Development of a Decision Support Tool to Test Energy Management Alarming Thresholds

Aaron Tarjan
University of Miami, aaron@tarjan.net

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_theses

Recommended Citation
https://scholarlyrepository.miami.edu/oa_theses/249
DEVELOPMENT OF A DECISION SUPPORT TOOL TO TEST ENERGY MANAGEMENT ALARMING THRESHOLDS

By

Aaron V. Tarjan

A THESIS

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Master of Science

Coral Gables, Florida

May 2011
UNIVERSITY OF MIAMI

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

DEVELOPMENT OF A DECISION SUPPORT TOOL TO TEST ENERGY MANAGEMENT ALARMING THRESHOLDS

Aaron V. Tarjan

Approved:

Shihab Asfour, Ph.D.
Professor and Chairman of Industrial Engineering

Terri A. Scandura, Ph.D.
Dean of the Graduate School

Murat Erkoc, Ph.D.
Assistant Professor of Industrial Engineering

Arzu Onar-Thomas, Ph.D.
Associate Member
Department of Biostatistics
St. Jude Children's Research Hospital
A novel model was developed to test the use of short data sets for testing various alarm thresholds as part of an energy management program. Several years of 15-minute interval data were utilized from five buildings in Jacksonville, Florida. The model aggregated the data by day type and occupancy so that there were four period types used. For all of the buildings’ meters, their daily usage by period type was tested against the threshold to determine if an alarm would be triggered, which would then be assigned a reward and cost based upon the type and duration of response. The risk management value was converted to dollars, in order to normalize the energy and time. It was determined that the 5-month short data set was the most appropriate choice for short data sets. In addition, it was concluded that the thresholds should be set between 0.8 and 1.0 standard deviation above the average of the short window. Several recommendations for further study are also enclosed.
I would like to thank my Committee – Dr. Shihab Asfour, Dr. Murat Erkoc and Dr. Arzu Onar-Thomas for assisting me in this endeavor. I’m especially grateful for the opportunity Dr. Asfour provided me by allowing to return to school and work in the Industrial Assessment Center. The training I received there was invaluable for my career. Also, the guidance from Dr. Onar-Thomas was extremely valuable as she was so able in dissecting my work and helping me patch the holes.

Along the way, a couple of professor friends provided some key guidance in helping me move the project along. Big thanks to Jan Schwartz and Sam Brody. Their simple tweaks came just at the right time. Some colleagues also proved crucial to the completion of the project. I want to thank Gary Reams, for allowing me to utilize all the data I needed. Cary Fukada proved invaluable as several lengthy discussions helped me to develop the “risk management” concept.

I can’t thank my father, Dr. Peter P. Tarjan, enough for all of his help during the final stages of this project. His commitment to excellence has been one of life’s greatest lessons.

Lastly, I want to thank Liz, my best friend, chief advisor, wife and love of my life. When I couldn’t see clearly, her patience, tolerance, strength and determination kept refocusing me.
# Table of Contents

List of Tables ....................................................................................................................... vi  
List of Figures ..................................................................................................................... vii  
List of Abbreviations ........................................................................................................... ix  
Chapter 1: Introduction ...................................................................................................... 1  
  1.1 The Business Case ............................................................................................. 1  
  1.2 The Problem ...................................................................................................... 4  
  1.3 The Study .......................................................................................................... 5  
Chapter 2: Background Information ............................................................................. 6  
  2.1 Performance Contracting .................................................................................. 6  
  2.2 The Energy Guarantee ...................................................................................... 7  
  2.3 International Performance Measurement and Verification Protocol .......... 9  
    2.3.1 Option A: Partially Measured Retrofit Isolation ...................................... 10  
    2.3.2 Option B: Retrofit Isolation ..................................................................... 11  
    2.3.3 Option C: Whole Building ........................................................................ 12  
    2.3.4 Option D: Calibrated Simulation ............................................................. 14  
  2.4 Data Diagnostics .............................................................................................. 14  
    2.4.1 Rules Engines ........................................................................................... 15  
    2.4.2 Fault Detection and Diagnostics .............................................................. 16  
  2.5 The Study ........................................................................................................ 17  
  2.6 Data Management .......................................................................................... 20  
    2.6.1 Calibration ............................................................................................... 21  
    2.6.2 Missing Data ............................................................................................ 21  
  2.7 Summary ......................................................................................................... 22  
Chapter 3: Methodology .................................................................................................. 24  
  3.1 Building Descriptions ...................................................................................... 24  
  3.2 Data Summation ............................................................................................. 25  
  3.3 The Decision Model ........................................................................................ 29  
    3.3.1 Loops ....................................................................................................... 30  
    3.3.2 Alarm Generation .................................................................................... 31  
    3.3.3 Multipliers ............................................................................................... 34  
    3.3.4 Risk Management .................................................................................... 36  
    3.3.5 Data Analysis ........................................................................................... 38  
    3.3.6 Test Runs ................................................................................................. 39  
Chapter 4: Results and Analysis ....................................................................................... 42  
  4.1 Weather Data .................................................................................................. 42
List of Tables

Table 1: Summary of Buildings ........................................................................................................ 24
Table 2: Monthly and Annual Mean Daily Temperatures ......................................................... 25
Table 3: Schedule of threshold increments .................................................................................. 30
Table 4: Sample of Weekly Occupied, 3-Month window data, along with daily
  usages by period type .............................................................................................................. 32
Table 5: Distribution for response types, and response lengths ................................................. 33
Table 6: Slopes of TTM building data for all daily energy usages over Base
  and Test period ......................................................................................................................... 64
Table 7: Correlations of Base and Test run result arrays for data averages ............................... 66
Table 8: Correlations of Base and Test run result arrays for standard
  deviations ................................................................................................................................. 66
Table 9: Weekday Occupied Test run results for the 5-month window .................................... 67
Table 10: Weekday Occupied Test run results for the 12-month window ................................. 69
List of Figures

Figure 1: Mean Daily Temperature for the Months and Years of 2006-2009 .......... 25
Figure 2: Sample of calculated data for the Weekday Occupied period type .......... 27
Figure 3: Sample of calculate data for the 1-month, 3-month and 5-month windows by period type .......................................................... 28
Figure 4: Decision model flowchart ................................................................. 29
Figure 5: Data used for the calculation of the Projected Multipliers .................... 35
Figure 6: Load Factor Multiplier data for Weekday Occupied meters ................. 36
Figure 7: Sample of risk management values for Threshold 1, Weekday Occupied, 3-month window ............................................................ 38
Figure 8: Sample of Weekday Occupied data for all 7 thresholds .................... 39
Figure 9: Base model window and test model dates for the 1-month and 3-month windows ................................................................................. 40
Figure 10: Base model window and test model dates for the 5-month and 12-month windows ..................................................................... 41
Figure 11: Weekday Occupied plotted against mean daily temperature, Building 1, Meter 1 ............................................................................. 42
Figure 12: Weekday Occupied plotted against mean daily temperature, Building 3, Meter 1 ............................................................................. 43
Figure 13: BASE run, Risk Management total, 1-month and 3-month ................ 44
Figure 14: BASE run, Risk Management total, 5-month and 12-month ............... 45
Figure 15: TEST run, Risk Management total, 1-month and 3-month ................ 46
Figure 16: TEST run, Risk Management total, 5-month and 12-month ............... 47
Figure 17: BASE run, Weekday Occupied, 1-month and 3-month .................... 48
Figure 18: BASE run, Weekday Occupied, 5-month and 12-month .................... 49
Figure 19: TEST run, Weekday Occupied, 1-month and 3-month .................... 50
Figure 20: TEST run, Weekday Occupied, 5-month and 12-month .................... 51
Figure 21: BASE run, Weekday Unoccupied, 1-month and 3-month ................. 52
Figure 22: BASE run, Weekday Unoccupied, 5-month and 12-month ................. 53
Figure 23: TEST run, Weekday Unoccupied, 1-month and 3-month ................. 54
Figure 24: TEST run, Weekday Unoccupied, 5-month and 12-month ................. 55
Figure 25: BASE run, Weekend Occupied, 1-month and 3-month .................... 56
Figure 26: BASE run, Weekend Occupied, 5-month and 12-month .................... 57
Figure 27: TEST run, Weekend Occupied, 1-month and 3-month .................... 58
Figure 28: TEST run, Weekend Occupied, 5-month and 12-month .................... 59
Figure 29: BASE run, Weekend Unoccupied, 1-month and 3-month ................. 60
Figure 30: BASE run, Weekend Unoccupied, 5-month and 12-month ................. 61
Figure 31: TEST run, Weekend Unoccupied, 1-month and 3-month .................... 62
Figure 32: TEST run, Weekend Unoccupied, 5-month and 12-month ............................. 63
Figure 33: Building 5, Meter 1, normal probability plot for the weekday occupied period ........................................................................................................................................ 65
Figure 34: Building 5, Meter 1, normal probability plot for the weekend unoccupied period ........................................................................................................................................ 65
Figure 35: Weekday Occupied Test run results for the 5-month window ......................... 68
Figure 36: Weekday Occupied Test run results for the 12-month window ....................... 69
Figure 37: Building 1, Meter 1, energy usage by period type .............................................. 78
Figure 38: Building 1, Meter 2, energy usage by period type .............................................. 80
Figure 39: Building 2, Meter 1, energy usage by period type .............................................. 81
Figure 40: Building 2, Meter 2, energy usage by period type .............................................. 82
Figure 41: Building 3, Meter 1, energy usage by period type .............................................. 83
Figure 42: Building 4, Meter 1, energy usage by period type .............................................. 84
Figure 43: Building 5, Meter 1, energy usage by period type .............................................. 85
Figure 44: Building 5, Meter 2, energy usage by period type .............................................. 86
Figure 45: Building 1, Meter 1, normal probability plots by period type ......................... 106
Figure 46: Building 1, Meter 2, normal probability plots by period type ......................... 107
Figure 47: Building 2, Meter 1, normal probability plots by period type ......................... 108
Figure 48: Building 2, Meter 2, normal probability plots by period type ......................... 109
Figure 49: Building 3, Meter 1, normal probability plots by period type ......................... 110
Figure 50: Building 4, Meter 1, normal probability plots by period type ......................... 111
Figure 51: Building 5, Meter 1, normal probability plots by period type ......................... 112
Figure 52: Building 5, Meter 2, normal probability plots by period type ......................... 113
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHU</td>
<td>air handling unit</td>
</tr>
<tr>
<td>ASHRAE</td>
<td>American Society of Heating, Refrigerating and Air-Conditioning Engineers</td>
</tr>
<tr>
<td>BAM</td>
<td>baseline adjustment multiplier</td>
</tr>
<tr>
<td>BR</td>
<td>billable rate</td>
</tr>
<tr>
<td>CV</td>
<td>constant volume</td>
</tr>
<tr>
<td>DAS</td>
<td>data acquisition system</td>
</tr>
<tr>
<td>ECM</td>
<td>energy conservation measure</td>
</tr>
<tr>
<td>ESCO</td>
<td>energy service company</td>
</tr>
<tr>
<td>FDD</td>
<td>fault detection and diagnostics</td>
</tr>
<tr>
<td>FEMP</td>
<td>Federal Energy Management Program</td>
</tr>
<tr>
<td>IPMVP</td>
<td>International Performance Measurement and Verification Protocol</td>
</tr>
<tr>
<td>LF</td>
<td>load factor</td>
</tr>
<tr>
<td>LFM</td>
<td>load factor multiplier</td>
</tr>
<tr>
<td>M&amp;V</td>
<td>measurement and verification</td>
</tr>
<tr>
<td>NG</td>
<td>natural gas</td>
</tr>
<tr>
<td>OAT</td>
<td>outside air temperature</td>
</tr>
<tr>
<td>PA</td>
<td>period average</td>
</tr>
<tr>
<td>PM</td>
<td>projected multiplier</td>
</tr>
<tr>
<td>PU</td>
<td>period usage</td>
</tr>
<tr>
<td>PV</td>
<td>projected value</td>
</tr>
<tr>
<td>RT</td>
<td>energy rate</td>
</tr>
<tr>
<td>RTU</td>
<td>Rooftop unit</td>
</tr>
<tr>
<td>SF</td>
<td>size factor</td>
</tr>
<tr>
<td>SOC</td>
<td>standard of occupancy and control</td>
</tr>
<tr>
<td>TM</td>
<td>time spent</td>
</tr>
<tr>
<td>TTM</td>
<td>twelve trailing months</td>
</tr>
<tr>
<td>US DOE</td>
<td>United States Department of Energy</td>
</tr>
<tr>
<td>VAV</td>
<td>variable air volume</td>
</tr>
<tr>
<td>VBA</td>
<td>Visual Basic for Applications</td>
</tr>
<tr>
<td>VEE</td>
<td>validation, estimating and editing</td>
</tr>
<tr>
<td>VFD</td>
<td>variable frequency drive</td>
</tr>
</tbody>
</table>

---

1. **ECM**: A set of activities designed to increase the energy efficiency of a facility (IPMVP, 2002).
2. **ESCO**: A firm which provides a range of energy efficiency and financing services and guarantees that the specified results will be achieved under an energy performance contract (IPMVP, 2002).
3. **M&V**: The process of determining savings using one of the four IPMVP Options (IPMVP, 2002).
Chapter 1: Introduction

1.1 The Business Case

The Energy Performance Contracting industry is a complex multi-billion dollar industry based upon considerable uncertainty in the data on which decisions are made. The basic model for an energy performance project is that an entity, typically a public one, originates a loan to pay an energy services company (ESCO) for capital equipment upgrades on its property. An ESCO is a firm which provides a range of energy efficiency and financing services and guarantees that the specified results will be achieved under an energy performance contract (IPMVP, 2002). The energy efficiency of the post-retrofit environment is guaranteed to generate an amount of financial savings over a given period so that the borrower, using those same savings, will be able to meet its debt service obligations. In the case that the performance of the energy savings is below the guaranteed amount, the ESCO would then make up the difference with a payment to the customer equal to the shortfall.

