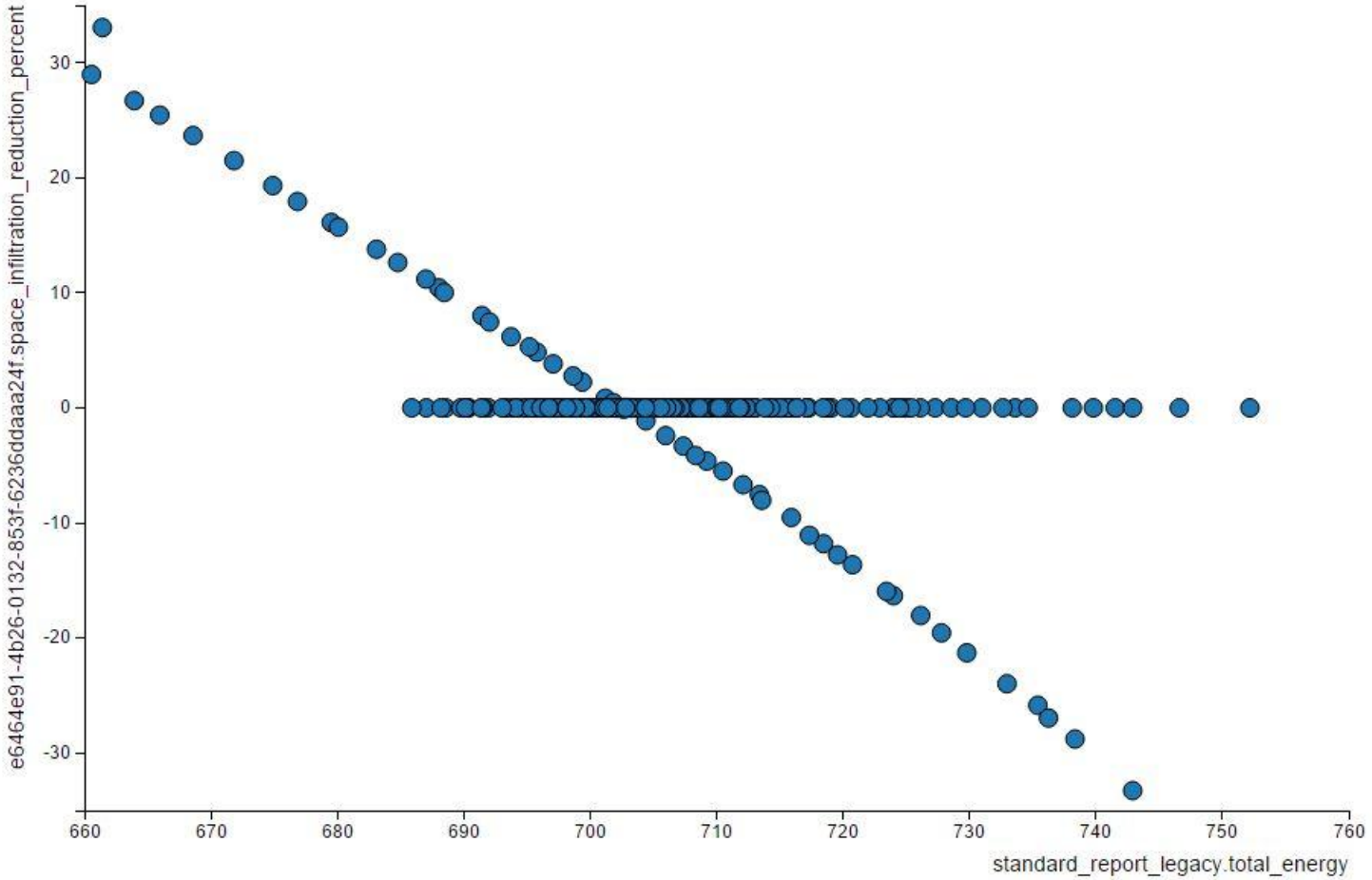
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



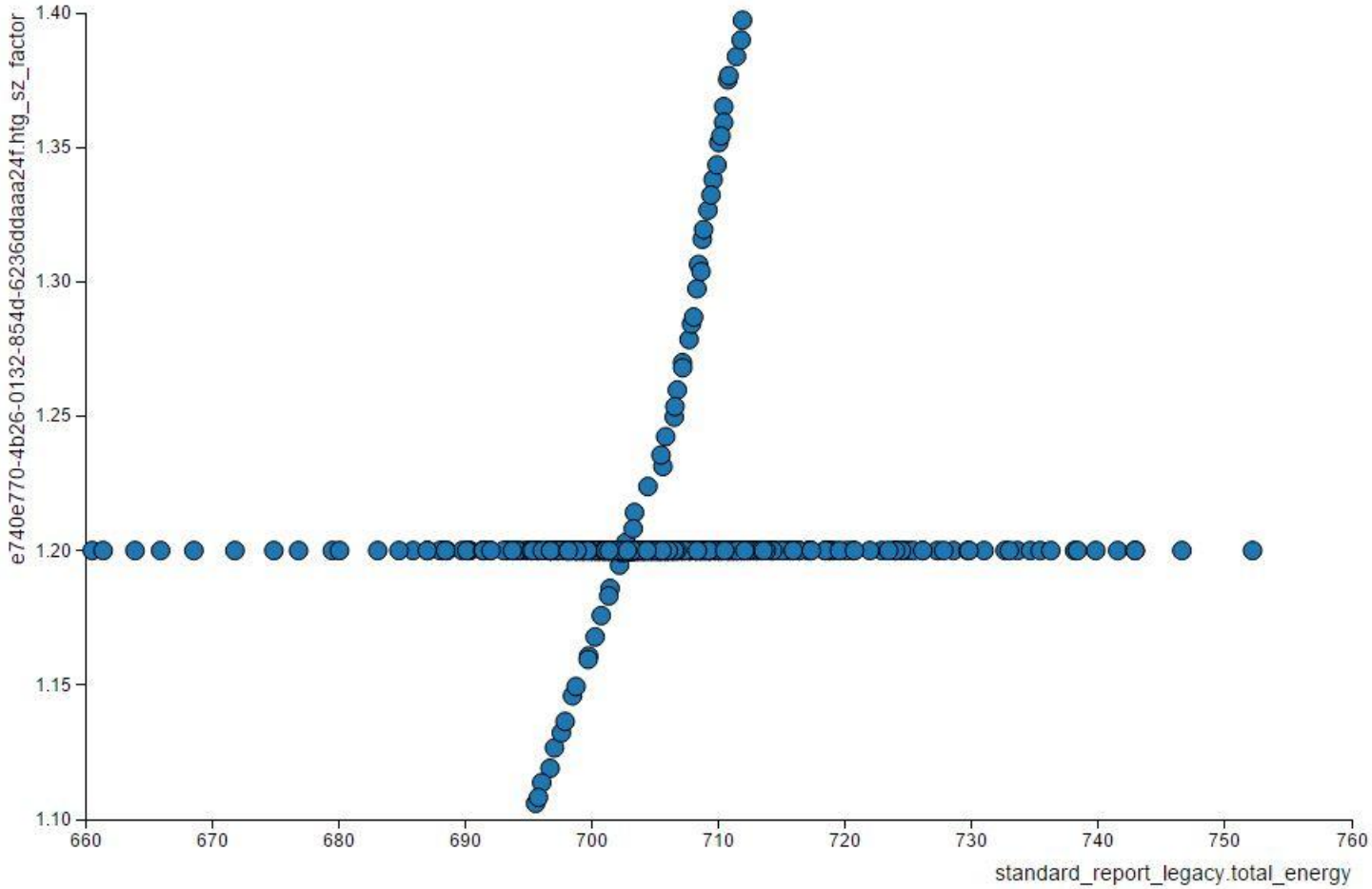
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



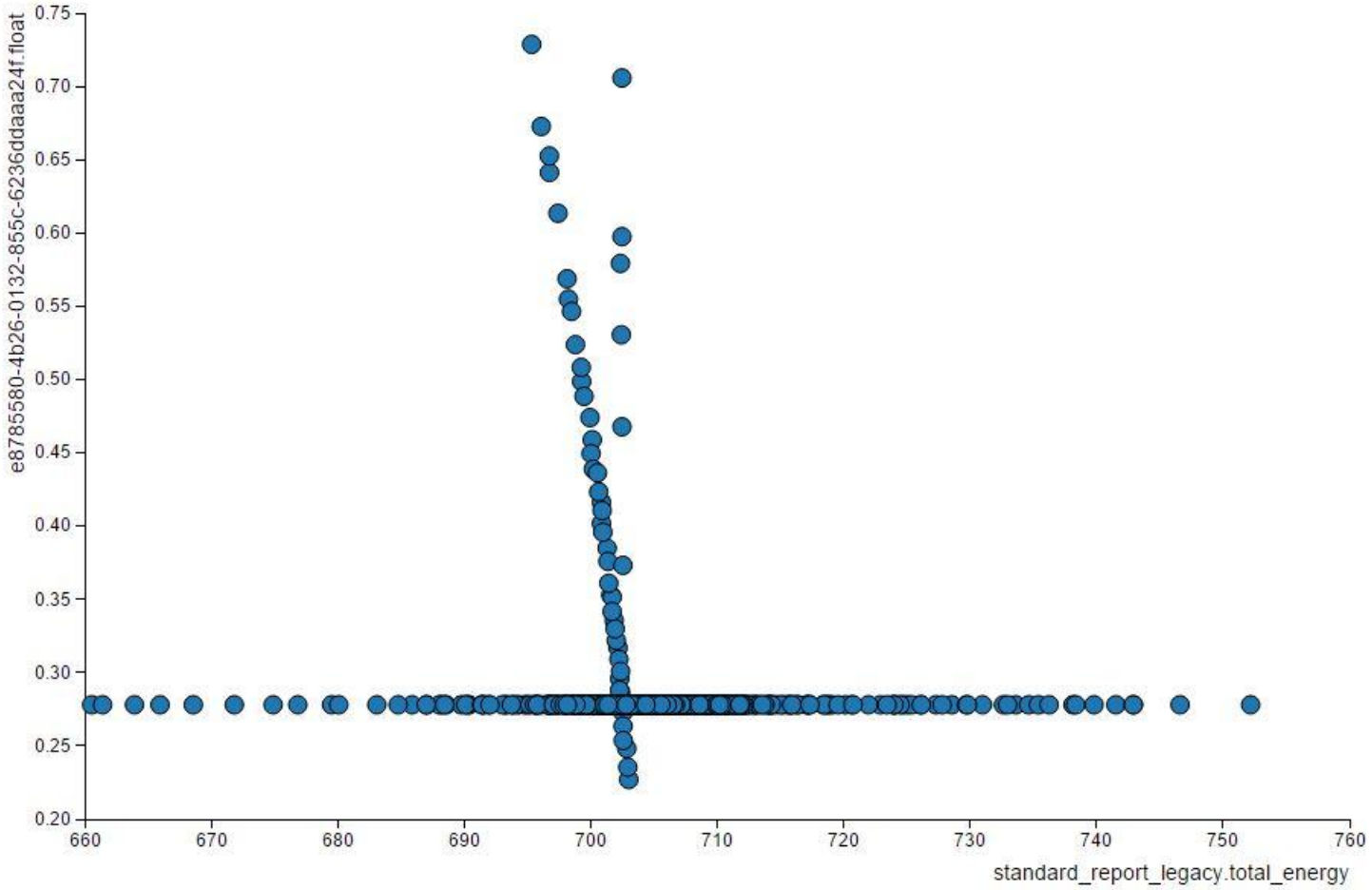
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



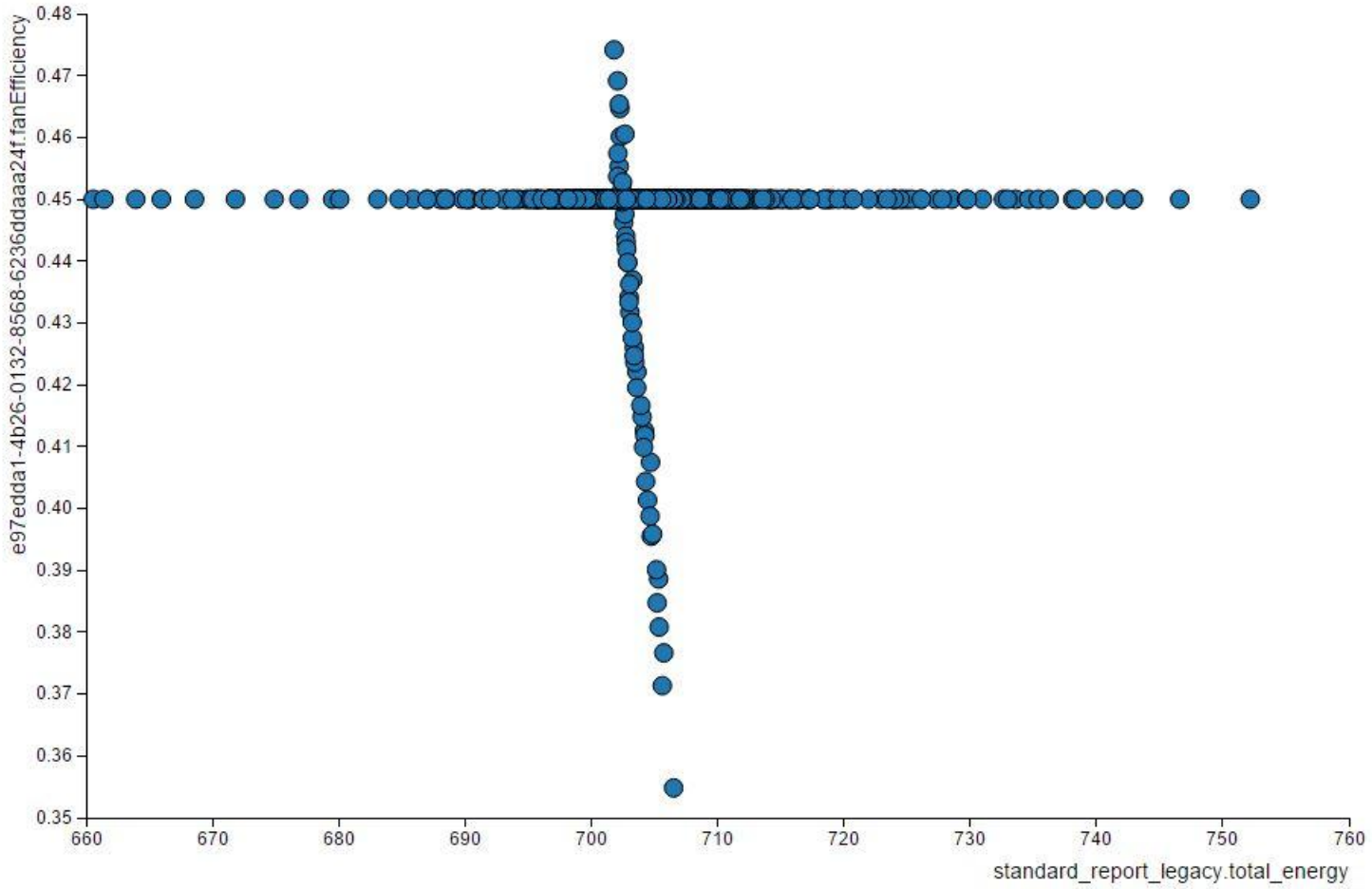
OpenStudio Cloud Management Console

Analysis Results — target lhs_test



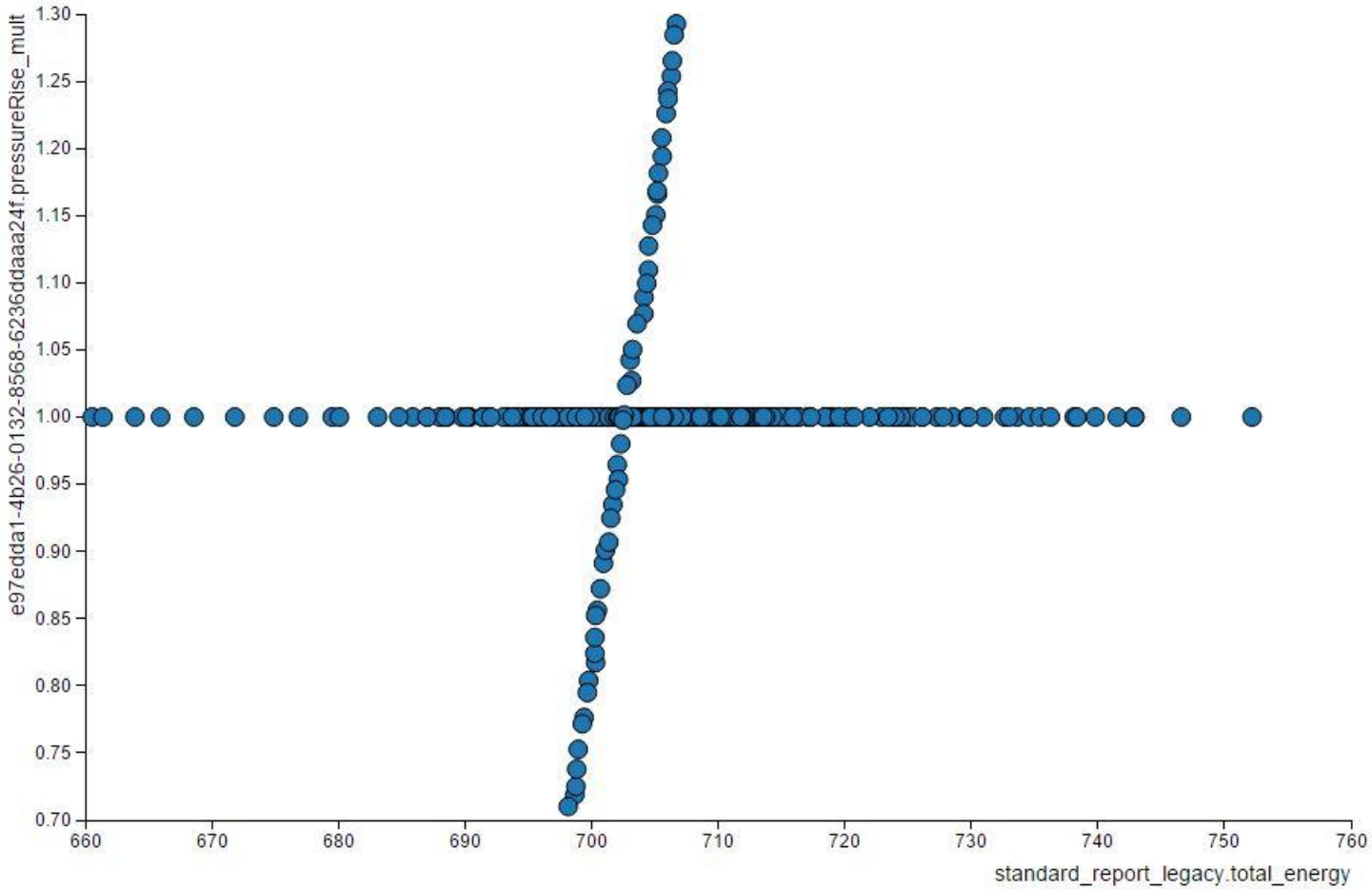
OpenStudio Cloud Management Console

Analysis Results — target lhs_test

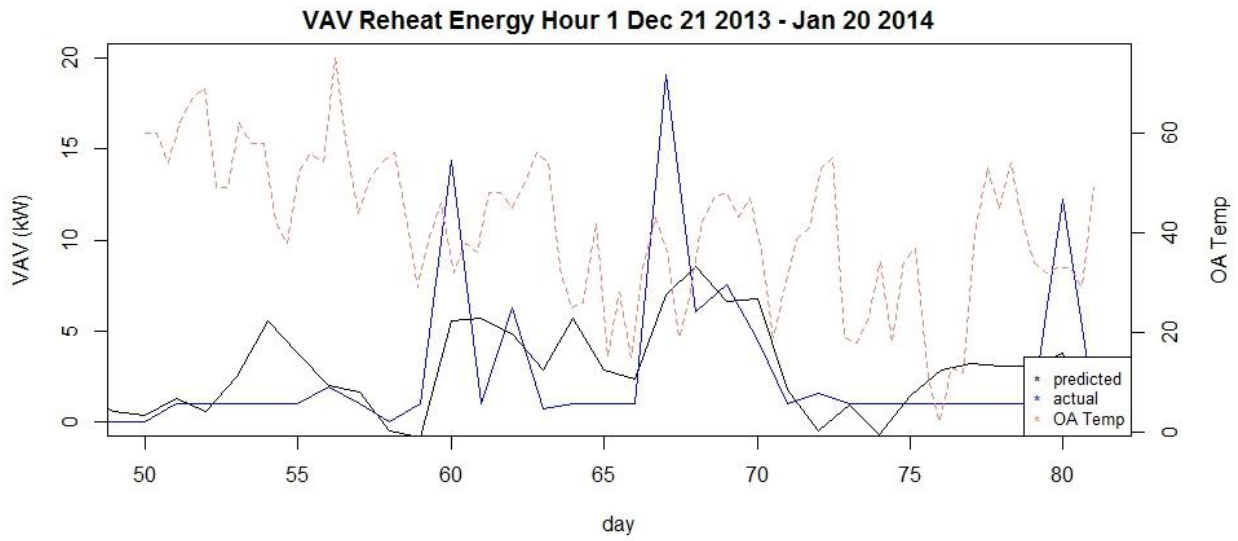


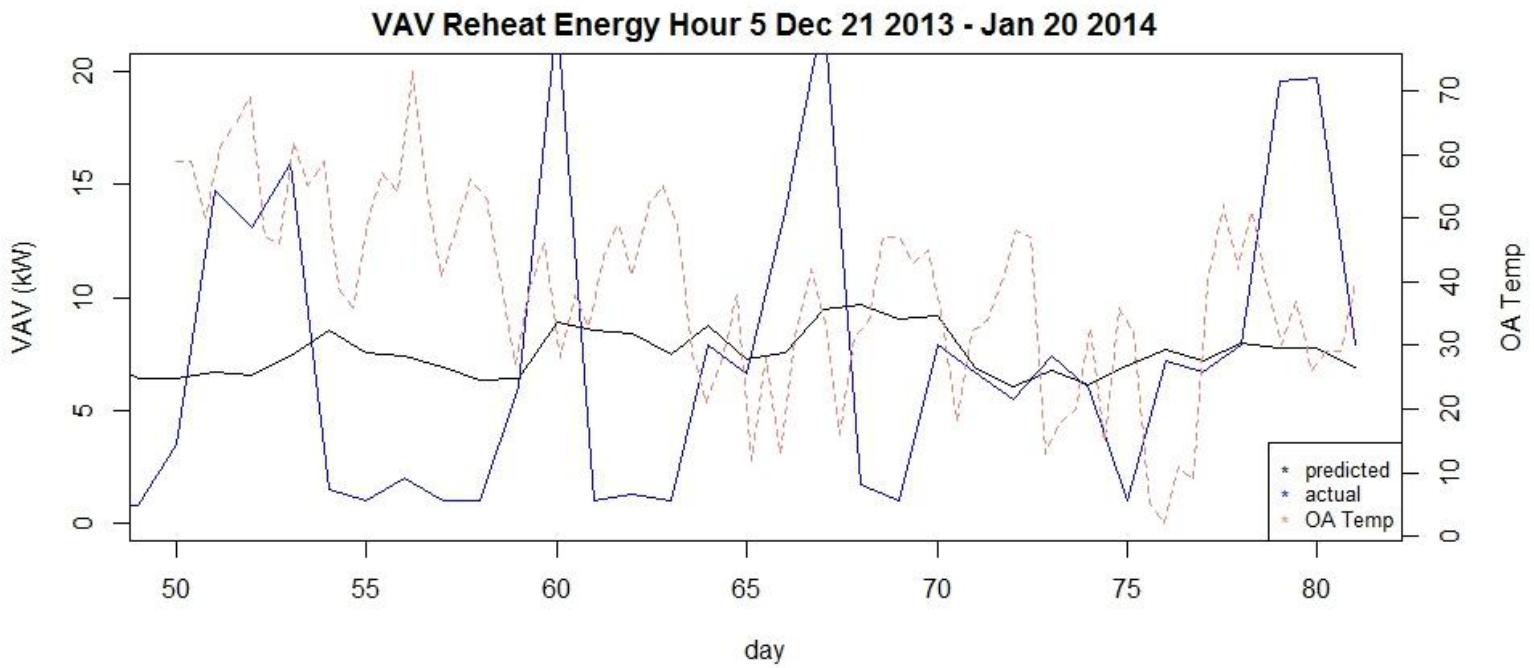
OpenStudio Cloud Management Console

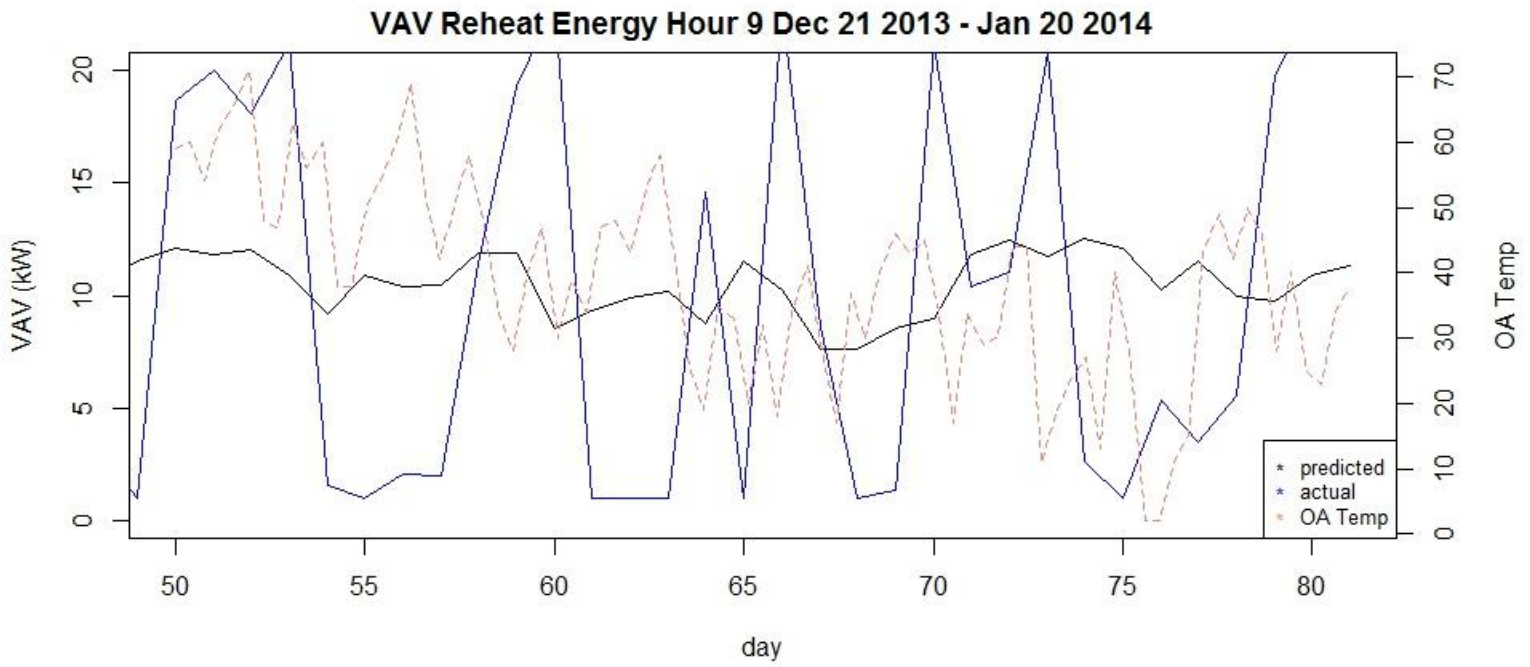
Analysis Results — target lhs_test

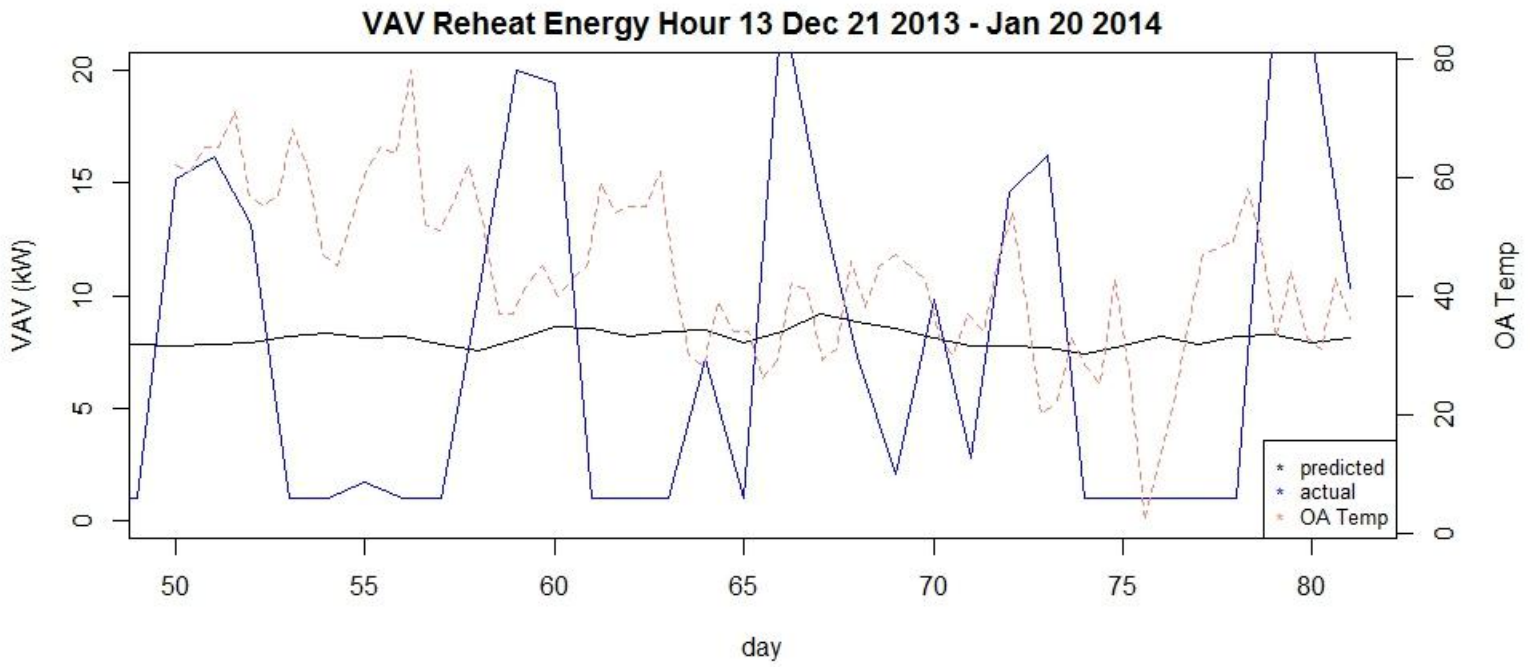


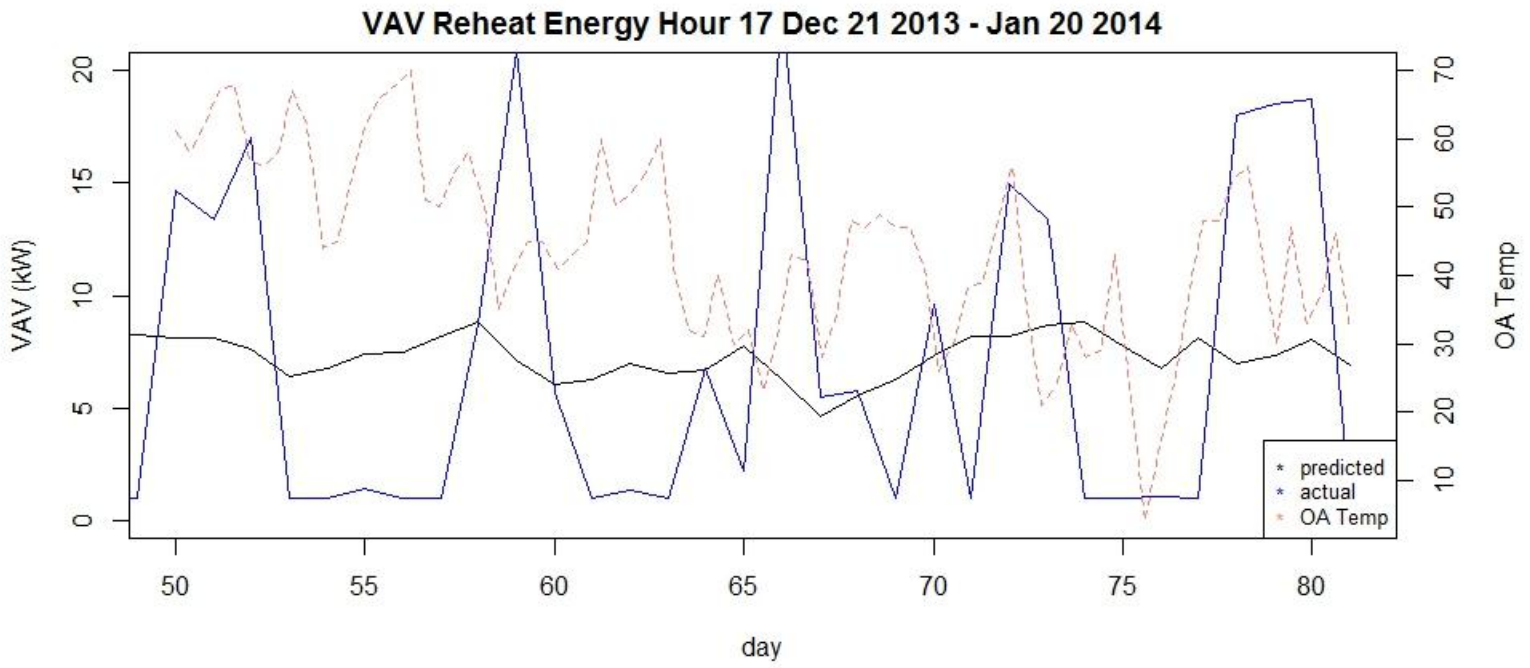
11.8 Appendix H: VAV Data Synthesis

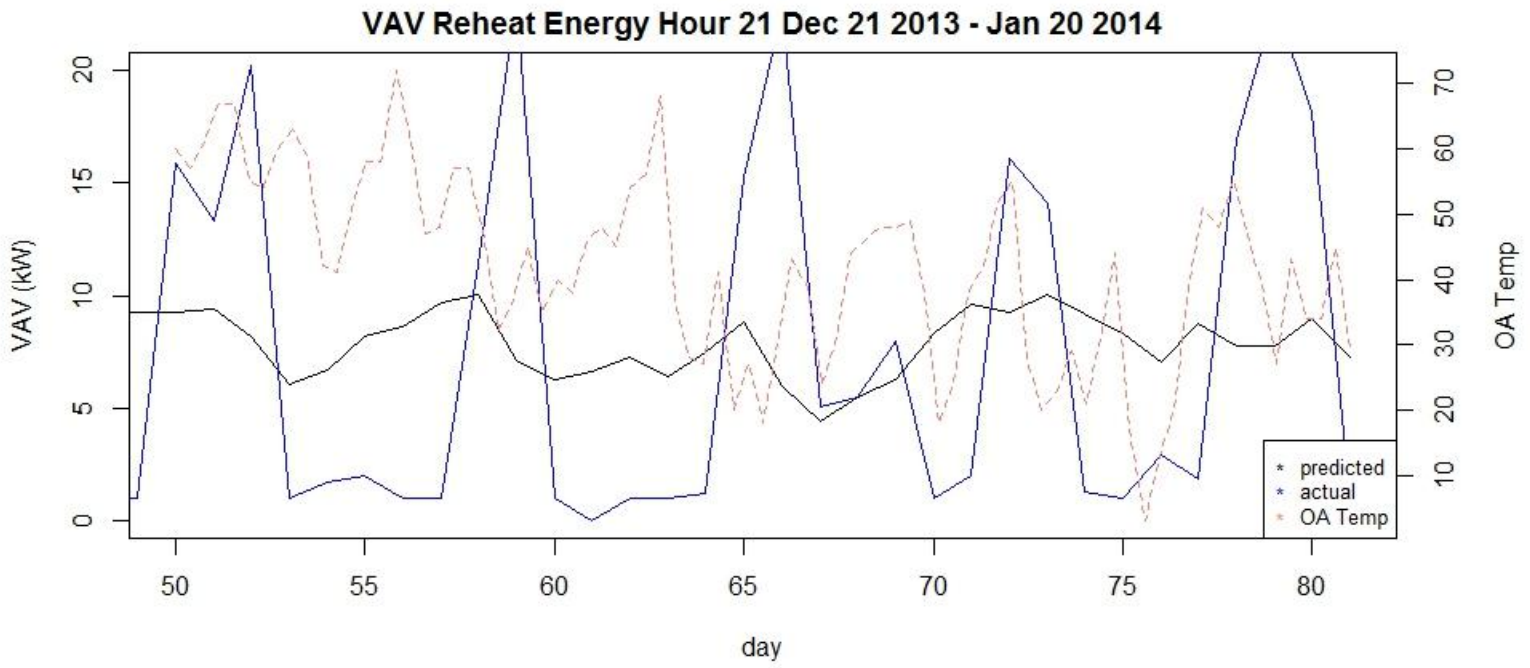



VAV Reheat Energy Hour 5 Dec 21 2013 - Jan 20 2014
VAV (kW)
OA Temp
day
predicted
actual
OA Temp
0
5
10
15
20
0
10
20
30
40
50
60
70
50
55
60
65
70
75
80


VAV Reheat Energy Hour 9 Dec 21 2013 - Jan 20 2014
VAV (kW)
OA Temp
day
predicted
actual
OA Temp
0
5
10
15
20
0
10
20
30
40
50
60
70
50
55
60
65
70
75
80


VAV Reheat Energy Hour 13 Dec 21 2013 - Jan 20 2014
VAV (kW)
OA Temp
day
predicted
actual
OA Temp
0
5
10
15
20
0
20
40
60
80
50
55
60
65
70
75
80


VAV Reheat Energy Hour 17 Dec 21 2013 - Jan 20 2014
VAV (kW)
OA Temp
day
predicted
actual
OA Temp
0
5
10
15
20
10
20
30
40
50
60
70
50
55
60
65
70
75
80


VAV Reheat Energy Hour 21 Dec 21 2013 - Jan 20 2014
VAV (kW)
OA Temp
day
predicted
actual
OA Temp
0
5
10
15
20
10
20
30
40
50
60
70
50
55
60
65
70
75
80

11.9 Appendix I: Cost Matrix Detailed Explainations

The cost function is the input going into the signal tool that captures the response that the facility manager would recommend for each unique scenario. For example, natural gas used is observed as much higher than expected – since this a significant operational expenditure, immediate action to address high gas use would be recommended. A unique cost matrix has been developed for each end use, capturing the utilization, maintenance cost and relative importance

When assigning costs to signals for a particular end-use, the user can consider:

1. Does the end use have relatively high cost to repair compared to its cost to operate?
    a. If so, then the magnitude of cost for showing any false "red" signal should be penalized more, since a red signal will place this system high on the list of action items. That is, action should only be taken if probability of faults is quite high.

    Examples: Cooling systems in a cool climate, heating systems in a warm climate

    b. If not, then the magnitude of the cost for showing a false low signal is penalized more; and a false high signal penalized less, since maintenance is inexpensive relative to the potential cost of operating at more than expected energy use. Also, the cost for showing a false green signal with high consumption is penalized more; since doing nothing may be nearly as bad as taking the wrong action.
2. Would an observed reading that is higher (or lower) than expected correspond to a problem in the building that has indirect costs associated with customer comfort or product stability?
    a. If higher energy use observation corresponds to problem, then penalize false low signals more heavily.

    Examples:

    b. If a lower energy use observation corresponds to problem, then penalize false high signals more heavily.

    Examples: Grocery freezer and refrigeration systems, cooling systems in a warm climate, ventilation systems

Secondary Cost Matrices:

NEUTRAL

	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	6	8	20
YH	2	0	2	4	6	14
G	4	2	0	2	4	12
YL	6	4	2	0	2	14
RL	8	6	4	2	0	20
Σ State Cost:	**20**	**14**	**12**	**14**	**20**	

Figure 101: The neutral cost matrix

Q 1: Cost ratio maintain/operate?

1- low	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	5	6	17
YH	2	0	2	4	5	13
G	4	2	0	2	4	12
YL	5	4	2	0	2	13
RL	6	5	4	2	0	17
Σ State Cost:	**17**	**13**	**12**	**13**	**17**	

1 - high	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	7	12	25
YH	2	0	2	5	7	16
G	4	2	0	2	4	12
YL	7	5	2	0	2	16
RL	12	7	4	2	0	25
Σ State Cost:	**25**	**16**	**12**	**16**	**25**	

Q 2: Does high or low consumption indicate a serious non-energy fault?

2 - low	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	7	10	23
YH	2	0	2	4	7	15
G	3.5	2	0	2.5	4.5	12.5
YL	5	4	2	0	2.5	13.5
RL	6	5	4	2	0	17
Σ State Cost:	**16.5**	**13**	**12**	**15.5**	**24**	

2 - high	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	5	6	17
YH	2	0	2	4	5	13
G	4	2	0	2	4	12
YL	6	4	2	0	2	14
RL	8	6	4	2	0	20
Σ State Cost:	**20**	**14**	**12**	**13**	**17**	

Figure 102: The secondary cost matrices

The final cost matrix is found via a calculation taking into account the answers to the two questions by the following formula, which gives equal weighting to the answers from both questions:

$$C_{Final} = \frac{C_1 + C_2}{2}$$