Much has been written about the modeling of energy savings, but most of the literature typically focuses on the pre-retrofit period and provides methods for calculating energy savings based upon those pre-existing conditions. Although an ESCO may have internal processes in place for reviewing and assessing the risk associated with a project prior to implementation, the engineering models project only best
estimates of potential performance. In other words, “when the rubber meets the road,” the assumptions in the models may fall short in the real world.

The retrofits, which are commonly referred to as energy conservation measures (ECMs), will have varying degrees of interdependence. ECMs are sets of activities designed to increase the energy efficiency of a facility (IPMVP, 2002). They could be operational changes, equipment upgrades, and improvements to maintenance practices. Accounting for the energy performance in a post-retrofit environment can be very problematic. This may be due to several factors, such as operating characteristics of the building, which were unknown to the engineering team during the development phase, changes to the physical building that may be made outside the scope of the project, modifications to the operating conditions of the building which affects the internal load, and many more. For the energy management team responsible for managing the risk associated with ten to twenty years of contingent liability, this can be extremely challenging.4

One powerful tool which can help a team with the ongoing monitoring of a building’s energy usage is a system of fault detection. This is typically accomplished using software, commonly known as a “rules engine,” which can evaluate data sets and generate alarms based upon pre-defined business rules. For example, if a building has several ECMs completed in it, and their interdependence is so great that it is impossible

4 Contingent liability is the liability resulting from the future risk associated with the possible payment due to a shortfall with the guarantee. The guarantee is reconciled annually, so the liability is therefore contingent upon the positive performance of the program in future years.
to calculate their savings individually, the best monitoring approach may be one involving the whole building using interval data that has been calibrated to the building’s utility meter to report energy usage.\(^5\) If this interval data were uploaded to a central server for analysis on a daily basis, ideally more frequently, a set of rules could be run against the data to determine the existence of any abnormal conditions. Some of these conditions might be strings of zeros, unusually high values, negative numbers and null values. If a condition considered abnormal were met, the system would then generate an alarm to notify the energy management team which might lead to possible investigation.

When dealing with an entire building which is very dynamic, the question becomes, what is the appropriate level for determining if a value is abnormal? The problem is that if a threshold is set too high, then an interval of high usage, which should be investigated, may not be detected, resulting in a lost opportunity. Conversely, if the threshold is set too low, then the alarm recipient might become numb to their arrival and ignore them, even though some percentage will be worthy of further investigation. The ideal situation is then to have the thresholds set at an appropriate point.

\(^5\) Interval data is time stamped data collected at regular intervals used to measure a commodity over time. The data could represent many things, such as kWh, power factor, voltage, gallons of water, pounds of steam, etc.
1.2 The Problem

Although the performance guarantee is based upon projected savings, those savings don’t always materialize, so it is very important to begin to monitor and track a building’s post-retrofit usage as early as possible in the first year of the performance period. This is especially critical since the building has just undergone many changes.

The pre-retrofit models used for the engineering of the projects are typically based upon at least one full year of usage data, often more. This usage period is known as the baseline period. Once the energy efficiency retrofits are complete, this baseline data is nothing more than a reference point, since the energy profile of the building will have been modified.

As mentioned previously, part of the ongoing monitoring program should include thresholds being set for the building for the alarming tool. Ideally, predicted energy usage will be based upon a model with at least one full year of data. Since the new energy usage profile for the building is going to have very little history, it is necessary to develop an improved method to determine where the thresholds should be set, so that the alarming tool would work appropriately during that first year. One of the keys to success is to have reasonable thresholds established in the beginning of the monitoring period so that potential performance issues can begin to be addressed as early as possible. Later on, once a full year’s worth of data is available, the process of determining and setting alarm thresholds can become an iterative process and should be revisited, as presumably it would be possible to make more accurate estimates.
1.3 The Study

The purpose of this study is to test a novel decision tool that was developed by the author, to evaluate how well energy usage thresholds can be estimated using short data sets.
Chapter 2: Background Information

2.1 Performance Contracting

The energy performance contracting business has its roots in the early 1970’s after the oil embargo, when the spike in energy prices drove up interest in conservation. At that time, among other concerns, such as national security, there was a realization that energy efficiency was not just about being “green,” but also about reducing the burden on constrained operating budgets that were being squeezed by higher energy costs.

Performance Contracting is a diverse industry with a multitude of contractors, who typically provide services to publicly funded entities. The range of services provided varies widely and includes energy efficiency, water conservation, and energy education. Over the years it has largely been driven by the need for public entities to upgrade capital equipment and better manage their resources.

Although the business has evolved over the years, and the specifics can vary by state and by project, the typical arrangement today is one of a performance contract. The energy performance contract is an agreement entered into between an owner/operator, which is typically a public entity, such as a city, a school district, a correctional organization, or housing authority, and an ESCO. The scope of the work will vary, but the Owner will receive equipment upgrades which increase the operating efficiency of the building. The list of ECMs often includes items such as lighting, boilers,
chillers, and building management systems. The financial savings that will be achieved by performing the retrofits are guaranteed to be great enough over the life of the project to pay for the costs of the retrofits, including the debt service on the construction loans. The guarantee component of the project differentiates these projects from capital expense efficiency projects. The third party in this arrangement is the financing institution that provides the funding to the Owner in order for the project to proceed.

2.2 The Energy Guarantee

The guarantee is the vehicle which separates a performance contract from a simple energy efficiency project. It is a separate component from the financing, and is used to reduce the risk of the owner/financier by ensuring the ability of the owner to perform on the debt service. It is based upon the financial performance of a project over its life and guarantees that the energy efficiency project will achieve a certain level of performance, enabling the borrower to perform on the debt service. If the guaranteed savings are not achieved, the shortfall would then be made up through financial remuneration from the ESCO to the owner.

Energy savings and energy cost savings are separate terms, which are usually defined within a contract, and refer to use avoidance and cost avoidance, respectively. These are important distinctions which have different impacts. For example, over the life of a project there will usually be a normal rate of degradation in the performance of the installed equipment. However, at the same time, there will also most likely be an
increase in the cost of energy over that same period. Ideally, over the long term, the two will balance each other out.

The pre-retrofit performance is considered to be the baseline period and can be thought of as a reference point for what would be consumed without the retrofit project. At its most basic level, the energy savings is calculated as the difference between the measured pre-retrofit performance of the equipment minus the actual post-retrofit usage of the equipment. This can be stated as:

\[
\text{Energy Savings} = (\text{Baseperiod Energy Use}) - (\text{Post} - \text{Retrofit Energy Use})
\]

\[\pm \text{ Adjustments}\]

The energy savings, or use avoidance, is then multiplied by the established energy unit rate to calculate the cost avoidance. Assumptions about how the equipment would have operated are part of the original energy savings modeling, and can include variables such as operating schedule, occupancy patterns, weather, and electrical loads. A change in any of these factors may create a need to adjust the baseline. Adjustments to the baseline are therefore made in order to standardize the two time periods to a similar set of conditions. Stated another way, a baseline adjustment is used to account for what the usage of the building would have been if the retrofit work had not taken place but the same change in operation had been made.
2.3 International Performance Measurement and Verification Protocol

The Energy Policy Act of 1992 mandated that the Federal Government achieve a certain level of efficiency in its energy utilization (EPAct 1992). In 1994, the U.S. Department of Energy, tasked with determining how to manage the large scope of this work began to lead the effort to reduce the uncertainty in the performance contracting business by bringing together many business partners and industry experts throughout North America. Two years later, the first edition of the collaborative work was released. This is known as the International Performance Measurement and Verification Protocol (IPMVP). Today, the U.S. government is the largest consumer of energy performance contracting work.

A major factor in the drive to create a framework was the need for one standard for the three parties - the owner, the ESCO and the financial institution - to utilize when negotiating details of a project. Some of the questions raised were: How would the owner and financer be certain that the projected savings would be achieved? What certainty would the financer have that the borrower would be able to generate the savings stream to repay the debt? How could the ESCO reduce the owner’s and lender’s uncertainty with their measurement and verification (M&V) plan, when there was no standard language or framework for defining and outlining an M&V Plan?

The development of a framework within which lenders, buyers and sellers could communicate helped to increase understanding among the parties and effectively reduce the risk and level of uncertainty being assumed by the parties. In the years that followed the release of the first edition, the IPMVP became the de facto reference for
outlining program savings. It has been translated into several languages and is used around the world. Although some other, more prescriptive M&V documents have been developed, such as the US Department of Energy’s Federal Energy Management Program (FEMP) and ASHRAE’s Guideline 14P, the basis of their contents comes from the IPMVP, maintaining the Protocol as a key reference for M&V.

When verifying the performance of an ECM, there are different approaches to measuring the energy savings. Each method will have different levels of accuracy and cost, and those two variables must be weighed against the amount of risk associated with the ECM in order to help determine the best approach. The methods would be laid out in the M&V Plan and agreed upon by all parties. The IPMVP therefore creates a common language to clarify expectations. The protocol defines four M&V Options.

2.3.1 **Option A: Partially Measured Retrofit Isolation**

For Option A, there is only a partial measurement of the ECM, with one or more of the parameters being stipulated. The agreed upon assumptions, or stipulations, would typically be well known and will be such that their impact on the ECMs savings will not be significant. In addition, the stipulations should be parameters that are outside the control of the ESCO and are assumed to be relatively constant. The measured variables would be those that are important to determine the performance of the ESCO’s work on the ECM. For example, if the reduced load comes from a lighting retrofit, the operating hours might be stipulated since they are controlled by the owner. However, the fixture count, and the loads coming from the ballasts and lamps would be measured.
One of the benefits in utilizing an Option A plan is that it reduces the overall cost of verifying the savings. Typically, the stipulated parameter will be one which is much less costly to obtain than the measured variable.

### 2.3.2 Option B: Retrofit Isolation

This Option is typically much more complicated than Option A. With Option B, there is ongoing measurement of the retrofit with all of the ECM’s parameters being measured. Since there are no stipulated parameters, Option B will typically be more costly than Option A. However, the additional cost of full measurement may be important when there is a higher degree of uncertainty in the potential savings as a result of the greater variance among the parameters.

For example, Option B might be preferable in the case of a photovoltaic installation, where there is net metering in place, and the impact of weather on the power generation is significant. The isolation of a chiller plant presents another example. This equipment will usually have its load vary throughout the day and year, and is dependent on many variables such as space occupancy and outside air temperatures.