Where, C_1 and C_2 are the respective cost matrices associated with the answers to questions one and two, as posed above. If either of those questions is answered with "N/A", then C_{Final} will be calculated with the neutral cost matrix substituted for the cost matrix corresponding to this question. If the answer to both questions is "N/A", then C_{Final} will simply be the neutral cost matrix. Cost matrices corresponding to these answers are found below. In some cases, the user may wish to modify these to reflect better information about energy management priorities.

The Eight Possible Final Cost Matrices:

2

N/A, Low	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	6.5	9	21.5
YH	2	0	2	4	6.5	14.5
G	3.75	2	0	2.25	4.25	12.25
YL	5.5	4	2	0	2.25	13.75
RL	7	5.5	4	2	0	18.5
Σ State Cost:	18.25	13.5	12	14.75	22	

3

N/A, High	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	5.5	7	18.5
YH	2	0	2	4	5.5	13.5
G	4	2	0	2	4	12
YL	6	4	2	0	2	14
RL	8	6	4	2	0	20
Σ State Cost:	20	14	12	13.5	18.5	

4

Low, N/A	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	5.5	7	18.5
YH	2	0	2	4	5.5	13.5
G	4	2	0	2	4	12
YL	5.5	4	2	0	2	13.5
RL	7	5.5	4	2	0	18.5
Σ State Cost:	18.5	13.5	12	13.5	18.5	

5

Low, Low	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	6	8	20
YH	2	0	2	4	6	14
G	3.75	2	0	2.25	4.25	12.25
YL	5	4	2	0	2.25	13.25
RL	6	5	4	2	0	17
Σ State Cost:	16.75	13	12	14.25	20.5	

6

Low, High	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	5	6	17
YH	2	0	2	4	5	13
G	4	2	0	2	4	12
YL	5.5	4	2	0	2	13.5
RL	7	5.5	4	2	0	18.5
Σ State Cost:	18.5	13.5	12	13	17	

7

High, Neut	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	6.5	10	22.5
YH	2	0	2	4.5	6.5	15
G	4	2	0	2	4	12
YL	6.5	4.5	2	0	2	15
RL	10	6.5	4	2	0	22.5
Σ State Cost:	22.5	15	12	15	22.5	

8

High, Low	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	7	11	24
YH	2	0	2	4.5	7	15.5
G	3.75	2	0	2.25	4.25	12.25
YL	6	4.5	2	0	2.25	14.75
RL	9	6	4	2	0	21
Σ State Cost:	20.75	14.5	12	15.75	24.5	

9

High, High	MH	SH	S	SL	ML	Σ Action Cost:
RH	0	2	4	6	9	21
YH	2	0	2	4.5	6	14.5
G	4	2	0	2	4	12
YL	6.5	4.5	2	0	2	15
RL	10	6.5	4	2	0	22.5
Σ State Cost:	22.5	15	12	14.5	21	

Table 34: All possible final cost matrices based on the answers to the two questions

When setting up the Energy Signal Tool for the case study building, answers were provided to the two cost matrix definition questions for each end use:

Whole building energy

1. **ratio $\frac{cost\ maint}{cost\ operate}$=Low**
   Justification: When considering the processes of the entire building, annual energy costs generally far outweigh annual maintenance costs.
2. **Is either high or low usage indicative of other problem? = N/A**
   Justification: Neither extreme is a definite indicator of a serious problem at the whole building level.

Natural Gas

1. **ratio $\frac{cost\ maint}{cost\ operate}$ = Low**
   Justification: Natural gas equipment is mainly gas burners for heating, which have low maintenance costs and high fuel costs.
2. **Is either high or low usage indicative of other problem? = N/A**
   Justification: Neither extreme is a definite indicator of a serious non-energy problem for natural gas use. Very low natural gas use that compromises heating would be indicated early on by comfort complaints.

HVAC

1. **ratio $\frac{cost\ maint}{cost\ operate}$ = N/A**
   Justification: The HVAC end use is comprised of cooling, fan, and some electric reheat energy. Fans function throughout the year, and have a relatively low cost to maintain. The cooling system, being in a cool climate of Wisconsin, has relatively little utility and high maintenance costs. The two effectively cancel out.
2. **Is either high or low usage indicative of other problem? = Low**
   Justification: Lower than expected HVAC energy usage would most likely be indicative of a fan failure and loss of ventilation air, which may go unnoticed by store occupants.

Refrigeration

1. **ratio $\frac{cost\ maint}{cost\ operate}$ = High**

Justification: The refrigeration end use is small relative to the sum of all other building systems in this retail store, and service calls require a high level of expertise. This answer might be different for a grocery store rather than big-box retail.

2. **Is either high or low usage indicative of other problem? = Low**
   Justification: Lower than expected Refrigeration energy use could indicate compressor failure. This would also probably be noticed in refrigeration case temperature monitoring.

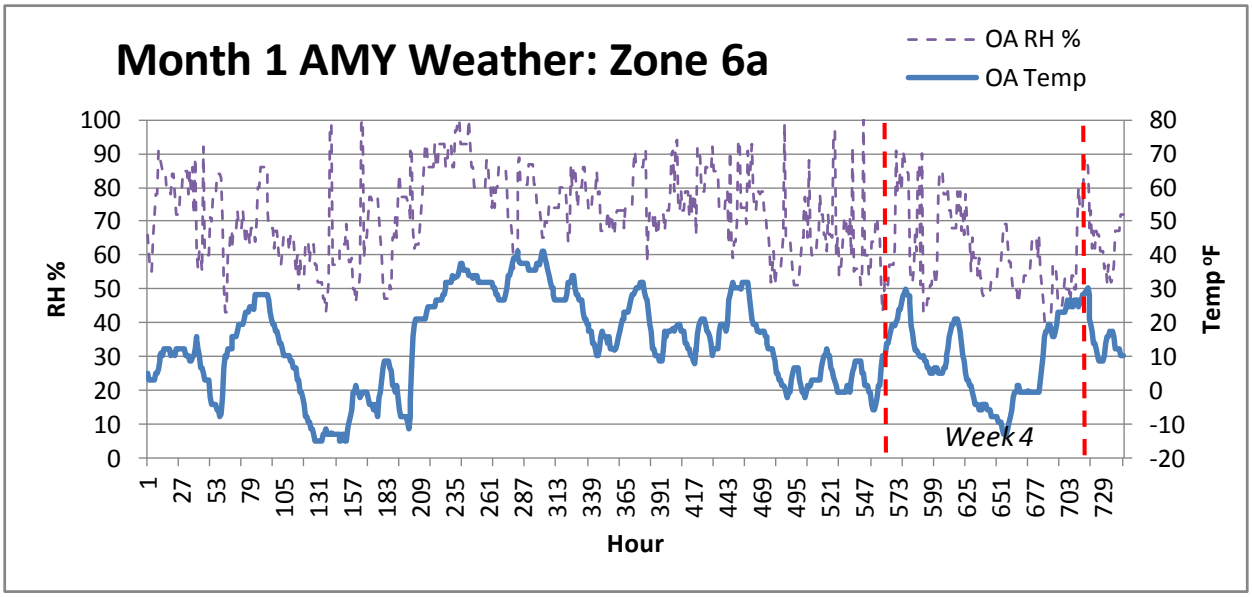
11.10 Appendix J: Weather Data Summary for Fault Testing Locations

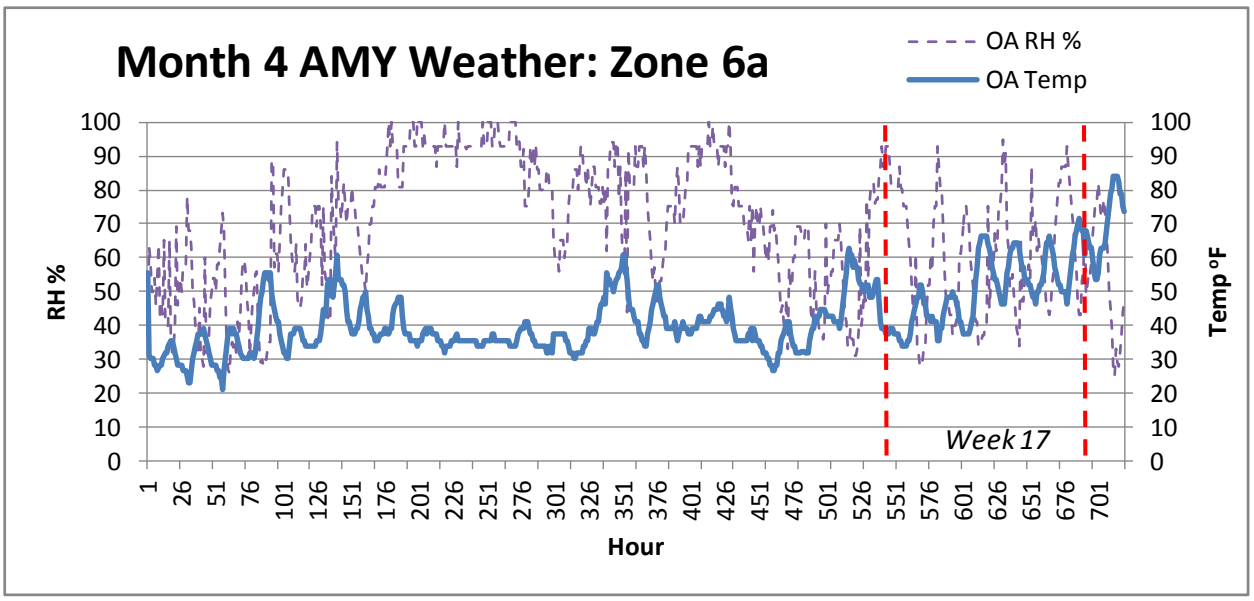
WHOLE BUILDING ENERGY	Month 1	Month 4	Month 7	Month 10	Year
Climate 3a	0.107	0.107	0.107	0.107	0.107
Climate 4c	0.104	0.104	0.104	0.104	0.104
Climate 6a	0.106	0.106	0.106	0.106	0.106

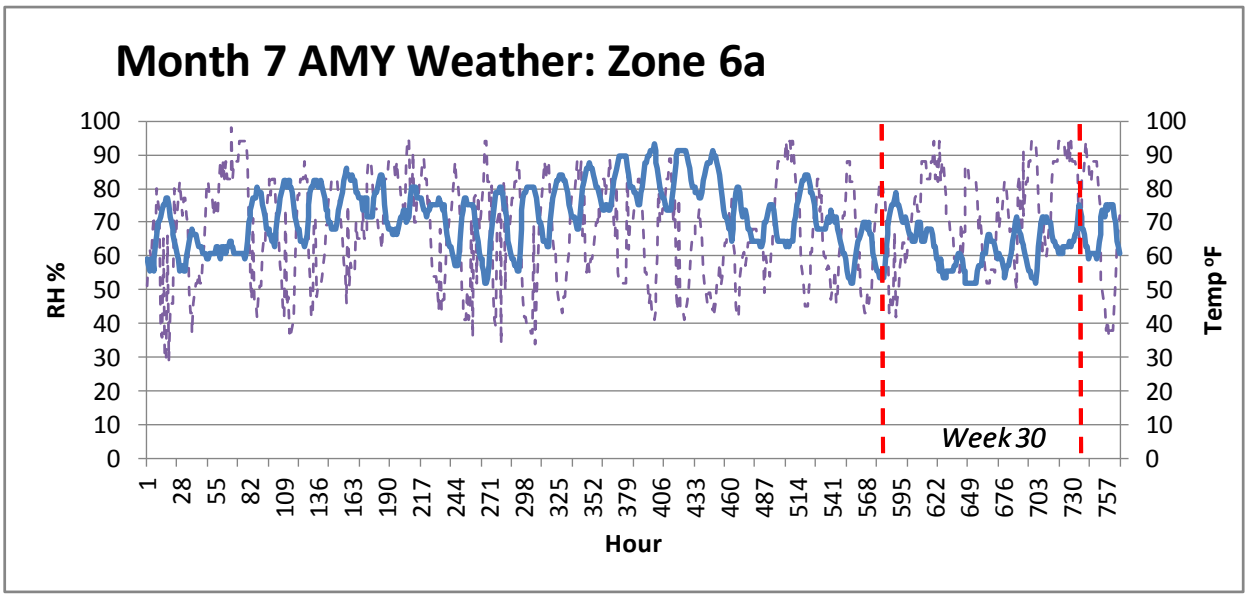
ATURAL GAS	Month 1	Month 4	Month 7	Month 10	Year
Climate 3a	0.099	0.118	0.158	0.127	0.099
Climate 4c	0.101	0.121	0.161	0.130	0.101
Climate 6a	0.106	0.106	0.127	0.169	0.137

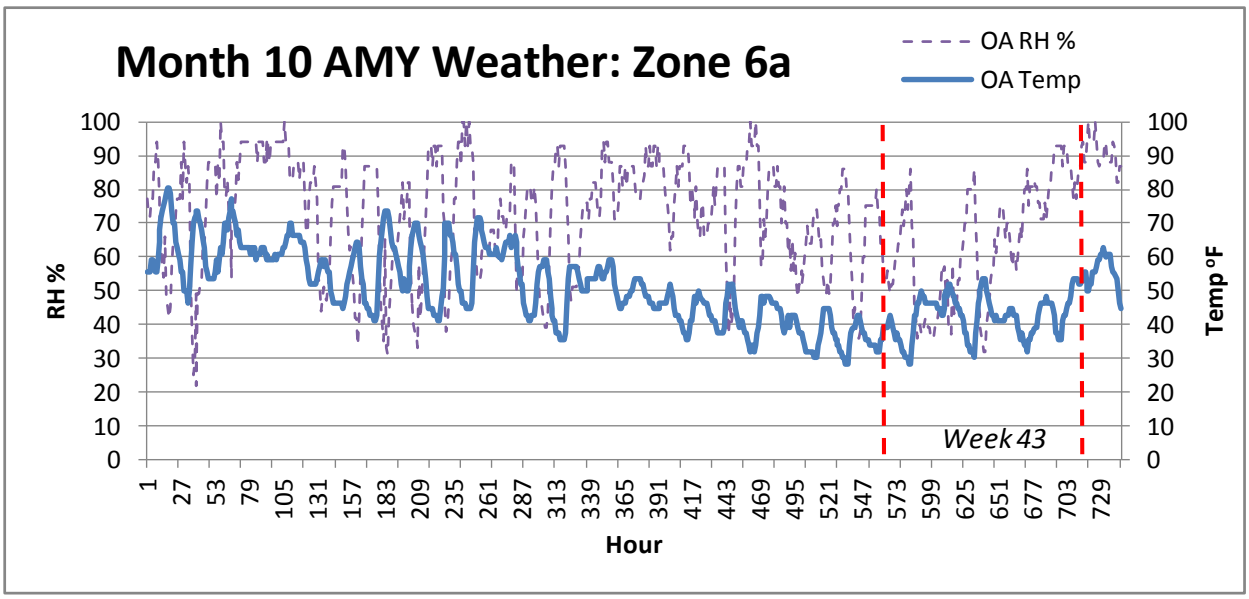
HVAC	Month 1	Month 4	Month 7	Month 10	Year
Climate 3a	0.139	0.150	0.122	0.144	0.122
Climate 4c	0.157	0.170	0.138	0.163	0.138
Climate 6a	0.124	0.141	0.152	0.124	0.146

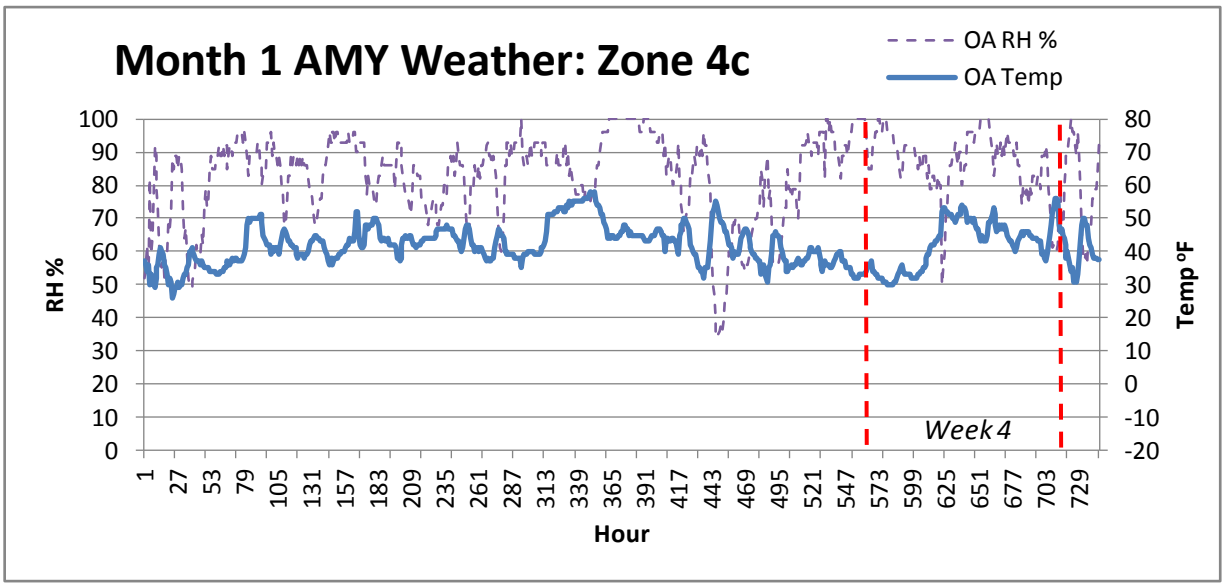
RIGERATION	Month 1	Month 4	Month 7	Month 10	Year
Climate 3a	0.105	0.105	0.105	0.105	0.105
Climate 4c	0.132	0.132	0.132	0.132	0.132
Climate 6a	0.100	0.100	0.100	0.100	0.100

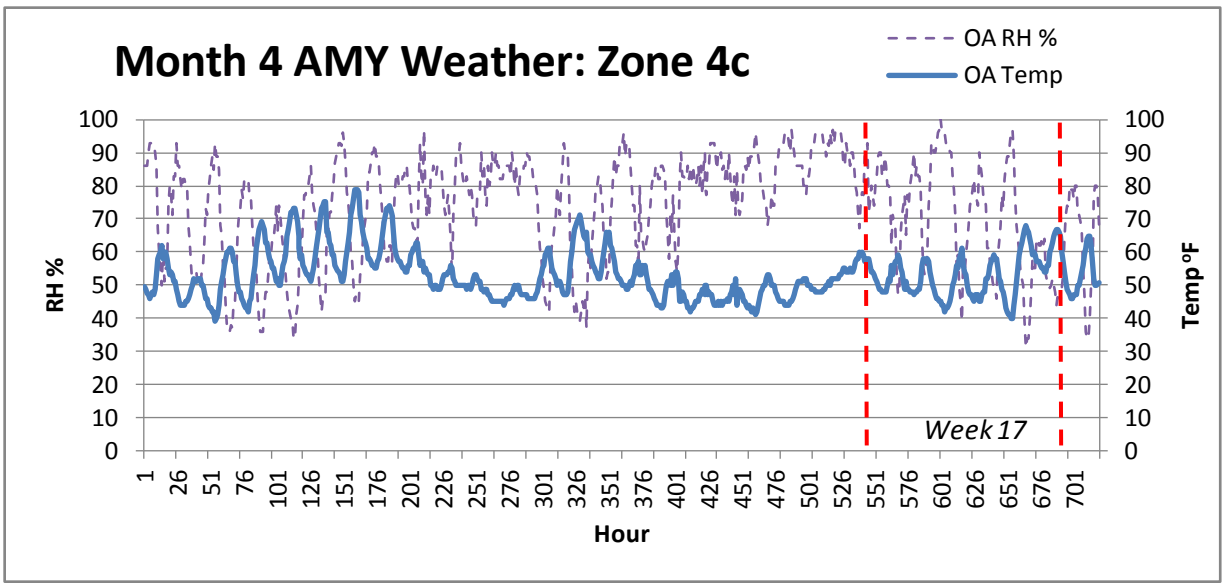


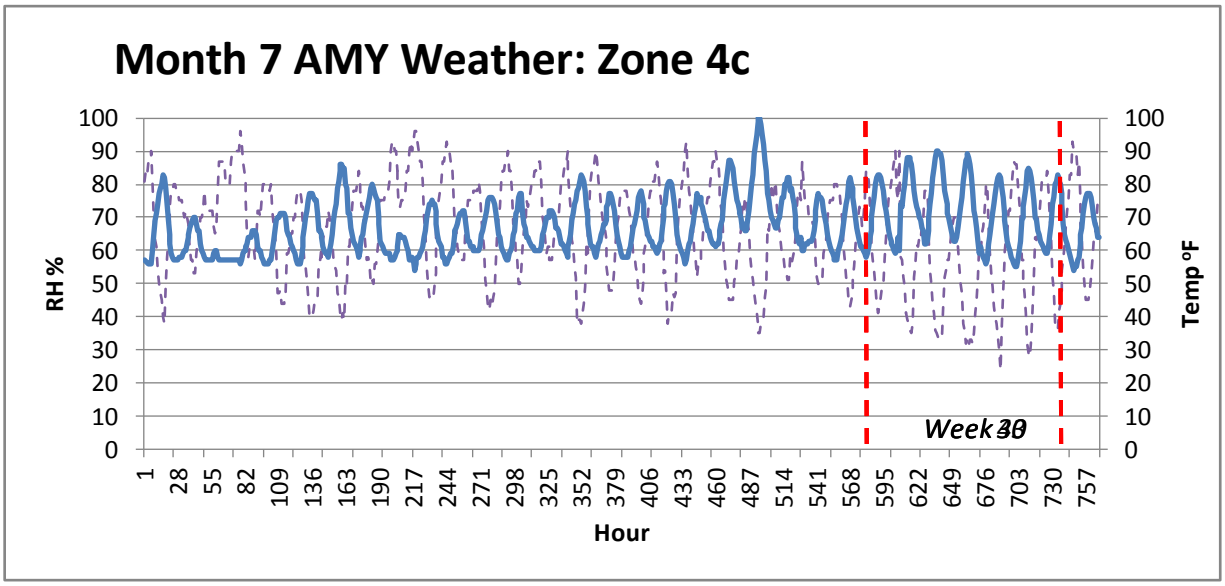

Month 4 AMY Weather: Zone 6a
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 17

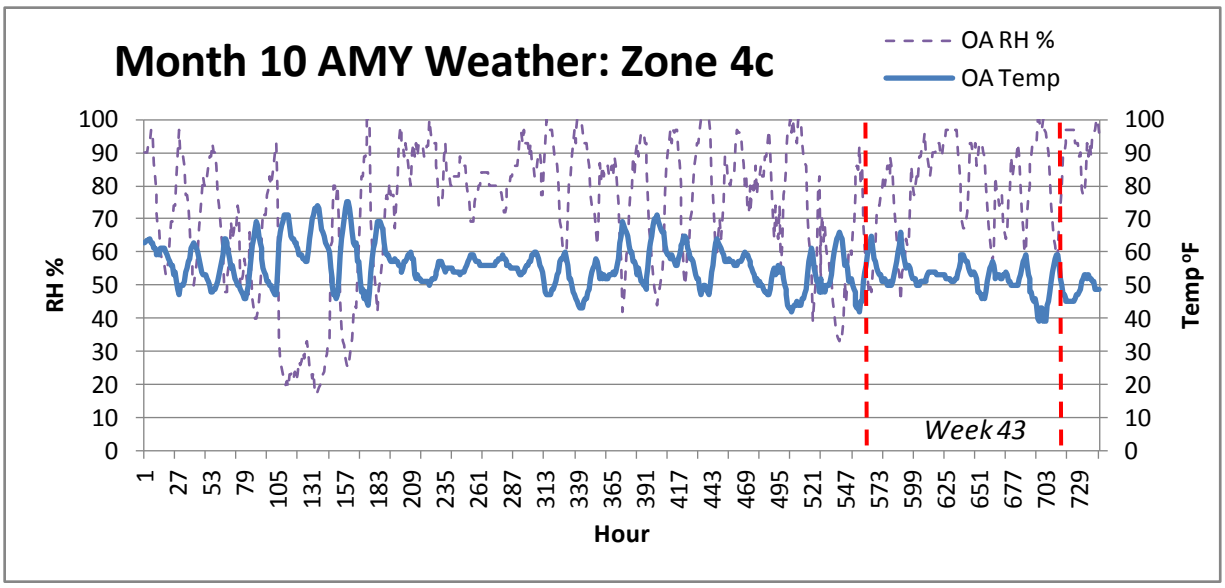

Month 7 AMY Weather: Zone 6a
RH %
Temp °F
Hour
Week 30

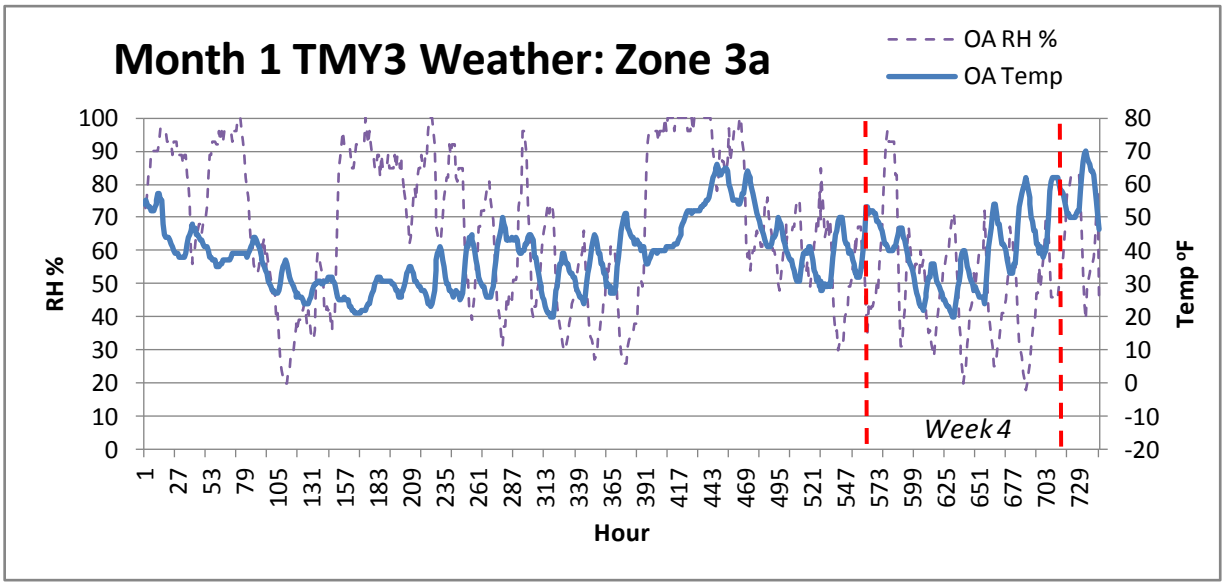

Month 10 AMY Weather: Zone 6a
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 43