There are several reasons why the partners may choose to use Option B. One is that the interaction of the ECM with the rest of the building may not be important. In addition, sub-metering may be necessary to guarantee state and utility rebates. Lastly, the potential savings from the single ECM are not great enough in proportion to the whole building’s usage to be a part of an Option C program.
2.3.3 Option C: Whole Building

This approach uses the metering of a whole building to determine the collective performance of the ECMs. Although there may be sub-meters within the building to isolate the performance of specific systems and help with the ongoing monitoring, the performance of individual ECMs is not reported; only the performance of the whole building is reported. Any changes to the building’s operating characteristics would need to be evaluated to make necessary adjustments to the energy baseline; that is whether they impact the baseline positively or negatively. Some of the building parameters to track would include issues such as operating schedule, occupancy, electrical plug load, and HVAC set-points and scheduling.

Option C would be used when the potential energy savings are significant enough compared to the energy consumption of the entire building. A general rule of thumb is that the energy savings should be, at the very least, 10% of the baseline energy usage. Option C might also be employed if there is considerable interaction between the boundaries of the individual systems. An example of this might be a building where multiple ECMs were performed, such as the installation of lighting, a new high efficiency chiller, variable air volume boxes, high efficiency water pump motors with variable frequency drives (VFD), and lastly, an energy management system to control operating set-points and schedules. In this case, although the energy savings may have been calculated and assessed separately for each ECM as part of the overall project, their interactions would be far too great to report on individually. Lastly, Option C may also
be employed if there are too many ECMs within a building such that reporting on them independently would not be cost effective.

The reporting of energy data can be performed by using interval data collected from a pulse relay connected to the utility’s energy meter, or with data from the utility bill. Each of these has benefits and disadvantages. When receiving a pulse from the utility’s meter, there is potential for a higher level of confidence in the reported performance of the building as there will be a lot of interval data which the ESCO and owner/operator can use for analysis. However, no system is perfect, and in the event of missing data, there is difficulty in reporting usage for time periods which do not overlap the meter’s service period. Missing data would then need to be estimated or interpolated through a predetermined procedure. This procedure would need to be explained in detail in the M&V Plan. In addition, if variables, such as peak demand or power factor are being reported as part of the savings program, it is imperative that the data be collected in the same manner that the utility captures them.

When using the utility bill data, there is an assumption that the bill data are accurate. Sometimes bills have estimates, which may not be clearly stated on the bill. For an individual month, the impact may be great, however, over time this would be corrected once the meter is actually read, and then actual usage could be reported by backfilling the missing data based upon meter reads. One of the biggest disadvantages of not having interval data and relying on the few data points provided by the utility’s bill is that there may be great uncertainty in how the building is actually performing. For
example, without interval data, there is no way to graphically view the building’s load to help determine if the characteristics of the load may have changed over time.

Independent variables, such as weather, non-ECM equipment load, occupancy and operating hours need to be specified in the M&V Plan. These factors will be critical to determining future baseline adjustments. Often these variables can be modeled using statistical methods such as regression analysis.

2.3.4 Option D: Calibrated Simulation

With Option D, the building or the ECM is modeled using computer simulation software to calculate the impact on energy usage for either the baseline period or the post-retrofit periods. Typically, the models employ hourly calculation techniques. In order to ensure that the estimates provided by the modeling are reasonable, there must be some method for calibrating the model for testing. The simulation can be used to test the individual ECMs or for all of them combined.

One of the keys to successful modeling is the user’s understanding of the software being employed. In addition, the adage, “garbage in, garbage out” applies here. The quality of the input data is critical and must be thoroughly reviewed. Any modifications or corrective procedures performed on the data should be clearly stated as part of the modeling process.

2.4 Data Diagnostics

Collecting energy interval data from utility meters or trend data from an EMS can lead to enormous quantities of data. When this amount is multiplied across hundreds of
meters along with a multitude of EMS systems, each with hundreds of points, the volume of data can quickly become too cumbersome to manage, visualize and analyze without some form of automation. A lack of automation causes the data’s value to diminish as they are rendered useless due to sheer magnitude.

2.4.1 Rules Engines

In recent years, a new class of software, called a rules engine, has been developed for business applications (Chisholm, 2004). This type of software has simplified the software development process by having a programming team tasked with only creating a platform, which can then have sets of business rules created and modified by an end user such as a business analyst (Chisholm, 2004). By removing the rules from within the software code and out of the control of IT professionals, it can improve the speed of deployment, enable rapid updates to the rules, and increase the flexibility of the software for the end user. A by-product of the rules is that based upon conditions being met, workflows can then be created. For example, if the engine discovers that a condition is met, such as high electrical usage in a building, an alarm might be sent via email to notify the appropriate user.

Rules engines may be confused with data mining and artificial intelligence tools, but are actually quite different. The former would be formed around known parameters defined by business practices, while the latter two are looking to extract patterns from large data sets or generate inferences without the engine necessarily posing a specific question (Chisholm, 2004).
2.4.2 Fault Detection and Diagnostics

Within the energy management industry, the application of these tools is commonly known as fault detection and diagnostics (FDD). A system fault is typically defined when a system, or one of its components, operates outside its design range. This may be caused by hardware failure, maintenance troubles, system programming errors or a number of other things. Some cases where a fault detection rule might be applied are:

- verifying building zone temperatures during an unoccupied state to confirm a building’s inside temperature is drifting to its unoccupied set point;
- checking the energy usage of an air handling unit (AHU) during an occupied period while the outside air temperature is below a certain range to verify that the system is cycling appropriately;
- ensuring that the flow rate in a steam condensate return meter is high enough to satisfy a minimum rate of steam loss within a distribution system.

An extensive study performed by the U.S. Department of Energy (US DOE, 2005) estimated that approximately 11% of the annual energy consumed by lighting, HVAC and refrigeration systems in large commercial buildings was a direct result of systems faults. This amount is quite significant and is projected to be equivalent to about one quad of energy.\(^6\) Although they analyzed over a dozen types of faults, they found that three faults alone – duct leakage, HVAC left on when space unoccupied, and lights being

\(^6\) A quad is shorthand for an amount of energy equal to a quadrillion or \(10^{15}\) BTU.
left on when space unoccupied – accounted for approximately 68% of that figure (DOE, 2005). These and other issues can go undetected for long periods of time, which makes their detection such an important part of the management of the energy performance guarantee. Although this study does not analyze these types of faults directly, with the first level of alarming that is being reviewed here, responding to those alarms more adequately is critical in facilitating fault detection.

2.5 The Study

One ESCO with a national market position utilizes a rules engine to test thresholds. However, the ESCO has a difficult time setting the thresholds at appropriate levels. The process currently involves not much more than best guesses, which are modified over time. They have developed their own rules engine software tool for performing a wide array of diagnostics called Facility Diagnostics. Although, they also collect trend data from various EMS systems and weather stations, the data, which are of primary interest, are the energy interval data received from all over the country by way of either a direct pulse from utility meters or meters installed by the ESCO. The greatest number of these points comes from electric meters, but depending upon the specific needs of the M&V plan, they may also include natural gas, propane and water meters.

For the ongoing monitoring of Option C sites, one of the most common diagnostic rules which the company’s energy management team has implemented for existing projects is a rule which checks if energy usage or demand has exceeded a pre-
specified threshold. This initial alarm provides opportunity for additional investigation.

A key factor to the successful implementation of this rule is that the thresholds must be set to appropriate levels. If the threshold for alarming is too low, then the people involved in support of the energy monitoring might begin to ignore the notifications. Conversely, if the threshold is too high, then abnormal usage patterns which should prompt further investigation would be missed. Therefore, finding the appropriate threshold level is critically important.

The equipment used to collect the interval data is typically installed during the energy project’s implementation period alongside all of the other energy conservation measures. A consequence of this is that the energy calculations performed during the development phase most often are not based upon any interval data, and almost all of the baseline data comes solely from utility bill data. Once one to two years of interval data have been collected and the energy manager has a strong working knowledge of the building’s operation, it is then obviously much easier to establish limits for the alarm thresholds by analyzing available historical data. Absent this, a key question then is how best to set the threshold levels on a new project, which has no more than a few months of regular interval data. To begin to try and answer this, the IPMVP provides the following statement:

“For buildings, one or more full years of energy use and weather data should be used to construct regression models. Shorter periods introduce more uncertainty through not having data on all operating modes. The best predictors of both cooling and heating annual energy use are models from data sets with mean temperatures close to the annual mean temperature. The range of variation of daily temperature values in the data set seems to be of secondary importance. One month data sets in spring and fall, when the above condition applies, can be
better predictors of annual energy use than five month data sets from winter and summer.” (IPMVP, 2002)

This statement came from a body of work by researchers at the Texas A&M Energy Systems Laboratory (Kissock, Reddy, Fletcher, and Claridge, 1993). Utilizing parts of Kissock’s procedure, the question becomes: Can this basic assumption be extended to other buildings, whose operating conditions differ from those used in the original study?

The most important component of this study that comes from the Kissock study is the use of short data sets. Short data sets are defined as 1-month, 3-month and 5-month sets of continuous data. Each of the data sets was compared to the annual usage.

There are several parameters for this study which differ from Kissock’s. The buildings used in the original study had constant volume (CV) air distribution, while the data in this study was collected from buildings in Jacksonville which all had variable air volume (VAV) boxes. In the original study, the analysis was performed by only looking specifically at heating and cooling energy use and removing the internal load of the building from their model. In this study, the whole building’s usage is utilized for the prediction method. While all of the heating in the original study was from natural gas, most of the heating and cooling needs for the buildings in this study are provided by electric power from the local utility. Lastly, Kissock did not segment the energy data. For this study, the data are segmented into four groups, based upon day type and occupancy status. These time periods are then: weekday occupied, weekday unoccupied, weekend occupied and weekend unoccupied.
2.6 Data Management

For all of the electric meters, time stamped 15-minute interval data are collected for calculation of energy savings. The data come from pulse initiators that are connected to the meter and are installed by the local utility. This relay sends a pulse based upon a fixed usage that is collected by a data acquisition system (DAS). The size of the pulse is predetermined, and is then sent to a central server for storage, analysis and reporting.

The quality of the data is an issue and involves an ongoing process. The first step is to confirm that the data are accurate through calibration. Using meter read dates and estimated meter read times, it is possible to sum up the interval data to get a value that can be tested against the utility’s bill. The ESCO that uses the data has an established procedure that, for reporting purposes, uses an acceptable error rate of +/-2%. The data are also re-commissioned regularly, typically at least once a year, to ensure that the pulse multiplier is accurate.7

On an ongoing basis the data are checked through a process that includes validation, estimating and editing (VEE). This is a set of requirements that verify the accuracy of interval data to make corrections on a frequency, ranging from real-time to daily. VEE is used to ensure data have been generated according to specifications, satisfy acceptance criteria, and are appropriate and consistent with their intended use.

7 The pulse multiplier is the predetermined value, depending on the utility meter that is assigned to a pulse.
2.6.1 Calibration

A primary concern with metered data is the accuracy of the readings. As mentioned above, the ESCO which collected the data utilized in this study has a method for calibrating the data, where the standard is to achieve readings within +/- 2% of the utility’s bill. The key factor impacting the calibration process is that it is being calibrated to the utility’s bill. This has a few main issues. The first is the accuracy of the utility’s own meters. The local utility, where the buildings are located for this study, states in its tariff sheets that they guarantee accuracy to within +/- 2%. Due to the inability to test the meters, the assumption is that the readings provided by the utility are indeed accurate. The second issue is that if a meter is being read in person, there is no way to know the exact time of day that the meter read is taking place. To provide a best guess, the estimate used for testing purposes is that the read occurs at noon. With a few to several months of data to calibrate, the slight variation will wash out over time. A third issue comes from estimated reads. Although these may occur, they are typically corrected in a month or two, once an actual read finally occurs. Again, these errors will rectify themselves with time.