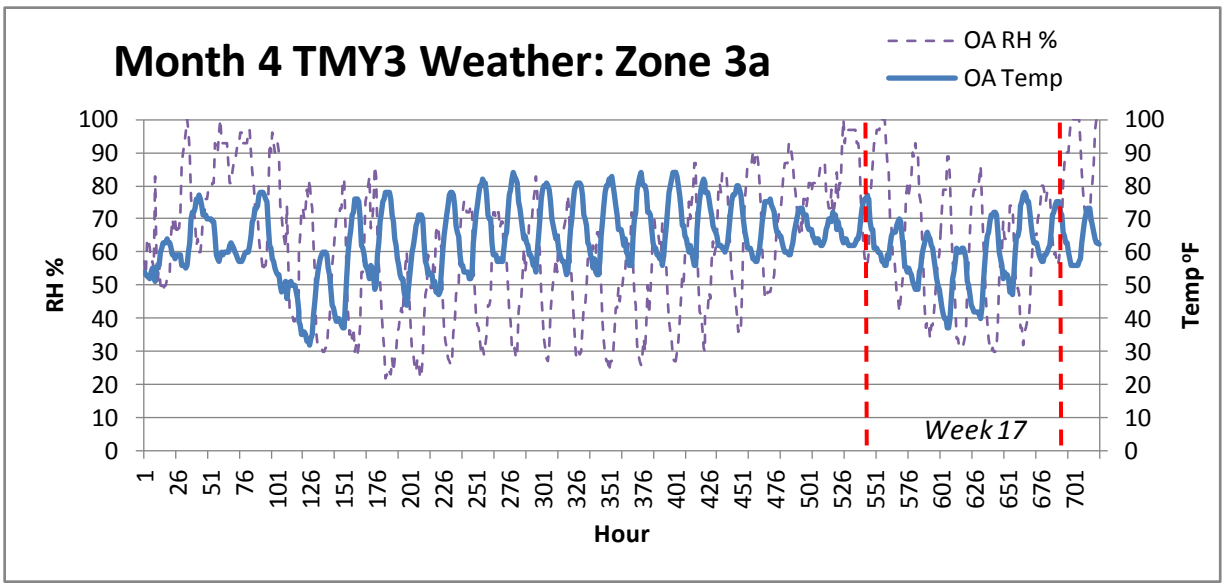

Month 1 AMY Weather: Zone 4c
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 4

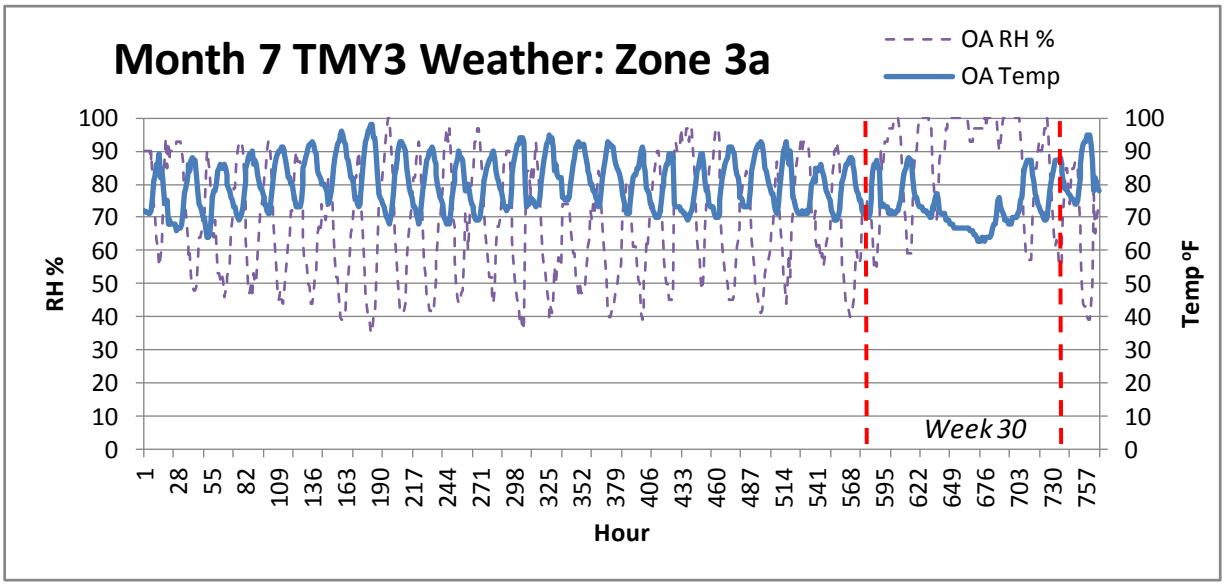

Month 4 AMY Weather: Zone 4c
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 17

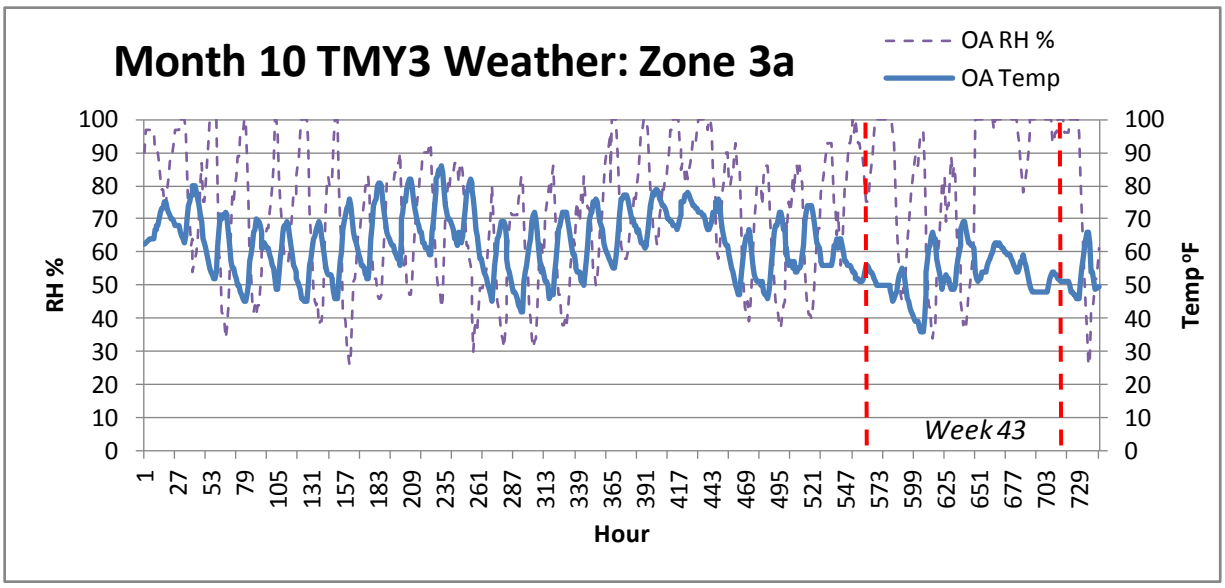

Month 7 AMY Weather: Zone 4c
RH %
Temp °F
Hour
Week 30


Month 10 AMY Weather: Zone 4c
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 43


Month 1 TMY3 Weather: Zone 3a
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 4


Month 4 TMY3 Weather: Zone 3a
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 17


Month 7 TMY3 Weather: Zone 3a
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 30


Month 10 TMY3 Weather: Zone 3a
OA RH %
OA Temp
RH %
Temp °F
Hour
Week 43
100
90
80
70
60
50
40
30
20
10
0
1
27
53
79
105
131
157
183
209
235
261
287
313
339
365
391
417
443
469
495
521
547
573
599
625
651
677
703
729

11.11 Appendix K: Detailed Fault Testing with Skill Metric Results

CLIMATE ZONE 6A (CASE STUDY LOCATION)

Base: As expected base building

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Buildng Energy	1.9	1.5	1.4	1.9	2.3	4.1	6.2	2.2	low ; 2.1	
Natural Gas	1.5	1.4	1.4	1.7	1.5	2.1	2.1	1.7	low ; 2.2	
HVAC	1.6	1.7	1.9	2	2.1	1.6	1.7	2	2	
Refrigeration	4.8	4	3.4	4.3	4.5	4.2	3.9	4.7	high ; 5.2	Σ:
TP										0
FP									3	3
TN	4	4	4	4	4	4	4	4	1	33
FN										0

Fault A: OA Dampers Stuck Open
(50% OA instead of 17% OA)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Building Energy	high ; inf	high ; inf	high ; 1700	high ; inf	high ; inf	high ; 3.6	high ; 4.9	high ; inf	high ; inf	
Natural Gas	high ; inf	high ; inf	high ; 1200	high ; inf	high ; inf	2.2	high ; 3.9	high ; inf	high ; inf	
HVAC	high ; 17	high ; 580	high ; 1700	high ; 3.1	2.1	high ; 5.1	high ; 5.8	high ; 2.5	high ; 3.3	
Refrigeration	4.8	low ; 14	low ; 17	high ; 4.4	high ; 5	4.1	high ; 5.1	high ; 4	high ; 4.3	Σ:
TP	*3*	*3*	*3*	*2*	*2*	*3*	*3*	*2*	*2*	***23***
FP	*0*	*1*	*1*	*0*	*1*	*0*	*1*	*0*	*1*	***5***
TN	*1*	*0*	*0*	*1*	*0*	*1*	*0*	*1*	*0*	***4***
FN	*0*	*0*		*1*	*1*	*0*	*0*	*1*	*1*	***4***

Fault B: Economizer broken
(no ecomomizer for one AHU)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	1.9	1.5	1.4	1.9	2.3	4.2	6	2.2	low ; 2.1	
Natural Gas	1.5	1.4	1.4	1.7	1.5	2.1	2.2	1.6	low ; 2.3	
HVAC	1.6	1.7	1.9	2	2.2	1.6	high ; 2	2	1.9	
Refrigeration	4.8	4	3.4	4.3	4.5	4.2	4	4.6	high ; 5.2	Σ:
TP							1			1
FP									3	3
TN	3	3	3	3	3	3	3	3		24
FN	1	1	1	1	1	1		1	1	8

Fault C: Fans broken (belts snapped)
(Sales1 AHU and Restroom AHU fans off)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	low ; 1300	low ; 430	low ; 580	low ; 350	low ; 73	low ; 5.4	low ; 6.3	low ; 85	low ; 3500	
Natural Gas	low ; 2600	low ; 690	low ; 700	low ; 120	low ; 30	2.2	low ; 3.8	low ; 32	low ; 870	
HVAC	low ; 91	low ; 4.5	low ; 3.8	low ; 210	low ; 580	low ; 120	low ; 270	low ; 4500	low ; 250	
Refrigeration	4.8	4.1	3.4	3.9	4.5	4.2	3.9	4.7	high ; 4.8	Σ:
TP	3	3	3	3	3	2	3	3	3	26
FP									1	1
TN	1	1	1	1	1	1	1	1		8
FN						1				1

Fault D: Refrigeration compressors broken
(Refrigeration Suction UA up by 300% on two racks)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	2.1	high ; 1.6	high ; 1.5	1.9	high ; 2.9	high ; 4.1	high ; 4.9	high ; 2.8	1.7	
Natural Gas	1.5	1.4	1.4	1.7	1.5	2.1	2.2	1.6	low ; 2.3	
HVAC	1.6	1.7	1.9	2	2.2	1.6	1.7	2	1.9	
Refrigeration	high ; Inf	high ; Inf	high ; Inf	high ; Inf	high ; Inf	high ; Inf	high ; Inf	high ; Inf	high ; Inf	Σ:
TP	1	2	2	1	2	2	2	2	1	15
FP									1	1
TN	2	2	2	2	2	2	2	2	1	17
FN	1			1					1	3

Fault E: Scheduling of operation errors:
(All AHUs turn on at 3am instead of 7am)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 15	high ; 11	high ; 8.6	high ; 17	high ; 52	high ; 17	high ; 12	high ; 28	high ; 11	
Natural Gas	high ; 9.3	high ; 7.9	high ; 6	high ; 9.6	high ; 26	high ;7.3	high ; 3.6	high ; 13	high ; 5.2	
HVAC	high ; 39	high ; 39	high ; 37	high ; 32	high ; 17	high ; 47	high ; 28	high ; 15	high ; 24	
Refrigeration	high ; 4.9	high ; 4	high ; 4.1	high ; 3.9	high ; 3.9	3.8	high ; 4.4	high ; 3.8	high ; 14	Σ:
TP	3	3	3	3	3	3	3	3	3	27
FP	1	1	1	1	1		1	1	1	8
TN						1				1
FN										0

Fault F: Thermostat setbacks not working :
(no night setbacks or setups)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 200	high ; 74	high ; 43	high ; 170	high ; 270	high ; 29	high ; 13	high ; 230	high ; 94	
Natural Gas	high ; 95	high ; 48	high ; 26	high ; 83	high ; 130	high ; 7.9	high ; 5.5	high ; 140	high ; 33	
HVAC	high ; 76	high ; 80	high ; 66	high ; 33	high ; 20	high ; 120	high ; 28	high ; 17	high ; 41	
Refrigeration	3.9	high ; 4.6	high ; 5.7	3.7	3.7	3.7	3.9	4	high ; 13	Σ:
TP	3	3	3	3	3	3	3	3	3	27
FP		1	1						1	3
TN	1			1	1	1	1	1		6
FN										0

Fault G: Cooling setpoint overrides:
(Cooling T;stats set down 1.8oC)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 5.6	high ; 3.8	high ; 4	high ; 2.9	high ; 4.1	high ; 180	high ; 500	high ; 5.3	high ; 1.9	
Natural Gas	high ; 2.8	high ; 3.2	high ; 3.1	high ; 2	high ; 2.2	high ; 52	high ; 10	high ; 2.6	1.6	
HVAC	high ; 370	high ; 18	high ; 13	high ; 32	high ; 23	high ; 250	high ; Inf	high ; 43	high ; 21	
Refrigeration	4.2	3.7	3	3.6	3.7	low ; 4.4	4.2	4.1	high ; 3.6	Σ:
TP	2	2	2	2	2	2	2	2	2	18
FP	1	1	1	1	1	2	1	1	1	10
TN	1	1	1	1	1		1	1	1	8
FN	2									2

Fault H: Heat recovery is being bypassed:
(Delete Air;to;Air HX on one AHU)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 62	high ; 23	high ; 11	high ; 1700	high ; 580	3.8	high ; 4.3	high ; 870	high ; 500	
Natural Gas	high ; 160	high ; 26	high ; 11	high ; Inf	high ; Inf	high ; 2.6	high ; 3	high ; Inf	high ; 870	
HVAC	high ; 1.9	high ; 2.4	high ; 2.5	1.9	2.1	high ; 1.8	high ; 2.1	1.8	1.8	
Refrigeration	4.6	3.9	3.4	high ; 4	high ; 4.6	4.1	high ; 4.9	4.3	high ; 6	Σ:
TP	3	3	3	2	2	2	3	2	2	22
FP				1	1		1		1	4
TN	1	1	1			1		1		5
FN				1	1	1		1	1	5

Fault I:Humidity controls fail or are shut off:
(EMS code for humidity Control is disabled)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng										
s to the root cause end-use fault	4	1.5	1.5	1.9	2.3	low ; 56	low ; 6.3	low ; 3	low ; 2.1	
rrespond to the root cause	9	1.4	1.4	1.7	1.5	low ; 51	low ; 140	low ; 2.3	low ; 2.3	
nd-use is not faulted		1.7	1.8	2	2.1	low ; Inf	low ; 71	low ; 26	2	
ually the end-use is faulted	.9	3.9	high ; 3.5	4.2	4.4	high ; 6.6	high ; 55	high ; 4.5	high ; 5.2	Σ:
	2					2	2	2		8
	2		1			2	2	2	3	12
TN		4	3	2	2				1	12
FN				2	2					4