2.6.2 Missing Data

With any network, there are multiple points of failure. When part of the network fails, the impact to the data quality can vary and will cause varying degrees of data loss. One common occurrence is that a meter will get replaced. Sometimes when

---

this happens, the wires providing the pulse from the meter’s contacts may not be
terminated correctly, or not at all. Another is that there is a transmission failure, which
causes data to be lost. This could be with a cable on site carrying a signal that may be
cut, a wireless radio failing, or a network issue, where the local IT team has modified its
network schema, rendering communication from the remote to the central server
impossible. Ideally, this is discovered within no more than a few days, and addressed
rapidly. However, the result is that the data on the central server during this down-time
are either non-existent, show up as a large spike or end up simply being a string of
zeros. In either case, the data for the days in question would need to be corrected.

The two most common occurrences are blank periods and usage spikes. With
blank periods, the data have to be filled in using an estimate. If the gap was less than a
month, then usage from the week or two immediately before and after was copied and
used to fill in the blank space. If the length was greater than a month, then usage from
the previous year in the same time frame was copied over. For the case of data spikes,
a process of linear interpolation is used. This simply counts the number of missing
intervals and divides the data spike by that number, effectively spreading the usage
uniformly over the time period.

2.7 Summary

The purpose of this thesis is to develop a procedure, using similar methodology
to the Kissock study which could be used to test alarming thresholds on interval data.
The model created uses Excel spreadsheets with Visual Basic for Applications (VBA)
programming. Although the basis of the Excel model created here differs from the original work by Kissock, et al., 1993, this may find many applications in the realm of decision making with threshold alarms.
Chapter 3: Methodology

3.1 Building Descriptions

The set of buildings being used for the study are five government buildings located in the downtown area of Jacksonville, Florida. Four of them have operating schedules and occupancy patterns which are fairly regular and consistent with energy load factors that are relatively stable throughout the year. The fifth building is a convention center and has a much more varied and unpredictable utilization which results in a much lower and more varied load factor.

Three of the buildings have their cooling needs met by chilled water, while the exterior zones of the building have VAV boxes that are fan powered with electric heating elements. The fourth building also has chillers providing chilled water for cooling, while two natural gas (NG) fired boilers provide hot water. Most years, the boilers operate less than 30 days per year. The last building has eight package rooftop units (RTU), which are also equipped with natural gas fired furnaces. The five buildings have a total of eight electric meters which are all being used in the study.

Table 1: Summary of Buildings

<table>
<thead>
<tr>
<th>Building</th>
<th>Cooling</th>
<th>Heating</th>
<th># of Meters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building 1</td>
<td>Chilled water</td>
<td>Electric heating element in VAV</td>
<td>2</td>
</tr>
<tr>
<td>Building 2</td>
<td>Chilled water</td>
<td>Electric heating element in VAV</td>
<td>2</td>
</tr>
<tr>
<td>Building 3</td>
<td>Chilled water</td>
<td>Electric heating element in VAV</td>
<td>1</td>
</tr>
<tr>
<td>Building 4</td>
<td>Chilled water</td>
<td>Hot water from NG boilers</td>
<td>1</td>
</tr>
<tr>
<td>Building 5</td>
<td>Rooftop package units</td>
<td>Natural gas furnaces in RTU</td>
<td>2</td>
</tr>
</tbody>
</table>
In a normal year, the heating degree days are 1,354 per year, while the cooling degree days are 2,627. For 2006 through 2009, the absolute deviation from the annual average temperature for the shoulder months of April and October were -2 and 2, respectively. These data can be seen in Table 2 and Figure 1.

Table 2: Monthly and Annual Mean Daily Temperatures

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly</td>
<td>54</td>
<td>54</td>
<td>61</td>
<td>67</td>
<td>74</td>
<td>80</td>
<td>82</td>
<td>83</td>
<td>79</td>
<td>71</td>
<td>61</td>
<td>58</td>
</tr>
<tr>
<td>Annual</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
</tr>
<tr>
<td>Difference</td>
<td>(15)</td>
<td>(14)</td>
<td>(8)</td>
<td>(2)</td>
<td>5</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>10</td>
<td>2</td>
<td>(8)</td>
<td>(11)</td>
</tr>
</tbody>
</table>

![Mean Daily Temperature, 2006-2009](image)

Figure 1: Mean Daily Temperature for the Months and Years of 2006-2009

3.2 Data Summation

Once the data were validated and any necessary corrections were made as discussed in Section 2.6, the interval data needed to be summed up into daily values so

---

9 Weather data from Jacksonville International Airport, NOAA Weather Station 13889
they could be used in the decision model. This was done by developing a tool in Excel which would aggregate the data using VBA. (See Appendix B.)

The interval data were summed up by day according to period type. The period types were based upon building occupancy and day type. The resulting four period types were:

- Weekday Occupied
- Weekday Unoccupied
- Weekend Occupied
- Weekend Unoccupied

The primary reason for wanting to study these four periods individually is that in a performance guarantee with ongoing Option C whole building measurement, there is a contractual document called the “Standards of Occupancy and Control” (SOC). This contract schedule is a table that states how the building is supposed to be operated in regard to HVAC operating schedules and temperature set-points. The values of these parameters are the same as those used for the post-retrofit period in the original energy savings calculations. Therefore, it is very important to break out the building’s energy profiles for the four period types. All of the occupancy schedules used in the Excel tool came from the SOC.

In addition to calculating the sum of the interval data for the daily period types, other values were calculated, such as the number of data points, the average for the interval data, the minimum and maximum values, plus the standard deviation. These were calculated on a daily basis as seen in Figure 2.
Figure 2: Sample of calculated data for the Weekday Occupied period type.

The data were then calculated for the 1, 3 and 5-month windows as shown in Figure 3.
Figure 3: Sample of calculate data for the 1-month, 3-month and 5-month windows by period type.
3.3 The Decision Model

A model was developed using Excel and VBA that calculates a “risk management” value for the energy program. (See Appendix C.) This value is based upon the alarms generated for all of the meters over a period of time along with the response to those alarms. This was done by having the predetermined values for the different sized short data set used as the threshold for triggering alarms on an entire year’s worth of data. A flowchart of the model follows:

![Decision model flowchart](image)

Figure 4: Decision model flowchart
3.3.1 Loops

The first loop was for testing the size of the thresholds. This loop has seven intervals which begin at 1 standard deviation and are uniformly distributed by an absolute change in probability with a maximum value of 2 standard deviations. Table 3 shows the thresholds used for the 7 increments.

<table>
<thead>
<tr>
<th>Increment</th>
<th>Probability Threshold</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8413</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>0.8640</td>
<td>1.098</td>
</tr>
<tr>
<td>3</td>
<td>0.8866</td>
<td>1.209</td>
</tr>
<tr>
<td>4</td>
<td>0.9093</td>
<td>1.336</td>
</tr>
<tr>
<td>5</td>
<td>0.9319</td>
<td>1.490</td>
</tr>
<tr>
<td>6</td>
<td>0.9546</td>
<td>1.691</td>
</tr>
<tr>
<td>7</td>
<td>0.9772</td>
<td>2.000</td>
</tr>
</tbody>
</table>

At the beginning of each increase in threshold the random number generator was reset with a seed of the same value. This was done in order to ensure that during the following three loops, for all of the trials, as the data are tested against the different window sizes and months, the random numbers used to determine the generation of alarms and the responses to those alarms remain the same.

The second loop was for testing the specified number of trials. For this study, all of the runs had 250 trials. The third loop was for the four window sizes. As mentioned previously, those sizes are 1-month, 3-month, 5-month, and 12-months. The last loop was for the individual month of the year being tested.
3.3.2 Alarm Generation

In order to generate alarms, every day’s four period type usages were tested against the daily threshold for each meter in the month being tested. Using the probability thresholds designated in the previous section, the daily threshold for each meter and period type was calculated using the Excel NORMINV function with the interval average and standard deviation for the respective period type. This number was then multiplied by the number of intervals in the period type to arrive at the threshold. If any of the daily period values being tested was greater than its respective threshold, then an alarm was generated.

The data shown in Table 4, provides a sampling of the data points used as inputs in the decision model for the 3-month window. The daily thresholds shown at the top of Table 4 were for the month of December. All of the thresholds were calculated using the Excel NORMINV function.
The next step was to determine the response to the alarm. This was done by assigning a random probability to the alarm which was used to determine the response type. The response type had four possibilities, which were to

- ignore the alarm,
- make a cursory response,
• perform an extended response, and

• have a response which ended up with a baseline adjustment.  

The distribution for determining the response type can be seen in Table 5. Using a second random probability, the duration of the response was calculated using a pre-defined distribution that was assigned to the response type. The distribution of response durations can also be found in Table 5.

Table 5: Distribution for response types, and response lengths

<table>
<thead>
<tr>
<th>Alarm Response</th>
<th>Value</th>
<th>Upper Limit Value</th>
<th>Length Min</th>
<th>Length Max</th>
<th>Distribution</th>
<th>Mean</th>
<th>St Dev</th>
<th>Mass Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Alarm</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ignore</td>
<td>1</td>
<td>50.0%</td>
<td>1</td>
<td>2</td>
<td>Uniform</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cursory</td>
<td>2</td>
<td>80.0%</td>
<td>2</td>
<td>15</td>
<td>Normal</td>
<td>8.5</td>
<td>2</td>
<td>99.899%</td>
</tr>
<tr>
<td>Extended</td>
<td>3</td>
<td>99.0%</td>
<td>2</td>
<td>15</td>
<td>Normal</td>
<td>8.0</td>
<td>2</td>
<td>99.996%</td>
</tr>
<tr>
<td>Baseline Adj</td>
<td>4</td>
<td>100.0%</td>
<td>15</td>
<td>75</td>
<td>Normal</td>
<td>45</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Baseline Adj Multiplier</td>
<td>3</td>
<td>15</td>
<td>Normal</td>
<td>9</td>
<td>3</td>
<td>98.200%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold Limits</td>
<td></td>
<td></td>
<td>1.00</td>
<td>2.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since the current system for generating the alarms is riddled with many operational challenges, it was impossible to review historical rates of response. Therefore, in order to arrive at the distributions listed in Table 5, several subject matter experts were consulted. These people were professional peers of the author who work with different types of alarms and have broad experience in the energy management field. The numbers used are best estimates of what the rates of response and consequent time spent should be.

---

10 A baseline adjustment is an accounting adjustment made to the energy baseline which accounts for a change as if the addition of or decrease to the building’s usage had occurred during the baseline period. One example might be additional equipment. The additional usage would be added to the baseline so that the energy could be accounted for without affecting the net savings that are part of the energy program.
3.3.3 Multipliers

There were three multipliers developed which were used to assist in calculating the risk management value. The first multiplier was the Baseline Adjustment Multiplier (BAM). The BAM was established in order to create additional value for the times when an extended alarm turns into a baseline adjustment, which results in reduced risk in managing the energy guarantee. The BAM was based upon a random probability and was calculated using the distribution shown in Table 5.

The second was the projected multiplier (PM). The intent of this multiplier was to provide weight to meters where there is a larger portion of the energy guarantee at risk. The PM was based upon the projected savings for each building and is independent of the period type. The percent of projected value is,

\[
%PV = \frac{\sum \text{Projected Meter Savings}}{\sum \text{Projected Program Savings}} \times 100
\]

and

\[
PM = \left( \frac{1}{1 - \frac{\%PV}{100}} \right)^{SF}
\]

where,

\[
SF \quad \text{Size factor}
\]

The size factor was set so that the meter with the greatest amount of projected savings would have a value of 3.00. All the other meters would then be relative to that one. In
In this study, the size factor was equal to 1.94. The data used in the calculation of the multiplier can be seen in Figure 5.

**Projected Multipliers**

<table>
<thead>
<tr>
<th>Projected Multiplier</th>
<th>Meter1-1</th>
<th>Meter1-2</th>
<th>Meter2-1</th>
<th>Meter2-2</th>
<th>Meter3-1</th>
<th>Meter3-2</th>
<th>Meter4-1</th>
<th>Meter4-2</th>
<th>Meter5-1</th>
<th>Meter5-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-month % Projected</td>
<td>1.99%</td>
<td>0.92%</td>
<td>12.89%</td>
<td>1.65%</td>
<td>43.14%</td>
<td>17.57%</td>
<td>19.36%</td>
<td>2.99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Projected Multiplier</td>
<td>1.032</td>
<td>1.018</td>
<td>1.239</td>
<td>1.037</td>
<td>3.000</td>
<td>1.457</td>
<td>1.520</td>
<td>1.061</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Projected Savings**

<table>
<thead>
<tr>
<th>All Blgs</th>
<th>Meter1-1</th>
<th>Meter1-2</th>
<th>Meter2-1</th>
<th>Meter2-2</th>
<th>Meter3-1</th>
<th>Meter3-2</th>
<th>Meter4-1</th>
<th>Meter4-2</th>
<th>Meter5-1</th>
<th>Meter5-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov-06</td>
<td>557,901</td>
<td>8,365</td>
<td>3,765</td>
<td>43,673</td>
<td>7,147</td>
<td>299,762</td>
<td>72,061</td>
<td>108,190</td>
<td>16,798</td>
<td></td>
</tr>
<tr>
<td>Dec-06</td>
<td>585,606</td>
<td>7,526</td>
<td>4,027</td>
<td>32,326</td>
<td>7,341</td>
<td>351,574</td>
<td>71,745</td>
<td>95,353</td>
<td>15,712</td>
<td></td>
</tr>
<tr>
<td>Jan-07</td>
<td>606,483</td>
<td>7,228</td>
<td>4,332</td>
<td>32,966</td>
<td>6,809</td>
<td>364,477</td>
<td>86,005</td>
<td>92,982</td>
<td>17,682</td>
<td></td>
</tr>
<tr>
<td>Feb-07</td>
<td>538,886</td>
<td>7,191</td>
<td>4,442</td>
<td>32,044</td>
<td>5,639</td>
<td>328,011</td>
<td>56,828</td>
<td>87,985</td>
<td>17,038</td>
<td></td>
</tr>
<tr>
<td>Mar-07</td>
<td>570,842</td>
<td>6,036</td>
<td>4,229</td>
<td>46,625</td>
<td>7,743</td>
<td>306,031</td>
<td>74,125</td>
<td>104,773</td>
<td>19,080</td>
<td></td>
</tr>
<tr>
<td>Apr-07</td>
<td>523,428</td>
<td>6,253</td>
<td>5,250</td>
<td>68,925</td>
<td>10,481</td>
<td>200,910</td>
<td>92,082</td>
<td>118,005</td>
<td>19,522</td>
<td></td>
</tr>
<tr>
<td>May-07</td>
<td>536,299</td>
<td>5,072</td>
<td>5,875</td>
<td>90,186</td>
<td>11,528</td>
<td>172,510</td>
<td>113,261</td>
<td>116,560</td>
<td>17,307</td>
<td></td>
</tr>
<tr>
<td>Jun-07</td>
<td>536,666</td>
<td>10,165</td>
<td>5,754</td>
<td>102,005</td>
<td>11,837</td>
<td>158,010</td>
<td>125,470</td>
<td>108,484</td>
<td>14,941</td>
<td></td>
</tr>
<tr>
<td>Jul-07</td>
<td>521,960</td>
<td>10,209</td>
<td>6,349</td>
<td>106,617</td>
<td>12,532</td>
<td>144,626</td>
<td>127,005</td>
<td>99,595</td>
<td>13,027</td>
<td></td>
</tr>
<tr>
<td>Aug-07</td>
<td>533,530</td>
<td>10,090</td>
<td>6,202</td>
<td>105,459</td>
<td>14,780</td>
<td>150,357</td>
<td>127,988</td>
<td>105,149</td>
<td>13,490</td>
<td></td>
</tr>
<tr>
<td>Sep-07</td>
<td>535,279</td>
<td>9,480</td>
<td>6,076</td>
<td>95,743</td>
<td>13,122</td>
<td>163,866</td>
<td>116,657</td>
<td>112,289</td>
<td>14,666</td>
<td></td>
</tr>
<tr>
<td>Oct-07</td>
<td>556,466</td>
<td>9,131</td>
<td>4,466</td>
<td>72,091</td>
<td>13,087</td>
<td>208,280</td>
<td>106,132</td>
<td>131,159</td>
<td>16,122</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5: Data used for the calculation of the Projected Multipliers**

The last multiplier was the Load Factor Multiplier (LFM). This multiplier was created with the goal of escalating the value of an alarm from a building with a lower load factor since a lower value is indicative of a greater degree of variability in the building’s usage.

The load factor (LF) of a building for a month is calculated as:

$$LF = \frac{\sum kWh}{Peak\ kW * 24 * (days\ in\ month)}$$

A building’s load factor is typically calculated using an entire month of data. However, for this study, the energy usages and peak demands for every meter were broken out for each month by period type. This then allowed for the calculation of a load factor for each meter’s period type by month.
\[ LF_{mp} = \frac{\sum kWh_{mp}}{\text{Peak } kW_{mp} \times \text{number of hours in month}_p} \]

where,

| \( m \) | Month |
| \( p \) | Period Type |

Load factor is a value between 0 and 1. In order to convert it to a number which would provide the desired effect, the load factor was inverted so as to arrive at the LFM.

\[ LFM = \frac{1}{LF} \]

Sample data of the load factor multipliers for one period type can be seen in Figure 6.

<table>
<thead>
<tr>
<th>Load Factor Multiplier</th>
<th>Weekday Occupied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Meter1-1</td>
</tr>
<tr>
<td>Nov-06</td>
<td>1.093</td>
</tr>
<tr>
<td>Dec-06</td>
<td>1.093</td>
</tr>
<tr>
<td>Jan-07</td>
<td>1.084</td>
</tr>
<tr>
<td>Feb-07</td>
<td>1.085</td>
</tr>
<tr>
<td>Mar-07</td>
<td>1.094</td>
</tr>
<tr>
<td>Apr-07</td>
<td>1.122</td>
</tr>
<tr>
<td>May-07</td>
<td>1.101</td>
</tr>
<tr>
<td>Jun-07</td>
<td>1.203</td>
</tr>
<tr>
<td>Jul-07</td>
<td>1.128</td>
</tr>
<tr>
<td>Aug-07</td>
<td>1.219</td>
</tr>
<tr>
<td>Sep-07</td>
<td>1.206</td>
</tr>
<tr>
<td>Oct-07</td>
<td>1.130</td>
</tr>
</tbody>
</table>

Figure 6: Load Factor Multiplier data for Weekday Occupied meters

### 3.3.4 Risk Management

In order to calculate the “risk management” value, alarms were assigned benefits and costs based upon the response type. The formulas for calculating the various production modes for an alarm are listed in the table below. Depending upon
the assigned response, an associated cost and benefit was calculated. Benefits are associated with an alarm being reviewed, while the cost associated with any investigation was equal to the billable time spent on the alarm. However, in the case of ignoring an alarm, while there was no possible reward, the cost was not measured in time, but in lost opportunity. Therefore, the cost of ignoring an alarm was set as equal to the reward of a cursory investigation.

\[ RM = \sum \text{Reward} - \sum \text{Cost} \]

The formulas for calculating the financial impact were:

<table>
<thead>
<tr>
<th>Production mode</th>
<th>Reward</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignore Alarm</td>
<td>0</td>
<td>((PU-PA)\times RT\times LFM\times PM)</td>
</tr>
<tr>
<td>Cursory Investigation</td>
<td>((PU-PA)\times RT\times LFM\times PM)</td>
<td>(TM\times BR)</td>
</tr>
<tr>
<td>Extended Investigation</td>
<td>(PU\times RT\times LFM\times PM)</td>
<td>(TM\times BR)</td>
</tr>
<tr>
<td>Baseline Adjustment</td>
<td>(PU\times RT\times LFM\times PM\times BAM)</td>
<td>(TM\times BR)</td>
</tr>
</tbody>
</table>

where,

\[ PU \] Period usage
\[ PA \] Period average
\[ RT \] Energy rate ($/kWh)\(^{11}\)
\[ TM \] Time spent on investigation
\[ BR \] Billable rate ($/minute = $1.25)
\[ LFM \] Load Factor Multiplier
\[ PM \] Projected Multiplier
\[ BAM \] Baseline adjustment multiplier

\(^{11}\) The energy rates used are the contractual base energy rates and were not escalated.
An example of the output data appears in Figure 7. The first five trials are shown here, however, data from all of the trials were used to calculate the statistics which appear at the top of the table. The statistics for the different thresholds were later compiled into a table with all of the thresholds for analysis.

<table>
<thead>
<tr>
<th>Weekday Occupied</th>
<th>Average</th>
<th>193</th>
<th>28</th>
<th>153</th>
<th>-41</th>
<th>267</th>
<th>40</th>
<th>93</th>
<th>157</th>
<th>37</th>
<th>45</th>
<th>40</th>
<th>115</th>
</tr>
</thead>
<tbody>
<tr>
<td>St Dev</td>
<td>-1.683</td>
<td>1.022</td>
<td>1.121</td>
<td>508</td>
<td>1.234</td>
<td>1.083</td>
<td>804</td>
<td>610</td>
<td>644</td>
<td>589</td>
<td>1.073</td>
<td>1.845</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>4,806</td>
<td>4,724</td>
<td>3,022</td>
<td>4,772</td>
<td>6,266</td>
<td>6,015</td>
<td>2,969</td>
<td>3,356</td>
<td>3,450</td>
<td>8,324</td>
<td>10,135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>-3,113</td>
<td>-2,113</td>
<td>-1,899</td>
<td>-1,109</td>
<td>-2,878</td>
<td>-2,327</td>
<td>-1,085</td>
<td>-890</td>
<td>-1,052</td>
<td>-1,052</td>
<td>-1,295</td>
<td>-3,366</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Run Time 3:01:48</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial #</td>
<td>Nov-06</td>
<td>Dec-06</td>
<td>Jan-07</td>
<td>Feb-07</td>
<td>Mar-07</td>
<td>Apr-07</td>
<td>May-07</td>
<td>Jun-07</td>
<td>Jul-07</td>
<td>Aug-07</td>
<td>Sep-07</td>
<td>Oct-07</td>
</tr>
<tr>
<td>1</td>
<td>(127)</td>
<td>2,114</td>
<td>1,752</td>
<td>(216)</td>
<td>(1,739)</td>
<td>934</td>
<td>365</td>
<td>(370)</td>
<td>457</td>
<td>82</td>
<td>505</td>
<td>2,642</td>
</tr>
<tr>
<td>2</td>
<td>1,052</td>
<td>617</td>
<td>(1,298)</td>
<td>506</td>
<td>393</td>
<td>666</td>
<td>(435)</td>
<td>366</td>
<td>(173)</td>
<td>277</td>
<td>(365)</td>
<td>(1,225)</td>
</tr>
<tr>
<td>3</td>
<td>341</td>
<td>(189)</td>
<td>(696)</td>
<td>8</td>
<td>(1,705)</td>
<td>638</td>
<td>758</td>
<td>(206)</td>
<td>(365)</td>
<td>276</td>
<td>(67)</td>
<td>510</td>
</tr>
<tr>
<td>4</td>
<td>(1,138)</td>
<td>149</td>
<td>(239)</td>
<td>(185)</td>
<td>(699)</td>
<td>(929)</td>
<td>(730)</td>
<td>(153)</td>
<td>275</td>
<td>151</td>
<td>985</td>
<td>(332)</td>
</tr>
<tr>
<td>5</td>
<td>(126)</td>
<td>(899)</td>
<td>(738)</td>
<td>(85)</td>
<td>727</td>
<td>(580)</td>
<td>756</td>
<td>454</td>
<td>(32)</td>
<td>(663)</td>
<td>(67)</td>
<td>851</td>
</tr>
</tbody>
</table>

Figure 7: Sample of risk management values for Threshold 1, Weekday Occupied, 3-month window

3.3.5 Data Analysis

The last step in the process was to aggregate all of the data from the different thresholds into a single data set for analysis. The data was assembled around the four period types, plus a total risk value, along with four window sizes. Figure 8 is a sample of the data for one window size and period type.
3.3.6 Test Runs

For the purpose of generating a base value, the model was run against itself.

This means that for each of the twelve months tested, the respective 1-month, 3-month
and 5-month windows were used to calculate the risk management metric within the same 12-month period.

In order to test the base model, the test months for the alarms were changed so that any test month would never run against itself using data from its base period. The thresholds for the test run were calculated based upon the base model data. However, for the purpose of generating alarms, the data used to test never overlapped with the base data. The dates used for the base windows and test model can be seen in Figure 9 and Figure 10.