Fault J:Thermostat Float temp increase:
(Increase float to 1.5 oC for all zones)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	2	1.5	high ; 1.6	1.9	low ; 3.3	low ; 4.2	low ; 4.2	low ; 2.7	low ; 2.2	
Natural Gas	low ; 1.5	1.5	high ; 1.6	1.7	low ; 3.2	low ; 2.8	low ; 4.3	low ; 2.2	low ; 2.4	
HVAC	low ; 5	1.8	1.8	low ; 2.9	low ; 6.3	low ; 36	low ; 11	low ; 5.5	low ; 2.5	
Refrigeration	4.6	3.8	high ; 3.6	high ; 4	high ; 4.1	4.1	high ; 4.4	4	high ; 10	Σ:
TP	2		2	1	3	3	3	3	3	20
FP			1	1	1		1		1	5
TN	1	1				1		1		4
FN	1	3	1	2						7

CLIMATE ZONE 4C (PORTLAND)

Base: As expected base building

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Buildng Energy	2	1.8	low ; 2.3	1.8	high ; 2.1	3.5	3.5	2.1	1.7	
Natural Gas	1.5	1.7	low ; 2.3	1.7	high ; 2.2	2.4	low ; 2.7	1.9	1.7	
HVAC	1.9	2.2	2.2	2.3	2.2	2	1.8	2	2.1	
Refrigeration	4.3	3.8	3.7	4	3.8	4.1	3.9	4	3.9	Σ:
TP										0
FP			2		2		1			5
TN	4	4	2	4	2	4	3	4	4	31
FN										0

Fault A: OA Dampers Stuck Open
(50% OA instead of 17% OA)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Building Energy	high ; inf	high ; inf	high ; inf	high ; inf	high ; inf	high ; 3.3	3.6	high ; inf	high ; inf	
Natural Gas	high ; inf	high ; inf	high ; inf	high ; inf	high ; inf	high ; 3.4	low ; 4.1	high ; inf	high ; inf	
HVAC	1.9	high ; 2.5	high ; 4.1	2.1	2	high ; 1.9	high ; 3.3	2	1.9	
Refrigeration	4.3	3.7	3.6	4.2	4.1	4.1	4	4.4	4	Σ:
TP	*2*	*3*	*3*	*2*	*2*	*3*	*2*	*2*	*2*	***21***
FP	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	***0***
TN	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	***9***
FN	*1*	*0*		*1*	*1*	*0*	*1*	*1*	*1*	***6***

Fault B: Economizer broken
(no ecomomizer for one AHU)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	2	1.8	low ; 2.3	1.8	high ; 2.1	3.5	3.5	2.1	1.7	
Natural Gas	1.5	1.7	low ; 2.3	1.7	high ; 2.2	2.4	low ; 2.8	1.9	1.7	
HVAC	1.9	2.2	2.2	2.3	2.2	1.9	1.8	2	2.1	
Refrigeration	4.3	3.8	3.7	4	3.8	4.1	3.9	4	3.9	Σ:
TP					1					1
FP			2		1		1			4
TN	2	4	2	2	1	2	1	2	4	20
FN	2			2	1	2	2	2		11

Fault C: Fans broken (belts snapped)
(Sales1 AHU and Restroom AHU fans off)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	low ; 2.2	high ; 2.8	1.5	high ; 4.1	high ; 40	low ; 7.5	low ; 3.4	2	high ; 4.3	
Natural Gas	high ; 7	high ; 6.4	high ; 1.8	high ; 18	high ; 800	low ; 3.8	low ; 5.3	high ; 3.2	high ; 13	
HVAC	low ; 21	low ; 10	low ; 9.3	low ; 13	low ; 11	low ; 28	low ; 3.9	low ; 15	low ; 12	
Refrigeration	4.3	3.8	3.7	4.2	4.3	3.3	4.2	4.5	4.2	Σ:
TP	2	1	1	1	1	2	2	1	1	12
FP	1	2	1	2	2	1	1	1	2	13
TN	1	1	1	1	1	1	1	1	1	9
FN			1					1		2

Fault D: Refrigeration compressors broken
(Refrigeration Suction UA up by 300% on two racks)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 2.3	1.5	low ; 1.6	high ; 2.4	high ; 3.3	high ; 3.8	high ; 3.2	high ; 2.8	high ; 2.2	
Natural Gas	1.5	1.7	low ; 2.3	1.7	high ; 2.2	2.4	high ; 2.7	1.9	1.7	
HVAC	1.9	2.2	2.2	2.3	2.2	2	1.8	2	2.1	
Refrigeration	high ; Inf	high ; 4500	high ; 4500	high ; Inf	high ; Inf	high ; 1100	high ; 2200	high ; Inf	high ; Inf	Σ:
TP	2	1	2	2	2	2	2	2	2	17
FP			1		1		1			3
TN	2	2	1	2	1	2	1	2	2	15
FN		1								1

Fault E: Scheduling of operation errors:
(All AHUs turn on at 3am instead of 7am)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 3.6	high ; 4	high ; 2.6	high ; 5	high ; 9.9	3.4	3.3	high ; 5	high ; 4.9	
Natural Gas	high ; 7.3	high ; 5.4	high ; 3	high ; 8.5	high ; 18	high ; 3	low ; 3.2	high ; 8.2	high ; 6.9	
HVAC	2	1.9	1.9	2.3	2.3	1.8	1.9	2.1	2.2	
Refrigeration	4	3.7	3.5	4	3.8	4.1	4	4.1	3.9	Σ:
TP	2	2	2	2	2	1		2	2	15
FP							1			1
TN	1	1	1	1	1	1	1	1	1	9
FN	1	1	1	1	1	2	2	1	1	11

Fault F: Thermostat setbacks not working :
(no night setbacks or setups)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 11	high ; 25	high ; 13	high ; 11	high ; 26	3.6	3.4	high ; 6.4	high ; 7.1	
Natural Gas	high ; 47	high ; 57	high ; 21	high ; 25	high ; 79	high ; 2.8	low ; 2.6	high ; 14	high ; 12	
HVAC	1.8	2	2	2.1	2	2	1.7	2.1	2.2	
Refrigeration	4.3	3.7	3.8	4	3.9	4.2	3.9	4.3	4.2	Σ:
TP	2	2	2	2	2	1		2	2	15
FP		1	1				1			3
TN	1	1	1	1	1	1	1	1	1	9
FN	1	1	1	1	1	2	1	1	1	10

Fault G: Cooling setpoint overrides:
(Cooling T;stats set down 1.8oC)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 2.8	1.5	low ; 1.7	high ; 2.7	high ; 4.2	high ; 31	high ; 21	high ; 3.4	high ; 2.3	
Natural Gas	high ; 2.8	1.5	low ; 1.7	high ; 3.3	high ; 4.8	high ; 210	high ; 21	high ; 2.1	high ; 2.6	
HVAC	high ; 4.5	2.1	2.1	high ;2.6	high ; 2.4	high ; 26	high ; 62	high ; 2.7	2.1	
Refrigeration	low ; 4.4	low ; 4.4	low ; 4.1	low ; 4.5	3.9	low ; 4	low ; 3.6	low ; 4	4	Σ:
TP	2			2	2	2	2	2		12
FP	2	1	3	2	1	2	2	2	2	17
TN		3	1		1				2	7
FN										0

Fault H: Heat recovery is being bypassed:
(Delete Air;to;Air HX on one AHU)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 17	high ; 10	high ; 5.8	high ; 130	high ; 270	high ; 3.3	3.2	high ; 120	high ; 84	
Natural Gas	high ; 86	high ; 17	high ; 8.1	high ; 660	high ; 1000	high ; 5.2	2.1	high ; 670	high ; 190	
HVAC	1.7	2.1	2	2.2	2.2	1.6	1.9	2	2	
Refrigeration	4.2	3.6	3.5	4.2	4.1	4.1	4	4.4	4.1	Σ:
TP	2	2	2	2	2	2	2	2	2	18
FP										0
TN	2	2	2	2	2	1	1	2	2	16
FN						1	1			2

Fault I:Humidity controls fail or are shut off:
(EMS code for humidity Control is disabled)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	low ; 2.9	low ; 1.8	low ; 2.3	2	1.9	low ; 51	low ; 17	low ; 2.5	1.9	
Natural Gas	low ; 2.1	1.7	low ; 2.3	1.7	1.8	low ; 59	low ; 150	low ; 2.4	1.8	
HVAC	low ; 8.1	2.2	2.2	low ; 3.1	2.5	low ; 750	low ; 55	low ; 3.6	low ; 2.8	
Refrigeration	4	3.7	3.7	4.2	4	high ; 5.5	high ; 14	3.9	4	Σ:
TP	2	1		1		2	2	2	1	11
FP	1		2			2	2	1		8
TN	1	3	2	2	2			1	3	14
FN				1	2					3

Fault J:Thermostat Float temp increase:
(Increase float to 1.5 oC for all zones)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	2.2	1.6	low ; 2.0	low ; 2.2	1.8	low ; 4.5	low ; 11	low ; 2.7	low ; 2.2	
Natural Gas	1.5	high ; 1.7	low ; 1.9	1.8	1.8	low ; 2.7	low ; 17	low ; 2.3	1.9	
HVAC	low ; 6	low ; 3	low ; 3	low ; 4.4	low ; 5	low ; 19	low ; 16	low ; 4.7	low ; 4.6	
Refrigeration	4.3	3.7	3.6	4.1	3.9	4.1	4	4.3	3.8	Σ:
TP	1	1	3	2	1	3	3	3	2	19
FP										0
TN	1	1	1	1	1	1	1	1	1	9
FN	2	2		1	2				1	8

CLIMATE ZONE 3A (ATLANTA)

Base: As expected base building

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Buildng Energy	2.4	1.7	low ; 2.5	1.9	3.5	3	2.8	3.2	high ; 2.7	
Natural Gas	1.7	1.9	low ; 3.2	1.7	2.1	3.7	3.9	1.8	high ; 2.6	
HVAC	1.4	1.6	1.7	2	1.6	1.5	1.4	1.7	high ; 1.9	
Refrigeration	3.7	4	3.5	4.3	3.9	3	2.6	3.9	4.3	Σ:
TP										0
FP			2						3	5
TN	4	4	2	4	4	4	4	4	1	31
FN										0

Fault A: OA Dampers Stuck Open
(50% OA instead of 17% OA)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Building Energy	high ; 150	high ; inf	high ; inf	high ; 6.2	high ; 12	3	2.7	high ; 120	high ; inf	
Natural Gas	high ; inf	high ; inf	high ; inf	high ; 180	high ; 240	low ; 5.7	low ; 5.1	high ; inf	high ; inf	
HVAC	high ; 1.9	high ; 4.4	high ; 3.2	1.8	1.8	high ; 1.6	high ; 1.8	1.7	high ; 3.9	
Refrigeration	3.7	3.5	high ; 4.5	3.9	4	3	2.4	high ; 4	3.6	Σ:
TP	*3*	*3*	*3*	*2*	*2*	*2*	*2*	*2*	*3*	***22***
FP	*0*	*0*	*1*	*0*	*0*	*0*	*0*	*1*	*0*	***2***
TN	*1*	*1*	*0*	*1*	*1*	*1*	*1*	*0*	*1*	***7***
FN	*0*	*0*	*0*	*1*	*1*	*1*	*1*	*1*	*0*	***5***

Fault B: Economizer broken
(no ecomomizer for one AHU)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	2.4	1.7	low ; 2.6	3.7	3.5	3	2.8	3.1	high ; 2.7	
Natural Gas	1.7	1.9	low ; 3.3	2.1	2.1	3.9	3.9	1.9	high ; 2.6	
HVAC	1.4	1.6	1.7	1.7	1.6	1.5	1.5	1.7	high ; 1.9	
Refrigeration	3.6	4	high ; 3.7	3.9	3.9	3	2.6	3.9	4.3	Σ:
TP									2	2
FP			3						1	4
TN	2	2	0	2	2	2	2	2		14
FN	2	2	1	2	2	2	2	2	1	16

Fault C: Fans broken (belts snapped)
(Sales1 AHU and Restroom AHU fans off)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	2.5	high ; 2.9	1.6	low ; 4,5	low ; 7.5	low ; 3.6	low ; 3.4	low ; 3.5	low ; 3	
Natural Gas	high ; 7.1	high ; 6.5	high ; 2.1	low ; 3.1	low ; 6.7	3.7	3.9	2	2.1	
HVAC	low ; 7.9	low ; 7.3	low ; 9.4	low ; 8.4	low ; 9.8	low ; 4.4	low ; 3.7	low ; 15	low ; 11	
Refrigeration	3.8	3.9	3.6	3.9	4.1	3.3	2.9	4.1	4.5	Σ:
TP	1	1	1	3	3	2	2	2	2	17
FP	1	2	1							4
TN	1	1	1	1	1	2	2	1	1	11
FN	1		1					1	1	4

Fault D: Refrigeration compressors broken
(Refrigeration Suction UA up by 300% on two racks)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 2.8	1.5	low ; 1.9	high ; 3.5	high ; 2.9	2.8	2.6	high ; 3.6	high ; 5.1	
Natural Gas	1.7	1.9	low ; 3.2	2.1	2.1	3.7	3.9	1.8	high ; 2.6	
HVAC	1.4	1.6	1.7	1.7	1.6	1.5	1.4	1.7	high ; 1.9	
Refrigeration	high ; 1500	high ; 2200	high ; Inf	high ; 4500	high ; Inf	high ; 48	high ; 30	high ; Inf	high ; Inf	Σ:
TP	2	1	1	2	2	1	1	2	2	14
FP			2						2	4
TN	2	2	1	2	2	3	3	2	0	17
FN		1							0	1

Fault E: Scheduling of operation errors:
(All AHUs turn on at 3am instead of 7am)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 3.7	high ; 5	high ; 2.6	3.2	high ; 3.6	2.9	2.6	high ; 3.6	high ; 7.6	
Natural Gas	high ; 9.5	high ; 7.5	high ; 2.9	high ; 5.3	high ; 7.1	high ; 3.1	2.2	high ; 5.7	high ; 17	
HVAC	high ; 1.8	1.7	1.7	1.8	1.8	high ; 2.4	high ; 2	1.6	high ; 1.8	
Refrigeration	3.7	4	high ; 3.8	3.9	3.8	3	2.5	4	4.4	Σ:
TP	3	3	2	1	2	2	1	2	3	19
FP			1							1
TN	1	1		1	1	1	1	1	1	8
FN			1	2	1	1	2	1		8

Fault F: Thermostat setbacks not working :
(no night setbacks or setups)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 9.9	high ; 46	high ; 23	high ; 3.6	high ; 3.7	high ; 3.5	2.6	high ; 3.6	high ; 9.1	
Natural Gas	high ; 55	high ; 120	high ; 37	high ; 9	high ; 7.3	high ; 3.2	2.2	high ; 7.7	high ; 26	
HVAC	high ; 2.9	high ; 1.9	high ; 1.7	1.8	high ; 2	high ; 6.7	high ; 3.4	1.7	1.7	
Refrigeration	3.8	3.9	high ; 4	4	4	3.2	2.5	4.1	4.1	Σ:
TP	3	3	3	2	3	3	1	2	2	22
FP			1							1
TN	1	1		1	1	1	1	1	1	8
FN				1			2	1	1	5

Fault G: Cooling setpoint overrides:
(Cooling T;stats set down 1.8oC)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 3.3	1.5	low ; 1.8	high ; 9.2	high ; 10	high ; 3	high ; 3.5	high ; 12	high ; 10	
Natural Gas	high ; 4.3	1.8	low ; 2.3	high ; 15	high ; 18	high ; 14	high ; 5.8	high ; 16	high ; 16	
HVAC	high ; 5.4	high ; 2.2	high ; 1.9	high ; 21	high ; 11	high ; 2	high ; 7.2	high ; 13	high ; 5.7	
Refrigeration	low ; 4.1	low ; 4.7	3.4	low ; 4	4	low ; 3.9	low ; 3.7	4.2	4.1	Σ:
TP	2	1	1	2	2	2	2	2	2	16
FP	2	1	2	2	1	2	2	1	1	14
TN		1	1		1			1	1	5
FN		1								1

Fault H: Heat recovery is being bypassed:
(Delete Air;to;Air HX on one AHU)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	high ; 7.8	high ; 20	high ; 6.7	high ; 3.9	high ; 4.4	2.9	2.8	high ; 12	high ; 1200	
Natural Gas	high ; 74	high ; 43	high ; 10	high ; 22	high ; 22	3.9	2.9	high ; 140	high ; Inf	
HVAC	1.5	1.6	1.6	1.7	1.7	1.4	1.5	1.6	high ; 2.6	
Refrigeration	3.7	3.8	high ; 3.9	3.9	4	2.9	2.4	high ; 4	3.9	Σ:
TP	2	2	2	2	2	0	0	2	2	14
FP			1					1		2
TN	2	2	1	2	2	2	2	1	1	15
FN						2	2		1	5

Fault I:Humidity controls fail or are shut off:
(EMS code for humidity Control is disabled)

End Use	Year Signal	Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Whole Buildng Energy	low ; 5	low ; 1.6	low ; 2.8	low ; 10	low ; 9.6	low ; 29	low ; 28	low ; 10	low ; 3.1	
Natural Gas	low ; 3.1	2	low ; 3.6	low ; 41	low ; 32	low ; 5.1	low ; 5.6	low ; 16	low ; 3.4	
HVAC	low ; 60	1.9	1.7	low ; 11	low ; 11	low ; 1000	low ; 990	low ; 74	low ; 3.8	
Refrigeration	high ; 4.3	3.6	high ; 3.7	high ; 4	high ; 4.8	high ; 4.3	high ; 4.3	high ; 9.5	high ; 4.6	Σ:
TP	2	1	1	2	2	2	2	2	2	16
FP	2		2	2	2	2	2	2	2	16
TN		2	1							3
FN		1								1

s to the root cause end-use fault
rrespond to the root cause
nd-use is not faulted
ually the end-use is faulted

Fault J:Thermostat Float temp increase:
(Increase float to 1.5 oC for all zones)

		Month 1 Signal	Week 4 Signal	Month 4 Signal	Week 17 Signal	Month 7 Signal	Week 30 Signal	Month 10 Signal	Week 43 Signal	
Energy	low ; 2.8	1.5	1.5	low ; 4.2	low ; 5.8	low ; 5.6	low ; 7.4	low ; 3,5	2.5	
Natural Gas	1.9	1.7	1.8	low ; 7.6	low ; 20	low ; 5.7	low ; 5.3	low ; 3.8	2.3	
HVAC	low ; 7.7	low ; 2.1	low ; 2.2	low ; 5.2	low ; 3.9	low ; 17	low ; 23	low ; 4.5	low ; 2.5	
Refrigeration	3.7	3.6	high ; 4.7	3.6	4	2.8	2.4	high ; 4	3.8	Σ:
TP	2	1	1	3	3	3	3	3	1	20
FP			1					1		2
TN	1	1		1	1	1	1		1	7
FN	1	2	2						2	7

Table 35

YEARLY faults, sorted by descending signal priority ratio											
Whole Building Energy			Natural Gas			HVAC			Refrigeration		
Location	95% Conf. Cost Dev (mean)	Signal Priority Ratio	Location	95% Conf. Cost Dev (mean)	Signal Priority Ratio	Location	95% Conf. Cost Dev (mean)	Signal Priority Ratio	Location	95% Conf. Cost Dev (mean)	Signal Priority Ratio
Climate 4c, Fault A	25086.5	6000	Climate 4c, Fault A	10648.5	6000	Climate 6a, Fault A	7002.5	430	Climate 4c, Fault D	3654.5	6000
Climate 6a, Fault A	42951	6000	Climate 3a, Fault A	6387	6000	Climate 6a, Fault G	6889	370	Climate 6a, Fault D	4742	6000
Climate 6a, Fault H	27176.5	370	Climate 6a, Fault A	18619	6000	Climate 6a, Fault F	5703.5	76	Climate 3a, Fault D	3818.5	1500
Climate 6a, Fault F	24198.5	200	Climate 6a, Fault H	11458	310	Climate 3a, Fault I	5978	60	Climate 6a, Fault I	1363	5.2
Climate 3a, Fault A	16870.5	150	Climate 6a, Fault F	9678	95	Climate 6a, Fault E	5117	39	Climate 4c, Fault G	854.5	4.4
Climate 6a, Fault C	15969.5	18	Climate 4c, Fault H	5382.5	85	Climate 6a, Fault H	4754	28	Climate 4c, Fault A	250	4.3
Climate 4c, Fault H	12465.5	17	Climate 3a, Fault H	3397.5	74	Climate 4c, Fault C	4248	21	Climate 4c, Fault B	63.5	4.3
Climate 6a, Fault E	15787.5	15	Climate 3a, Fault F	3277	55	Climate 6a, Fault B	4341	20	Climate 4c, Fault C	378.5	4.3
Climate 4c, Fault F	10858	11	Climate 4c, Fault F	4942.5	47	Climate 6a, Fault D	4309	19	Climate 4c, Fault E	85	4.3
Climate 3a, Fault F	9701.5	9.9	Climate 6a, Fault C	8084	35	Climate 4c, Fault I	3268	8.1	Climate 4c, Fault F	459.5	4.3
Climate 3a, Fault H	8860	7.8	Climate 3a, Fault E	2018	9.5	Climate 3a, Fault C	3728	7.9	Climate 4c, Fault J	256.5	4.3
Climate 6a, Fault D	10244.5	5.6	Climate 6a, Fault E	5701	9.3	Climate 3a, Fault J	3695.5	7.7	Climate 3a, Fault I	1191.5	4.3
Climate 6a, Fault G	10265.5	5.6	Climate 4c, Fault E	3058.5	8.8	Climate 4c, Fault J	2845	6	Climate 6a, Fault A	237.5	4.3
Climate 3a, Fault I	6867.5	5	Climate 3a, Fault C	1767	7.1	Climate 3a, Fault G	3064.5	5.4	Climate 4c, Fault H	178.5	4.2
Climate 4c, Fault E	6817.5	4.2	Climate 4c, Fault C	2780.5	7	Climate 6a, Fault J	2601	5.1	Climate 6a, Fault G	322.5	4.2
Climate 3a, Fault E	5965	3.7	Climate 3a, Fault G	1291	4.3	Climate 4c, Fault G	2251	4.5	Climate 3a, Fault G	874.5	4.1
Climate 3a, Fault G	5283.5	3.3	Climate 3a, Fault I	1158	3.1	Climate 3a, Fault F	1982	2.9	Climate 6a, Fault C	458.5	4.1
Climate 6a, Fault B	7154.5	3.3	Climate 4c, Fault G	1513.5	2.8	Climate 6a, Fault I	1233	2.1	Climate 6a, Fault J	790.5	4.1
Climate 4c, Fault I	4905	2.9	Climate 6a, Fault G	2701.5	2.8	Climate 4c, Fault A	545	1.9	Climate 4c, Fault I	584.5	4
Climate 4c, Fault G	4886.5	2.8	Climate 4c, Fault I	1273.5	2.1	Climate 4c, Fault B	535	1.9	Climate 6a, Fault E	737	4
Climate 3a, Fault D	3184	2.8	Climate 6a, Fault B	1759	2	Climate 4c, Fault D	565	1.9	Climate 6a, Fault H	773	4
Climate 3a, Fault J	3705	2.8	Climate 6a, Fault D	1759	2	Climate 3a, Fault A	1235	1.9	Climate 6a, Fault B	637	3.9
Climate 3a, Fault C	808	2.5	Climate 3a, Fault J	377	1.9	Climate 4c, Fault E	365.5	1.8	Climate 6a, Fault F	639	3.9
Climate 6a, Fault J	5431.5	2.5	Climate 6a, Fault J	1502	1.9	Climate 4c, Fault F	262.5	1.8	Climate 3a, Fault C	323.5	3.8
Climate 3a, Fault B	354.5	2.4	Climate 3a, Fault B	208	1.7	Climate 3a, Fault E	1014	1.8	Climate 3a, Fault F	175	3.8
Climate 4c, Fault D	2773	2.3	Climate 3a, Fault D	188	1.7	Climate 4c, Fault H	184	1.7	Climate 3a, Fault A	254.5	3.7
Climate 4c, Fault C	2634.5	2.2	Climate 6a, Fault I	791	1.6	Climate 6a, Fault C	264	1.7	Climate 3a, Fault E	2.5	3.7
Climate 4c, Fault J	2400	2.2	Climate 4c, Fault B	54.5	1.5	Climate 3a, Fault H	351	1.5	Climate 3a, Fault H	201	3.7
Climate 4c, Fault B	366.5	2	Climate 4c, Fault D	54.5	1.5	Climate 3a, Fault B	68	1.4	Climate 3a, Fault J	340.5	3.7
Climate 6a, Fault I	3391.5	2	Climate 4c, Fault J	204.5	1.5	Climate 3a, Fault D	20.5	1.4	Climate 3a, Fault B	14.5	3.6

Table 36

MONTH 1 faults, sorted by descending signal priority ratio											
Whole Building Energy			Natural Gas			HVAC			Refrigeration		
Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio
Climate 4c, Fault A	41784	6000	Climate 4c, Fault A	20520	6000	Climate 6a, Fault A	11088	6000	Climate 6a, Fault D	4398	6000
Climate 3a, Fault A	46302	6000	Climate 3a, Fault A	22602	6000	Climate 6a, Fault F	6732	80	Climate 4c, Fault D	3546	4500
Climate 6a, Fault A	69222	6000	Climate 6a, Fault A	38394	6000	Climate 6a, Fault E	5988	39	Climate 3a, Fault D	3336	2200
Climate 6a, Fault H	40524	96	Climate 3a, Fault F	10788	120	Climate 6a, Fault H	5670	29	Climate 6a, Fault A	1272	7.1
Climate 6a, Fault F	39012	74	Climate 6a, Fault H	22206	67	Climate 6a, Fault G	5136	18	Climate 3a, Fault G	924	4.7
Climate 3a, Fault F	21702	46	Climate 4c, Fault F	9366	57	Climate 6a, Fault I	4680	13	Climate 6a, Fault F	1110	4.6
Climate 6a, Fault C	32592	26	Climate 6a, Fault F	20742	48	Climate 6a, Fault B	4578	12	Climate 4c, Fault G	924	4.4
Climate 4c, Fault F	18414	25	Climate 3a, Fault H	9126	43	Climate 6a, Fault D	4578	12	Climate 6a, Fault E	882	4.2
Climate 3a, Fault H	18222	20	Climate 6a, Fault C	21054	41	Climate 6a, Fault J	4500	12	Climate 3a, Fault B	252	4
Climate 6a, Fault E	23298	11	Climate 4c, Fault H	7302	17	Climate 4c, Fault C	3522	10	Climate 3a, Fault E	228	4
Climate 4c, Fault H	14352	10	Climate 6a, Fault E	11538	7.9	Climate 3a, Fault C	3144	7.3	Climate 4c, Fault E	222	3.9
Climate 3a, Fault E	10776	5	Climate 3a, Fault E	5256	7.5	Climate 3a, Fault A	2616	4.4	Climate 3a, Fault C	426	3.9
Climate 6a, Fault D	14052	4.4	Climate 3a, Fault C	4902	6.5	Climate 6a, Fault C	2466	3.3	Climate 3a, Fault F	372	3.9
Climate 4c, Fault E	9060	4.1	Climate 4c, Fault C	4848	6.4	Climate 4c, Fault J	1530	3	Climate 6a, Fault J	756	3.9
Climate 6a, Fault G	12540	3.8	Climate 4c, Fault E	4572	5.7	Climate 4c, Fault A	1356	2.5	Climate 4c, Fault B	138	3.8
Climate 6a, Fault I	12408	3.8	Climate 6a, Fault G	5730	3.2	Climate 4c, Fault B	504	2.2	Climate 4c, Fault C	534	3.8
Climate 6a, Fault J	12360	3.8	Climate 6a, Fault J	5478	3.1	Climate 4c, Fault D	504	2.2	Climate 3a, Fault H	108	3.8
Climate 6a, Fault B	11748	3.6	Climate 6a, Fault I	5454	3	Climate 4c, Fault I	618	2.2	Climate 6a, Fault H	714	3.8
Climate 3a, Fault C	7128	2.9	Climate 6a, Fault B	5106	2.9	Climate 3a, Fault G	1392	2.2	Climate 6a, Fault I	714	3.8
Climate 4c, Fault C	6618	2.8	Climate 6a, Fault D	5106	2.9	Climate 4c, Fault G	828	2.1	Climate 4c, Fault A	294	3.7
Climate 4c, Fault B	3096	1.8	Climate 3a, Fault I	1728	2	Climate 4c, Fault H	300	2.1	Climate 4c, Fault F	594	3.7
Climate 4c, Fault I	3348	1.8	Climate 3a, Fault B	1254	1.9	Climate 3a, Fault J	924	2.1	Climate 4c, Fault I	78	3.7
Climate 3a, Fault I	3684	1.8	Climate 3a, Fault D	1206	1.9	Climate 4c, Fault F	432	2	Climate 4c, Fault J	210	3.7
Climate 3a, Fault B	2418	1.7	Climate 3a, Fault G	90	1.8	Climate 4c, Fault E	198	1.9	Climate 6a, Fault B	666	3.7
Climate 4c, Fault J	624	1.6	Climate 4c, Fault B	1350	1.7	Climate 3a, Fault F	984	1.9	Climate 6a, Fault G	282	3.7
Climate 4c, Fault D	366	1.5	Climate 4c, Fault D	1350	1.7	Climate 3a, Fault I	600	1.9	Climate 4c, Fault H	150	3.6
Climate 4c, Fault G	42	1.5	Climate 4c, Fault I	1458	1.7	Climate 3a, Fault E	696	1.7	Climate 3a, Fault I	36	3.6
Climate 3a, Fault D	294	1.5	Climate 4c, Fault J	786	1.7	Climate 3a, Fault B	84	1.6	Climate 3a, Fault J	144	3.6
Climate 3a, Fault G	684	1.5	Climate 3a, Fault J	114	1.7	Climate 3a, Fault D	72	1.6	Climate 3a, Fault A	24	3.5
Climate 3a, Fault J	636	1.5	Climate 4c, Fault G	30	1.5	Climate 3a, Fault H	396	1.6	Climate 6a, Fault C	522	3.5

Table 37

MONTH 4 faults, sorted by descending signal priority ratio											
Whole Building Energy			Natural Gas			HVAC			Refrigeration		
Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio
Climate 4c, Fault A	27108	6000	Climate 4c, Fault A	11520	6000	Climate 6a, Fault A	4788	36	Climate 4c, Fault D	3600	6000
Climate 6a, Fault A	52410	6000	Climate 6a, Fault A	23838	6000	Climate 6a, Fault F	4716	33	Climate 6a, Fault D	4758	6000
Climate 6a, Fault H	31242	3500	Climate 6a, Fault H	13722	6000	Climate 6a, Fault G	4680	32	Climate 3a, Fault D	3762	4500
Climate 6a, Fault F	23634	170	Climate 4c, Fault H	7392	660	Climate 6a, Fault E	4326	22	Climate 6a, Fault A	1254	5
Climate 4c, Fault H	17268	130	Climate 3a, Fault A	2112	180	Climate 3a, Fault G	4704	21	Climate 6a, Fault H	1074	4.6
Climate 6a, Fault C	15000	17	Climate 6a, Fault F	9786	83	Climate 4c, Fault C	3750	13	Climate 6a, Fault J	1074	4.6
Climate 6a, Fault E	15732	17	Climate 3a, Fault I	1680	41	Climate 6a, Fault H	3744	12	Climate 4c, Fault G	888	4.5
Climate 4c, Fault F	10716	11	Climate 4c, Fault F	4860	25	Climate 3a, Fault I	4002	11	Climate 4c, Fault E	204	4.3
Climate 3a, Fault I	7170	10	Climate 3a, Fault H	1464	22	Climate 6a, Fault B	3456	9.6	Climate 4c, Fault A	264	4.2
Climate 3a, Fault G	7092	9.2	Climate 4c, Fault C	4344	18	Climate 6a, Fault D	3456	9.5	Climate 4c, Fault C	480	4.2
Climate 3a, Fault A	6138	6.2	Climate 6a, Fault C	7560	16	Climate 6a, Fault I	3342	8.9	Climate 4c, Fault H	240	4.2
Climate 4c, Fault E	8292	6.1	Climate 3a, Fault G	1308	15	Climate 3a, Fault C	3726	8.4	Climate 4c, Fault I	216	4.2
Climate 3a, Fault C	5412	4.5	Climate 4c, Fault E	3660	11	Climate 3a, Fault J	2970	5.2	Climate 4c, Fault J	336	4.1
Climate 3a, Fault J	5208	4.2	Climate 6a, Fault E	5862	9.6	Climate 6a, Fault J	2580	5	Climate 4c, Fault B	48	4
Climate 4c, Fault C	6504	4.1	Climate 3a, Fault F	1086	9	Climate 4c, Fault J	2382	4.4	Climate 4c, Fault F	612	4
Climate 3a, Fault H	4074	3.9	Climate 3a, Fault J	1014	7.6	Climate 4c, Fault I	1500	3.1	Climate 3a, Fault F	246	4
Climate 6a, Fault D	7878	3.8	Climate 3a, Fault E	840	5.3	Climate 4c, Fault G	1170	2.6	Climate 3a, Fault G	786	4
Climate 3a, Fault B	690	3.7	Climate 4c, Fault G	1860	3.3	Climate 4c, Fault B	480	2.3	Climate 3a, Fault I	918	4
Climate 3a, Fault F	3234	3.6	Climate 3a, Fault C	594	3.1	Climate 4c, Fault D	486	2.3	Climate 6a, Fault E	816	4
Climate 3a, Fault D	2934	3.5	Climate 3a, Fault B	228	2.1	Climate 4c, Fault H	318	2.2	Climate 6a, Fault I	822	4
Climate 3a, Fault E	2340	3.2	Climate 3a, Fault D	234	2.1	Climate 4c, Fault A	138	2.1	Climate 3a, Fault A	234	3.9
Climate 6a, Fault G	6414	2.9	Climate 6a, Fault G	1398	2	Climate 4c, Fault E	168	2.1	Climate 3a, Fault B	30	3.9
Climate 4c, Fault G	4680	2.7	Climate 4c, Fault J	696	1.8	Climate 4c, Fault F	186	2.1	Climate 3a, Fault C	414	3.9
Climate 4c, Fault D	3576	2.4	Climate 6a, Fault B	786	1.8	Climate 6a, Fault C	414	2	Climate 3a, Fault E	54	3.9
Climate 6a, Fault B	4896	2.3	Climate 6a, Fault D	786	1.8	Climate 3a, Fault A	546	1.8	Climate 3a, Fault H	198	3.9
Climate 4c, Fault J	3348	2.2	Climate 4c, Fault B	324	1.7	Climate 3a, Fault E	234	1.8	Climate 6a, Fault B	780	3.9
Climate 6a, Fault I	4704	2.2	Climate 4c, Fault D	324	1.7	Climate 3a, Fault F	714	1.8	Climate 3a, Fault J	168	3.8
Climate 6a, Fault J	4368	2.1	Climate 4c, Fault I	468	1.7	Climate 3a, Fault B	132	1.7	Climate 6a, Fault C	432	3.7
Climate 4c, Fault I	2136	2	Climate 6a, Fault I	720	1.7	Climate 3a, Fault D	114	1.7	Climate 6a, Fault F	474	3.7
Climate 4c, Fault B	372	1.8	Climate 6a, Fault J	744	1.7	Climate 3a, Fault H	138	1.7	Climate 6a, Fault G	42	3.6