<table>
<thead>
<tr>
<th>Test Length</th>
<th>24</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Base Model Months</th>
<th>1-Month</th>
<th>3-Month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base Model Window</td>
<td>Test Model Dates</td>
</tr>
<tr>
<td></td>
<td>Start - End</td>
<td>Start - End</td>
</tr>
<tr>
<td>Nov-06</td>
<td>11/1/06 - 11/30/06</td>
<td>12/1/06 - 11/30/06</td>
</tr>
<tr>
<td>Dec-06</td>
<td>12/1/06 - 12/31/06</td>
<td>1/1/07 - 12/31/06</td>
</tr>
<tr>
<td>Jan-07</td>
<td>1/1/07 - 1/31/07</td>
<td>2/1/07 - 1/31/09</td>
</tr>
<tr>
<td>Mar-07</td>
<td>3/1/07 - 3/31/07</td>
<td>4/1/07 - 3/31/09</td>
</tr>
<tr>
<td>Apr-07</td>
<td>4/1/07 - 4/30/07</td>
<td>5/1/07 - 4/30/09</td>
</tr>
<tr>
<td>May-07</td>
<td>5/1/07 - 5/31/07</td>
<td>6/1/07 - 5/31/09</td>
</tr>
<tr>
<td>Jun-07</td>
<td>6/1/07 - 6/30/07</td>
<td>7/1/07 - 6/30/09</td>
</tr>
<tr>
<td>Jul-07</td>
<td>7/1/07 - 7/31/07</td>
<td>8/1/07 - 7/31/09</td>
</tr>
<tr>
<td>Aug-07</td>
<td>8/1/07 - 8/31/07</td>
<td>9/1/07 - 8/31/09</td>
</tr>
<tr>
<td>Sep-07</td>
<td>9/1/07 - 9/30/07</td>
<td>10/1/07 - 9/30/09</td>
</tr>
<tr>
<td>Oct-07</td>
<td>10/1/07 - 10/31/07</td>
<td>11/1/07 - 10/31/09</td>
</tr>
</tbody>
</table>

Figure 9: Base model window and test model dates for the 1-month and 3-month windows
One last point about the test model is that the base month was run against 2 years of data. In order to normalize the calculated “risk management” values to the 12-month base model, all of the calculated values in the test run were divided by two.
Chapter 4: Results and Analysis

4.1 Weather Data

During the development of the interval data summation tool, described in Section 3.2, weather data were included for graphing purposes. The weather data came from the NOAA website (www.noaa.gov), and each day’s energy usage was graphed against the mean daily temperature by period type. Although the decision model does not account for weather, the graphs demonstrate the usage patterns of the buildings. These graphs could be used by an energy manager to better understand a building’s usage by period type based upon outside air temperature (OAT). All thirty-two graphs can be seen in Appendix A. Two interesting results are shown below. The first has a clear inflection point around 73 degrees. This well defined “hockey stick” is indicative of a large piece of equipment, such as a chiller, not being utilized below a certain OAT.

Figure 11: Weekday Occupied plotted against mean daily temperature, Building 1, Meter 1
The second one is worthy of note due to the way that the usage converges above 70 degrees. This building is one which has electric heat in the building’s exterior zones. Once the OAT is great enough, the graph suggests that the building is in cooling mode as the usage converges.

![Weekday Occupied](image)

*Figure 12: Weekday Occupied plotted against mean daily temperature, Building 3, Meter 1*

### 4.2 Decision Model Results

The following graphs are the results of the decision model Base and Test runs.
Figure 13: BASE run, Risk Management total, 1-month and 3-month
Figure 14: BASE run, Risk Management total, 5-month and 12-month
Figure 15: TEST run, Risk Management total, 1-month and 3-month
Figure 16: TEST run, Risk Management total, 5-month and 12-month
Figure 17: BASE run, Weekday Occupied, 1-month and 3-month
### Figure 18: BASE run, Weekday Occupied, 5-month and 12-month

<table>
<thead>
<tr>
<th>Weekday Occupied</th>
<th>5-Month Windows</th>
<th>Weekday Occupied</th>
<th>12-Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>3.238</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>0.986</td>
<td>Average</td>
<td></td>
</tr>
<tr>
<td>Avg</td>
<td>1.523</td>
<td>Std Dev</td>
<td>0.600</td>
</tr>
<tr>
<td>Row 1</td>
<td>6.963</td>
<td>Row 2</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 2</td>
<td>6.963</td>
<td>Row 3</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 3</td>
<td>6.963</td>
<td>Row 4</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 4</td>
<td>6.963</td>
<td>Row 5</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 5</td>
<td>6.963</td>
<td>Row 6</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 6</td>
<td>6.963</td>
<td>Row 7</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 7</td>
<td>6.963</td>
<td>Row 8</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 8</td>
<td>6.963</td>
<td>Row 9</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 9</td>
<td>6.963</td>
<td>Row 10</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 10</td>
<td>6.963</td>
<td>Row 11</td>
<td>6.963</td>
</tr>
<tr>
<td>Row 11</td>
<td>6.963</td>
<td>Row 12</td>
<td>6.963</td>
</tr>
</tbody>
</table>

**Description:**
- **Weekday Occupied 5-Month Windows**
  - Max: 3.238
  - Min: 0.986
- **Weekday Occupied 12-Month**
  - Lower: 0.986
  - Average: 1.523
  - Std Dev: 0.600

**Notes:**
- The figure shows a comparison of weekday occupancy for 5-month and 12-month periods.
- The data is presented in a table format with max, min, and average values.
- The standard deviation is also provided for the 12-month period.
Figure 19: TEST run, Weekday Occupied, 1-month and 3-month
### Figure 20: TEST run, Weekday Occupied, 5-month and 12-month

<table>
<thead>
<tr>
<th>Weekday Occupied</th>
<th>5 Month Windows</th>
<th>12 Month</th>
<th>12 Month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Nov 06</td>
<td>Dec 06</td>
<td>Jan 07</td>
</tr>
<tr>
<td>1</td>
<td>0.0035449</td>
<td>5.458</td>
<td>5.727</td>
</tr>
<tr>
<td>2</td>
<td>0.003983</td>
<td>5.497</td>
<td>5.762</td>
</tr>
<tr>
<td>3</td>
<td>0.004429</td>
<td>5.536</td>
<td>5.802</td>
</tr>
<tr>
<td>4</td>
<td>0.004873</td>
<td>5.576</td>
<td>5.842</td>
</tr>
<tr>
<td>5</td>
<td>0.005319</td>
<td>5.615</td>
<td>5.882</td>
</tr>
<tr>
<td>6</td>
<td>0.005766</td>
<td>5.655</td>
<td>5.922</td>
</tr>
<tr>
<td>7</td>
<td>0.006212</td>
<td>5.695</td>
<td>5.962</td>
</tr>
</tbody>
</table>

Additional data and graphs are provided in the accompanying figures and tables.
Figure 21: BASE run, Weekday Unoccupied, 1-month and 3-month
Figure 22: BASE run, Weekday Unoccupied, 5-month and 12-month
Figure 23: TEST run, Weekday Unoccupied, 1-month and 3-month

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
<td>1.93</td>
</tr>
</tbody>
</table>

**Weekday Unoccupied**

<table>
<thead>
<tr>
<th>1-Month Windows</th>
<th>3-Month Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.667</td>
</tr>
<tr>
<td>2</td>
<td>2.667</td>
</tr>
<tr>
<td>3</td>
<td>2.667</td>
</tr>
<tr>
<td>4</td>
<td>2.667</td>
</tr>
<tr>
<td>5</td>
<td>2.667</td>
</tr>
<tr>
<td>6</td>
<td>2.667</td>
</tr>
<tr>
<td>7</td>
<td>2.667</td>
</tr>
<tr>
<td>8</td>
<td>2.667</td>
</tr>
<tr>
<td>9</td>
<td>2.667</td>
</tr>
<tr>
<td>10</td>
<td>2.667</td>
</tr>
<tr>
<td>11</td>
<td>2.667</td>
</tr>
<tr>
<td>12</td>
<td>2.667</td>
</tr>
</tbody>
</table>

**Weekday Unoccupied**
Figure 24: TEST run, Weekday Unoccupied, 5-month and 12-month
### Table 1: Weekend Occupied

<table>
<thead>
<tr>
<th>Month</th>
<th>1-Month Windows</th>
<th>3-Month Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max</td>
<td>Min</td>
</tr>
<tr>
<td>Nov 06</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Dec 06</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Jan 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Feb 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Mar 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Apr 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>May 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Jun 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Jul 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Aug 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Sep 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
<tr>
<td>Oct 07</td>
<td>1.605</td>
<td>0.84</td>
</tr>
</tbody>
</table>

### Figure 25: BASE run, Weekend Occupied, 1-month and 3-month

- **1-Month**
  - Occupied rate from November 2022 to October 2023.
  - Peaks in the middle of the year.
- **3-Month**
  - Occupied rate from November 2022 to October 2023.
  - Higher peaks compared to 1-month due to long-term trends.

The graphs illustrate the occupancy trends over the specified periods, showing fluctuations and potential peak times.
<table>
<thead>
<tr>
<th>Max</th>
<th>Min</th>
<th>Average</th>
<th>Std (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0054</td>
<td>1.0054</td>
<td>1.0054</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.9881</td>
<td>0.9881</td>
<td>0.9881</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

**Figure 26:** BASE run, Weekend Occupied, 5-month and 12-month
<table>
<thead>
<tr>
<th>Min</th>
<th>Week</th>
<th>1-Month Windows</th>
<th>3-Month Windows</th>
<th>Min</th>
<th>Week</th>
<th>1-Month Windows</th>
<th>3-Month Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.84134</td>
<td>1,033</td>
<td>1,265</td>
<td>157</td>
<td>1,015</td>
<td>50</td>
<td>551</td>
</tr>
<tr>
<td>2</td>
<td>0.88400</td>
<td>1,378</td>
<td>1,486</td>
<td>140</td>
<td>938</td>
<td>70</td>
<td>372</td>
</tr>
<tr>
<td>3</td>
<td>0.88665</td>
<td>1,371</td>
<td>2,226</td>
<td>57</td>
<td>598</td>
<td>76</td>
<td>926</td>
</tr>
<tr>
<td>4</td>
<td>0.90039</td>
<td>1,369</td>
<td>1,234</td>
<td>74</td>
<td>904</td>
<td>76</td>
<td>419</td>
</tr>
<tr>
<td>5</td>
<td>0.93195</td>
<td>1,340</td>
<td>1,188</td>
<td>17</td>
<td>824</td>
<td>71</td>
<td>282</td>
</tr>
<tr>
<td>6</td>
<td>0.95460</td>
<td>1,201</td>
<td>1,100</td>
<td>25</td>
<td>67</td>
<td>54</td>
<td>135</td>
</tr>
<tr>
<td>7</td>
<td>0.97705</td>
<td>1,268</td>
<td>900</td>
<td>74</td>
<td>469</td>
<td>62</td>
<td>167</td>
</tr>
</tbody>
</table>

Figure 27: TEST run, Weekend Occupied, 1-month and 3-month
<table>
<thead>
<tr>
<th>Min</th>
<th>1</th>
<th>1.433</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>1</td>
<td>0.6600</td>
<td>0.6665</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week</th>
<th>Occupied</th>
<th>5-Month Windows</th>
<th>12-Month Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 06</td>
<td>540</td>
<td>527</td>
<td>506</td>
</tr>
<tr>
<td>Dec 06</td>
<td>585</td>
<td>327</td>
<td>748</td>
</tr>
<tr>
<td>Jan 07</td>
<td>531</td>
<td>310</td>
<td>416</td>
</tr>
<tr>
<td>Feb 07</td>
<td>327</td>
<td>200</td>
<td>416</td>
</tr>
<tr>
<td>Mar 07</td>
<td>247</td>
<td>246</td>
<td>389</td>
</tr>
<tr>
<td>Apr 07</td>
<td>310</td>
<td>127</td>
<td>346</td>
</tr>
<tr>
<td>May 07</td>
<td>244</td>
<td>59</td>
<td>116</td>
</tr>
<tr>
<td>Jul 07</td>
<td>116</td>
<td>59</td>
<td>116</td>
</tr>
<tr>
<td>Sep 07</td>
<td>116</td>
<td>59</td>
<td>116</td>
</tr>
</tbody>
</table>

Figure 28: TEST run, Weekend Occupied, 5-month and 12-month
Figure 29: BASE run, Weekend Unoccupied, 1-month and 3-month
### Figure 30: BASE run, Weekend Unoccupied, 5-month and 12-month

<table>
<thead>
<tr>
<th>Max</th>
<th>4.048</th>
<th>2.244</th>
</tr>
</thead>
</table>

#### Weekend Unoccupied

**5-Month Windows**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**12-Month**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

*Note: The table and diagram are not transcribed here as they are visual representations.*
Figure 31: TEST run, Weekend Unoccupied, 1-month and 3-month
Figure 32: TEST run, Weekend Unoccupied, 5-month and 12-month
4.3 Data Quality

In order to perform a review of the data from the base and test periods, the daily total usage data for the eight buildings were plotted over a four-year period. Next, a 12 trailing month (TTM) average was calculated for the daily usages and a linear trend was added to determine the slope of the line. A positive slope indicates an increase in energy over time, while a negative one would imply a decreasing trend. The slopes of the lines for the TTM data are in Table 6. All eight of the graphs can be seen in Appendix D. Four of the buildings had negative slopes, while the remaining four had positive slopes. Using Meter 1-1, the slope would suggest a 5.4% increase in usage per year.