Table 38

MONTH 7 faults, sorted by descending signal priority ratio												
Whole Building Energy			Natural Gas			HVAC			Refrigeration			
Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio
Climate 6a, Fault G	15366	180	Climate 4c, Fault G	2412	210	Climate 6a, Fault G	11202	3500	Climate 6a, Fault D	4668	6000
Climate 4c, Fault I	11064	51	Climate 4c, Fault I	2004	59	Climate 3a, Fault I	13026	1000	Climate 4c, Fault D	3936	1100
Climate 6a, Fault D	11982	39	Climate 6a, Fault G	1872	52	Climate 4c, Fault I	7914	750	Climate 3a, Fault D	3972	48
Climate 4c, Fault G	10638	31	Climate 3a, Fault G	522	14	Climate 6a, Fault F	8118	120	Climate 6a, Fault I	2256	8.8
Climate 3a, Fault I	12456	29	Climate 6a, Fault I	1632	10	Climate 6a, Fault A	8010	110	Climate 4c, Fault I	1740	5.5
Climate 6a, Fault F	11322	29	Climate 6a, Fault F	1008	7.9	Climate 6a, Fault E	6876	47	Climate 6a, Fault G	1044	4.4
Climate 6a, Fault A	10386	20	Climate 6a, Fault E	960	7.3	Climate 6a, Fault H	6348	34	Climate 3a, Fault I	1890	4.3
Climate 6a, Fault E	10110	17	Climate 3a, Fault A	852	5.7	Climate 4c, Fault C	5310	28	Climate 4c, Fault C	60	4.2
Climate 6a, Fault H	9222	12	Climate 3a, Fault J	858	5.7	Climate 4c, Fault G	5472	26	Climate 4c, Fault F	84	4.2
Climate 4c, Fault C	6930	7.5	Climate 6a, Fault H	792	5.5	Climate 6a, Fault B	5814	22	Climate 4c, Fault A	312	4.1
Climate 6a, Fault B	7662	6.9	Climate 4c, Fault H	1200	5.2	Climate 6a, Fault D	5748	21	Climate 4c, Fault B	12	4.1
Climate 3a, Fault J	8262	5.6	Climate 3a, Fault I	828	5.1	Climate 4c, Fault J	4830	19	Climate 4c, Fault E	216	4.1
Climate 4c, Fault J	5592	4.5	Climate 6a, Fault A	612	4.2	Climate 3a, Fault J	6978	17	Climate 4c, Fault H	300	4.1
Climate 4c, Fault D	4332	3.8	Climate 3a, Fault B	690	3.9	Climate 3a, Fault F	5610	6.7	Climate 4c, Fault J	180	4.1
Climate 4c, Fault E	1266	3.6	Climate 3a, Fault H	690	3.9	Climate 6a, Fault I	3192	4.8	Climate 4c, Fault G	792	4
Climate 4c, Fault F	936	3.6	Climate 4c, Fault C	696	3.8	Climate 3a, Fault C	3990	4.4	Climate 3a, Fault G	1230	3.9
Climate 3a, Fault C	5868	3.6	Climate 3a, Fault C	678	3.7	Climate 3a, Fault E	2778	2.4	Climate 6a, Fault A	372	3.9
Climate 4c, Fault B	612	3.5	Climate 3a, Fault D	672	3.7	Climate 4c, Fault E	834	2.1	Climate 6a, Fault C	36	3.8
Climate 3a, Fault F	4938	3.5	Climate 4c, Fault A	876	3.4	Climate 4c, Fault D	582	2	Climate 6a, Fault E	288	3.8
Climate 6a, Fault I	4446	3.5	Climate 6a, Fault B	468	3.3	Climate 4c, Fault F	720	2	Climate 6a, Fault H	348	3.8
Climate 4c, Fault A	3402	3.3	Climate 6a, Fault D	468	3.3	Climate 3a, Fault G	2262	2	Climate 6a, Fault J	264	3.8
Climate 4c, Fault H	3426	3.3	Climate 3a, Fault F	246	3.2	Climate 4c, Fault A	894	1.9	Climate 6a, Fault B	144	3.7
Climate 6a, Fault C	474	3.2	Climate 3a, Fault E	258	3.1	Climate 4c, Fault B	486	1.9	Climate 6a, Fault F	180	3.7
Climate 6a, Fault J	1266	3.2	Climate 4c, Fault E	732	3	Climate 6a, Fault J	942	1.9	Climate 3a, Fault C	828	3.3
Climate 3a, Fault A	84	3	Climate 4c, Fault F	672	2.8	Climate 6a, Fault C	318	1.7	Climate 3a, Fault F	564	3.2
Climate 3a, Fault B	1632	3	Climate 4c, Fault J	438	2.7	Climate 4c, Fault H	66	1.6	Climate 3a, Fault A	420	3
Climate 3a, Fault G	3012	3	Climate 4c, Fault B	420	2.4	Climate 3a, Fault A	1122	1.6	Climate 3a, Fault B	168	3
Climate 3a, Fault E	2418	2.9	Climate 4c, Fault D	432	2.4	Climate 3a, Fault B	294	1.5	Climate 3a, Fault E	264	3
Climate 3a, Fault H	684	2.9	Climate 6a, Fault C	54	2.1	Climate 3a, Fault D	210	1.5	Climate 3a, Fault H	336	2.9
Climate 3a, Fault D	2562	2.8	Climate 6a, Fault J	228	1.9	Climate 3a, Fault H	174	1.4	Climate 3a, Fault J	192	2.8

Table 39

MONTH 10 faults, sorted by descending signal priority ratio											
Whole Building Energy			Natural Gas			HVAC			Refrigeration		
Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio	Location	95% Conf. Cost Dev, μ	Signal Priority Ratio
Climate 4c, Fault A	25056	6000	Climate 4c, Fault A	10158	6000	Climate 3a, Fault I	5508	74	Climate 4c, Fault D	3786	6000
Climate 6a, Fault A	42246	6000	Climate 3a, Fault A	4530	6000	Climate 6a, Fault G	4932	43	Climate 3a, Fault D	4464	6000
Climate 6a, Fault H	26010	3500	Climate 6a, Fault A	16794	6000	Climate 6a, Fault A	4428	25	Climate 6a, Fault D	4998	6000
Climate 6a, Fault F	21042	230	Climate 6a, Fault H	9882	6000	Climate 6a, Fault F	4056	17	Climate 3a, Fault I	1974	9.5
Climate 4c, Fault H	16398	120	Climate 4c, Fault H	6672	670	Climate 4c, Fault C	4026	15	Climate 6a, Fault I	1254	4.9
Climate 3a, Fault A	13116	120	Climate 3a, Fault H	2982	140	Climate 3a, Fault C	4128	15	Climate 4c, Fault C	234	4.5
Climate 6a, Fault E	15024	28	Climate 6a, Fault F	7620	140	Climate 6a, Fault E	3900	15	Climate 6a, Fault A	1026	4.5
Climate 3a, Fault G	8766	12	Climate 3a, Fault G	1926	16	Climate 3a, Fault G	4146	13	Climate 4c, Fault A	270	4.4
Climate 3a, Fault H	8886	12	Climate 3a, Fault I	1974	16	Climate 6a, Fault H	3552	11	Climate 4c, Fault E	72	4.4
Climate 3a, Fault I	8016	10	Climate 4c, Fault F	3576	14	Climate 6a, Fault B	3336	8.9	Climate 4c, Fault H	234	4.4
Climate 6a, Fault C	9570	7.6	Climate 6a, Fault E	4878	13	Climate 6a, Fault D	3288	8.7	Climate 6a, Fault J	1002	4.4
Climate 4c, Fault F	7800	6.4	Climate 4c, Fault E	3192	10	Climate 3a, Fault J	2628	4.5	Climate 4c, Fault F	432	4.3
Climate 6a, Fault D	9468	6.4	Climate 6a, Fault C	4206	7.8	Climate 4c, Fault J	2490	4.4	Climate 4c, Fault J	420	4.3
Climate 4c, Fault E	7620	6	Climate 3a, Fault F	1410	7.7	Climate 4c, Fault I	2094	3.6	Climate 3a, Fault G	282	4.2
Climate 6a, Fault G	8466	5.3	Climate 3a, Fault E	1212	5.7	Climate 6a, Fault J	1944	3.4	Climate 6a, Fault H	870	4.2
Climate 3a, Fault F	4626	3.8	Climate 4c, Fault G	1908	4.1	Climate 4c, Fault G	1338	2.7	Climate 3a, Fault C	156	4.1
Climate 3a, Fault D	4134	3.6	Climate 3a, Fault J	1098	3.8	Climate 6a, Fault C	1104	2.4	Climate 3a, Fault F	324	4.1
Climate 3a, Fault E	4140	3.6	Climate 4c, Fault C	1560	3.2	Climate 4c, Fault F	522	2.1	Climate 6a, Fault G	102	4.1
Climate 3a, Fault C	4116	3.5	Climate 6a, Fault G	1842	2.6	Climate 4c, Fault A	378	2	Climate 4c, Fault B	24	4
Climate 3a, Fault J	4134	3.5	Climate 4c, Fault I	1308	2.4	Climate 4c, Fault B	294	2	Climate 4c, Fault G	750	4
Climate 4c, Fault G	5196	3.4	Climate 4c, Fault J	1200	2.3	Climate 4c, Fault D	294	2	Climate 3a, Fault A	912	4
Climate 6a, Fault B	5934	3.2	Climate 6a, Fault B	996	2.1	Climate 4c, Fault H	102	2	Climate 3a, Fault E	606	4
Climate 3a, Fault B	438	3.1	Climate 6a, Fault D	996	2.1	Climate 6a, Fault I	804	2	Climate 3a, Fault H	894	4
Climate 4c, Fault D	4068	2.8	Climate 3a, Fault C	258	2	Climate 4c, Fault E	78	1.9	Climate 3a, Fault J	888	4
Climate 4c, Fault J	4662	2.7	Climate 4c, Fault B	354	1.9	Climate 3a, Fault A	612	1.7	Climate 6a, Fault C	456	4
Climate 4c, Fault I	4404	2.6	Climate 4c, Fault D	354	1.9	Climate 3a, Fault B	144	1.7	Climate 6a, Fault F	456	4
Climate 4c, Fault B	684	2.1	Climate 3a, Fault B	126	1.9	Climate 3a, Fault D	90	1.7	Climate 4c, Fault I	600	3.9
Climate 4c, Fault C	24	2	Climate 3a, Fault D	114	1.8	Climate 3a, Fault F	516	1.7	Climate 3a, Fault B	648	3.9
Climate 6a, Fault J	1830	2	Climate 6a, Fault I	498	1.7	Climate 3a, Fault E	252	1.6	Climate 6a, Fault B	720	3.8
Climate 6a, Fault I	954	1.9	Climate 6a, Fault J	432	1.7	Climate 3a, Fault H	264	1.6	Climate 6a, Fault E	702	3.8

11.12 Appendix L: Additional Figures Describing ESTool Implementation

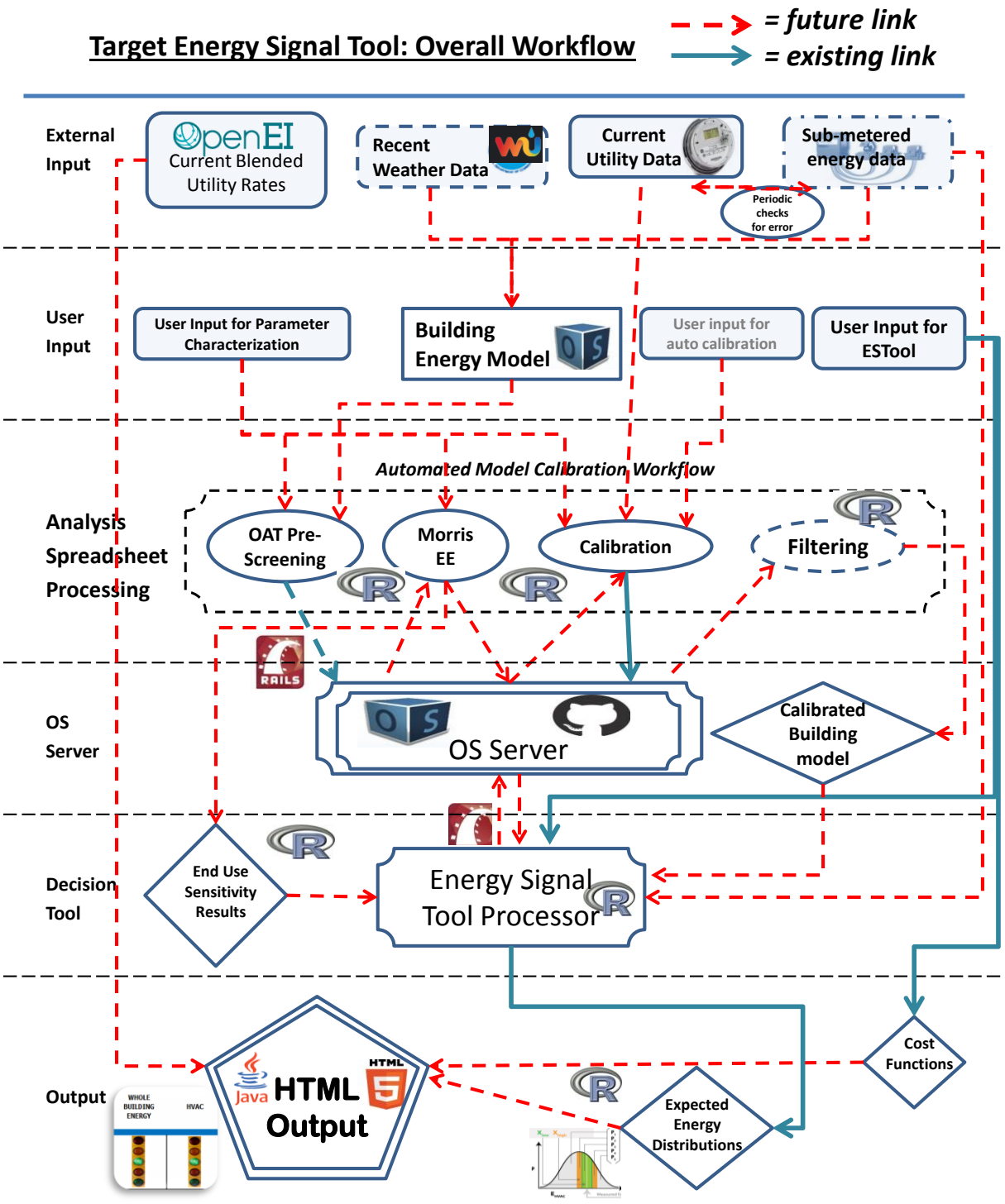


Figure 103: Present and future workflow visualization for a working Energy Signal Tool