Table 6: Slopes of TTM building data for all daily energy usages over Base and Test period.

<table>
<thead>
<tr>
<th>Meter</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>0.774</td>
</tr>
<tr>
<td>1-2</td>
<td>(1.606)</td>
</tr>
<tr>
<td>2-1</td>
<td>0.948</td>
</tr>
<tr>
<td>2-2</td>
<td>0.122</td>
</tr>
<tr>
<td>3-1</td>
<td>(0.172)</td>
</tr>
<tr>
<td>4-1</td>
<td>(2.445)</td>
</tr>
<tr>
<td>5-1</td>
<td>0.146</td>
</tr>
<tr>
<td>5-2</td>
<td>(0.105)</td>
</tr>
</tbody>
</table>

In addition, normal probability plots of the base period energy usage were created for all of the meters by period type. They can be seen in Appendix E. The plots utilized the same 16-months of data that from the base case. These plots were generated since the assumption in the model was that the daily usage values were normally distributed. As discussed previously, the thresholds for triggering the alarms were calculated under that assumption.

Approximately one-third of the meters have profiles with good fits to a normal distribution. However, in the majority of the cases, it is clear that above and below 2 standard deviations, as the data approaches the tails, the plots broke down and
changed their slope. For example, this can be demonstrated by using two graphs from Building 5, Meter1 as seen in Figure 33 and Figure 34. The weekday occupied graph has an excellent fit as it is very straight. However, looking at the lower end of the weekend unoccupied, the data clearly drops off which would suggest that the data is skewed.

Figure 33: Building 5, Meter 1, normal probability plot for the weekday occupied period

Figure 34: Building 5, Meter 1, normal probability plot for the weekend unoccupied period
4.4 Analysis of Results

By observation, the graphs of the data from the base and test runs for all of the windows have very similar profiles. A simple test of the resulting data shows a high degree of correlation between the Base and Test results. All of the result arrays for the averages and standard deviations were tested for correlation. The results can be seen in Table 7 and Table 8.

### Table 7: Correlations of Base and Test run result arrays for data averages

<table>
<thead>
<tr>
<th>Correlation</th>
<th>1-month</th>
<th>3-month</th>
<th>5-month</th>
<th>12-month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Management</td>
<td>0.976</td>
<td>0.980</td>
<td>0.972</td>
<td>0.977</td>
</tr>
<tr>
<td>Weekday Occupied</td>
<td>0.942</td>
<td>0.951</td>
<td>0.972</td>
<td>0.903</td>
</tr>
<tr>
<td>Weekday Unoccupied</td>
<td>0.927</td>
<td>0.939</td>
<td>0.952</td>
<td>0.936</td>
</tr>
<tr>
<td>Weekend Occupied</td>
<td>0.952</td>
<td>0.828</td>
<td>0.632</td>
<td>0.932</td>
</tr>
<tr>
<td>Weekend Unoccupied</td>
<td>0.981</td>
<td>0.963</td>
<td>0.940</td>
<td>0.975</td>
</tr>
</tbody>
</table>

### Table 8: Correlations of Base and Test run result arrays for standard deviations

<table>
<thead>
<tr>
<th>Correlation</th>
<th>1-month</th>
<th>3-month</th>
<th>5-month</th>
<th>12-month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Management</td>
<td>0.953</td>
<td>0.970</td>
<td>0.957</td>
<td>0.957</td>
</tr>
<tr>
<td>Weekday Occupied</td>
<td>0.863</td>
<td>0.921</td>
<td>0.938</td>
<td>0.900</td>
</tr>
<tr>
<td>Weekday Unoccupied</td>
<td>0.954</td>
<td>0.949</td>
<td>0.950</td>
<td>0.906</td>
</tr>
<tr>
<td>Weekend Occupied</td>
<td>0.948</td>
<td>0.876</td>
<td>0.745</td>
<td>0.923</td>
</tr>
<tr>
<td>Weekend Unoccupied</td>
<td>0.988</td>
<td>0.993</td>
<td>0.982</td>
<td>0.962</td>
</tr>
</tbody>
</table>

For the data average array, with the exception of the 3-month and 5-month windows for the Weekend Occupied period type, all of the correlations were above 0.9. For the standard deviation array, all of the correlations were above 0.9, with the exception of the 3-month and 5-month windows for the Weekend Occupied period type, plus the Weekday Occupied 1-month window.

For all of the period types, there is a smoothing of the profiles as the window size increases. In all cases, the 1-month period had the most dramatic changes with
many peaks and valleys. The size of the threshold had little to no effect on this. For the 12-month window, the result was an essentially flat profile. The 5-month window had the smoothest profile of the three short data sets.

The 5-month window most clearly demonstrates the seasonality of the various buildings’ energy profiles. The risk management values in the winter are mostly lower. This can be explained by the increased electrical usage of several of the buildings in the winter. With an increased usage during those months, the alarms throughout the year would not be triggered much less frequently.

It is clear that when utilizing this model, the shorter the data set, the less predictive value there is for setting alarming thresholds for an entire year. One example of this can be seen by reviewing the Weekday Occupied 5-month Test results, as seen in Table 9 and Figure 35. For the months of June through August, the results for the seventh threshold were all greater than the first threshold for the months of January and February.

Table 9: Weekday Occupied Test run results for the 5-month window

<table>
<thead>
<tr>
<th>Weekday Occupied</th>
<th>5-Month Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nov-06</td>
</tr>
<tr>
<td>1</td>
<td>0.8413</td>
</tr>
<tr>
<td>2</td>
<td>0.0640</td>
</tr>
<tr>
<td>3</td>
<td>0.8666</td>
</tr>
<tr>
<td>4</td>
<td>0.9053</td>
</tr>
<tr>
<td>5</td>
<td>0.9319</td>
</tr>
<tr>
<td>6</td>
<td>0.9546</td>
</tr>
<tr>
<td>7</td>
<td>0.9772</td>
</tr>
</tbody>
</table>
Figure 35: Weekday Occupied Test run results for the 5-month window

While the 5-month window clearly has the least variability of the three short data sets, it also appears to be more valuable in setting alarm thresholds than the 12-month window. Due to the loss of seasonality and flatness in the profiles, if the 12-month window were used, there would be lost opportunities for investigating alarms in the summer and winter. For example, by utilizing the weekday occupied results again, the rate of alarming in the summer would be lower than needed, leading to missed opportunities to investigate usage.
The ultimate question for future testing and examination of other building portfolios is, what should the appropriate level for setting thresholds be? In order to draw a final conclusion, the arrays of the values found in the results of Figure 13 through Figure 32 were simplified by averaging the values over the twelve test months for each threshold. These averages were then plotted against their respective threshold values. Next, a second order polynomial was calculated for its best curve fit. Using the equation...
for the three window sizes, the parabolic curves were extended by calculating their theoretical values using the two probability values of 0.500 and 1.000. Then, for each equation, the derivative was taken. For any 2nd order equation of the form

\[ y = ax^2 + bx + c, \]

by taking the derivative and then by setting the resulting equation to be equal to 0, the maxima value is found to be:

\[ x^* = \frac{-b}{2a} \]

This optimal value was calculated for the Base and Test cases in each of the three short window sizes and four period types plus the total Risk Management type. Finally, using the optimal probabilities, the standard deviation was calculated. All of the graphs and results can be found in the tables on the following five pages.

An examination of the results shows that the optimal point is found to be in the range between 0.8 and 1.0.
Window | Best Fit | \( x^* \) | \( z(x^*) \)
--- | --- | --- | ---
1-month | \( y = -88,250x^2 + 142,507x + -49,537 \) | 0.8074 | 0.8684 |
3-month | \( y = -77,225x^2 + 119,992x + -38,779 \) | 0.7769 | 0.7618 |
5-month | \( y = -120,632x^2 + 194,936x + -71,341 \) | 0.8080 | 0.8705 |

Window | Best Fit | \( x' \) | \( z(x') \)
--- | --- | --- | ---
1-month | \( y = -75,194x^2 + 115,707x + -34,140 \) | 0.7694 | 0.7368 |
3-month | \( y = -99,097 x^2 + 156,058x + -51,626 \) | 0.7874 | 0.7974 |
5-month | \( y = -102,364 x^2 + 158,581x + -51,761 \) | 0.7746 | 0.7541 |
### BASE Weekday Occupied

<table>
<thead>
<tr>
<th>Window</th>
<th>Best Fit</th>
<th>x*</th>
<th>z (x*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>$y = -42,429x^2 + 72,159x + -28,214$</td>
<td>0.8503</td>
<td>1.0379</td>
</tr>
<tr>
<td>3-month</td>
<td>$y = -30,496x^2 + 50,090x + -18,204$</td>
<td>0.8213</td>
<td>0.9202</td>
</tr>
<tr>
<td>5-month</td>
<td>$y = -39,777x^2 + 65,442x + -24,620$</td>
<td>0.8226</td>
<td>0.9254</td>
</tr>
</tbody>
</table>

### TEST Weekday Occupied

<table>
<thead>
<tr>
<th>Window</th>
<th>Best Fit</th>
<th>x*</th>
<th>z (x*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>$y = -39,579x^2 + 64,039x + -22,146$</td>
<td>0.7988</td>
<td>0.8372</td>
</tr>
<tr>
<td>3-month</td>
<td>$y = -40,081x^2 + 64,030x + -21,883$</td>
<td>0.7988</td>
<td>0.8372</td>
</tr>
<tr>
<td>5-month</td>
<td>$y = -40,643x^2 + 63,227x + -20,901$</td>
<td>0.7778</td>
<td>0.7649</td>
</tr>
</tbody>
</table>
### BASE Weekday Unoccupied

Window | Best Fit | $x^*$ | $z (x^*)$
--- | --- | --- | ---
1-month | $y = -14,776x^2 + 20,663x + -4,603$ | 0.6992 | 0.5221
3-month | $y = -25,851x^2 + 39,753x + -12,918$ | 0.7689 | 0.7351
5-month | $y = -44,106x^2 + 72,499x + -27,690$ | 0.8219 | 0.9225

### TEST Weekday Unoccupied

Window | Best Fit | $x^*$ | $z (x^*)$
--- | --- | --- | ---
1-month | $y = -11,063x^2 + 14,253x + -1,399$ | 0.6442 | 0.3696
3-month | $y = -27,203x^2 + 42,058x + -13,368$ | 0.7730 | 0.7489
5-month | $y = -44,928x^2 + 73,194x + -27,098$ | 0.8146 | 0.8948
### BASE Weekend Occupied

<table>
<thead>
<tr>
<th>Window</th>
<th>Best Fit</th>
<th>$x^*$</th>
<th>$z(x^*)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>$y = -10,994x^2 + 17,652x - 6,272$</td>
<td>0.8028</td>
<td>0.8516</td>
</tr>
<tr>
<td>3-month</td>
<td>$y = -14,438x^2 + 23,568x - 8,902$</td>
<td>0.8162</td>
<td>0.9010</td>
</tr>
<tr>
<td>5-month</td>
<td>$y = -27,394x^2 + 46,992x - 19,482$</td>
<td>0.8577</td>
<td>1.0700</td>
</tr>
</tbody>
</table>

### TEST Weekend Occupied

<table>
<thead>
<tr>
<th>Window</th>
<th>Best Fit</th>
<th>$x^*$</th>
<th>$z(x^*)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>$y = -9,462x^2 + 15,246x - 5,404$</td>
<td>0.8057</td>
<td>0.8620</td>
</tr>
<tr>
<td>3-month</td>
<td>$y = -14,478x^2 + 23,585x - 8,998$</td>
<td>0.8145</td>
<td>0.8946</td>
</tr>
<tr>
<td>5-month</td>
<td>$y = -2,367x^2 + 1,569x + 910$</td>
<td>0.3315</td>
<td>0.4359</td>
</tr>
</tbody>
</table>
### BASE Weekend Unoccupied

<table>
<thead>
<tr>
<th>Window</th>
<th>Best Fit</th>
<th>$x^*$</th>
<th>$z (x^*)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>$y = -20,049x^2 + 32,030x + -10,447$</td>
<td>0.7988</td>
<td>0.8373</td>
</tr>
<tr>
<td>3-month</td>
<td>$y = -6,441x^2 + 6,584x + 1,244$</td>
<td>0.5111</td>
<td>0.0277</td>
</tr>
<tr>
<td>5-month</td>
<td>$y = -9,355x^2 + 10,002x + 452$</td>
<td>0.5346</td>
<td>0.0869</td>
</tr>
</tbody>
</table>

### TEST Weekend Unoccupied

<table>
<thead>
<tr>
<th>Window</th>
<th>Best Fit</th>
<th>$x^*$</th>
<th>$z (x^*)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>$y = -15,089x^2 + 22,170x + -5,191$</td>
<td>0.7346</td>
<td>0.6268</td>
</tr>
<tr>
<td>3-month</td>
<td>$y = -17,334x^2 + 26,382x + -7,375$</td>
<td>0.7610</td>
<td>0.7095</td>
</tr>
<tr>
<td>5-month</td>
<td>$y = -14,424x^2 + 20,587x + -4,671$</td>
<td>0.7136</td>
<td>0.5641</td>
</tr>
</tbody>
</table>
Chapter 5: Conclusions

5.1 Summary

The goal of the study was to determine what an optimal threshold value would be for setting alarm threshold to test energy usage in a building using short data sets. Several years of 15-minute interval data was utilized from five buildings in Jacksonville, Florida. For this paper, the concept of “risk management” was developed in order to calculate a value, in dollars which would normalize the cost of the time associated with investigating alarms, and the benefit of energy “saved” which would result from the same investigation.

The results of the decision model developed for this study indicate that when utilizing short data sets for predicting annual alarming thresholds, a 5-month window provides better results than a 1-month or 3-month window. An assumption in the original formulation of the model was that the optimal value for establishing alarm thresholds would most likely fall somewhere between 1.0 and 2.0 standard deviations. In the end, it was observed that the optimal value for setting the thresholds may be in the 0.8 to 1.0 standard deviation range above the average for this data set.
5.2 Recommendations for Future Research

Due to the nature of this project, this study was limited in scope and in several of its assumptions. Based upon the work done, there are various items that could be addressed in future studies.

1) Modifications could be made to the calculation of the risk management value. In this paper, the projected savings and load factor multipliers were factored in after the alarm response and alarm length were determined. This variable should be part of the formula for determining response type and length, so that critical buildings would be considered more carefully.

2) The interval data could be corrected so that it is more normal. There are various statistical procedures that could be used, but one might be to perform a log normal transformation on the data.

3) The thresholds were calculated using the same probability for all the meters at any given time during the runs. Efforts to optimize the thresholds would provide further benefit. The result would be lower thresholds for larger buildings, while smaller, less “risky” buildings would have higher thresholds.

4) Variations in the time horizons might provide better insight on how useful short data sets are. Instead of running the model by testing the windows against twelve or twenty-four months of data, further study could result from testing the short data sets against shorter time horizons.

5) A sensitivity analysis could be performed on the various response type and duration distributions to see how their change might impact the results.
Appendix A: Building Profiles by Occupancy Type

Figure 37: Building 1, Meter 1, energy usage by period type
Figure 38: Building 1, Meter 2, energy usage by period type

Weekday Occupied

Weekday Unoccupied

Weekend Occupied

Weekend Unoccupied
Figure 39: Building 2, Meter 1, energy usage by period type
Figure 40: Building 2, Meter 2, energy usage by period type

- **Weekday Occupied**
  - kWh vs. Average Daily Temp
  - Data points show energy usage variation with temperature changes.

- **Weekday Unoccupied**
  - kWh vs. Average Daily Temp
  - Data points indicate energy usage with temperature, showing different patterns from occupied periods.

- **Weekend Occupied**
  - kWh vs. Average Daily Temp
  - Data points reflect energy usage on weekends, correlating with temperature.

- **Weekend Unoccupied**
  - kWh vs. Average Daily Temp
  - Data points illustrate energy usage on weekends without occupancy, showing a distinct trend compared to occupied periods.
Figure 41: Building 3, Meter 1, energy usage by period type

Weekday Occupied

Weekday Unoccupied

Weekend Occupied

Weekend Unoccupied
Figure 42: Building 4, Meter 1, energy usage by period type
Figure 43: Building 5, Meter 1, energy usage by period type
Figure 44: Building 5, Meter 2, energy usage by period type

- **Weekday Occupied**
- **Weekday Unoccupied**
- **Weekend Occupied**
- **Weekend Unoccupied**
Appendix B:
VBA for Interval Data Aggregation

Sub Step1_SummarizeIntervalData()

Application.ScreenUpdating = False

'***Clear Workbook

Sheets("DailyPeriodStats").Select
Range("A4:AD503").Select
Selection.ClearContents

Sheets("MonthlyPeriodStats").Select
Range("B4:AD20").Select
Selection.ClearContents

'***This section loads the interval data into the RawData array

Dim rowRaw As Long, rowDaily As Long
Dim row As Long, col As Byte
Dim StartDate As Single, EndDate As Single
Dim NumberOfIntervals As Long
Dim StartTime As Single, EndTime As Single, RunTime As Single

StartTime = Time
StartDate = Sheets("inputs").Range("D4")
EndDate = Sheets("inputs").Range("D5")
NumberOfIntervals = (EndDate - StartDate) * 96 + 1

Dim RawData() As Variant
ReDim RawData(1 To NumberOfIntervals, 1 To 3) As Variant

For rowRaw = LBound(RawData) To UBound(RawData)
    For col = 1 To 2
        If Sheets("IntervalData").Cells(rowRaw + 5, col) = " "
        Then RawData(rowRaw, col) = 0
        Else RawData(rowRaw, col) = 
            Sheets("IntervalData").Cells(rowRaw + 5, col)
    Next col
Next rowRaw

'***This section assigns a period type To the interval. This is for use later on with both the Daily & Monthly summaries.

Dim WkDayOcc As Single, WkDayUnocc As Single
Dim WkEndOcc As Single, WkEndUnocc As Single
Dim DayType As Byte, OccState As Byte

Sheets("inputs").Select
WKDayOcc = Range("D8").Value        'Case 3 (11)
WKDayUnocc = Range("E8").Value      'Case 2 (10)
WKEndOcc = Range("D9").Value        'Case 1 (01)
WKEndUnocc = Range("E9").Value      'Case 0 (00)
Sheets("IntervalData").Select
For rowRaw = LBound(RawData) To UBound(RawData)
   OccState = 0
   'Determines Weekday or Weekend
   If (Int(RawData(rowRaw, 1)) Mod 7) <= 1
cntclm  Then DayType = 0
   Else DayType = 1 'WkEnd=0to1, WkDay=2to6 (Sa=0,...,F=6)
   'Sets Mon 12AM to Weekend
   If RawData(rowRaw, 1) Mod 7 = 2
      And RawData(rowRaw, 1) = Int(RawData(rowRaw, 1))
      Then DayType = 0
      'Sets Sat 12AM to Weekday
   If RawData(rowRaw, 1) Mod 7 = 0
      And RawData(rowRaw, 1) = Int(RawData(rowRaw, 1))
      Then DayType = 1
   'Determines Occupied or Unoccupied  (Occ=1, Unocc=0)
   If DayType = 0
      And ((RawData(rowRaw, 1) - Int(RawData(rowRaw, 1))) >
      WkEndOcc) _
      And ((RawData(rowRaw, 1) - Int(RawData(rowRaw, 1))) <=
      WkEndUnocc)
      Then OccState = 1
   If DayType = 1
      And ((RawData(rowRaw, 1) - Int(RawData(rowRaw, 1))) >
      WkDayOcc) _
      And ((RawData(rowRaw, 1) - Int(RawData(rowRaw, 1))) <=
      WkDayUnocc)
      Then OccState = 1
   'Calculates Period Type
   RawData(rowRaw, 3) = DayType * 2 + OccState * 1
Next rowRaw
   ***This calculates the basic DAILY stats For the 4 period types
   'They are Count, Sum, Average, Min, Max, StDev
   Dim NumberOfDays As Integer
   'Weekday Occupied
   Dim WdOccCount As Single, WdOccSum As Single
   Dim WdOccAvg As Single, WdOccStDev As Single
   Dim WdOccMin As Single, WdOccMax As Single
   Dim WdOccSumSqr As Single
   'Weekday Unoccupied
   Dim WdUnoccCount As Single, WdUnoccSum As Single
   Dim WdUnoccAvg As Single, WdUnoccStDev As Single
   Dim WdUnoccMin As Single, WdUnoccMax As Single
   Dim WdUnoccSumSqr As Single
   'Weekend Occupied
   Dim WeOccCount As Single, WeOccSum As Single
   Dim WeOccAvg As Single, WeOccStDev As Single
   Dim WeOccMin As Single, WeOccMax As Single
   Dim WeOccSumSqr As Single
'Weekend Unoccupied
Dim WeUnoccCount As Single, WeUnoccSum As Single
Dim WeUnoccAvg As Single, WeUnoccStDev As Single
Dim WeUnoccMin As Single, WeUnoccMax As Single
Dim WeUnoccSumSqr As Single

Dim junk As Single
Dim Counter As Integer

NumberOfDays = EndDate - StartDate

Dim DailyData() As Variant
ReDim DailyData(1 To NumberOfDays, 1 To 30) As Variant

'Loads the dates into column 1 of the DailyData array
For rowDaily = LBound(DailyData) To UBound(DailyData)
    DailyData(rowDaily, 1) = StartDate + (rowDaily - 1)
    DailyData(rowDaily, 2) = _
        Sheets("inputs").Cells(rowDaily + 3, 13)
Next rowDaily

For rowDaily = LBound(DailyData) To UBound(DailyData)

    'Reset variables for next day
    WdOccCount = 0
    WdOccSum = 0
    WdOccMin = 1000000
    WdOccMax = 0
    WdOccSumSqr = 0
    WdOccAvg = 0
    WdOccStDev = 0

    WdUnoccCount = 0
    WdUnoccSum = 0
    WdUnoccMin = 1000000
    WdUnoccMax = 0
    WdUnoccSumSqr = 0
    WdUnoccAvg = 0
    WdUnoccStDev = 0

    WeOccCount = 0
    WeOccSum = 0
    WeOccMin = 1000000
    WeOccMax = 0
    WeOccSumSqr = 0
    WeOccAvg = 0
    WeOccStDev = 0

    WeUnoccCount = 0
    WeUnoccSum = 0
    WeUnoccMin = 1000000
    WeUnoccMax = 0
    WeUnoccSumSqr = 0
    WeUnoccAvg = 0
    WeUnoccStDev = 0

    For rowRaw = (((rowDaily - 1) * 96) + 2) To _
Select Case RawData(rowRaw, 2)
    Case Is > 0
        Select Case RawData(rowRaw, 3)
            Case 3          'Weekday Occupied
                WdOccCount = WdOccCount + 1
                WdOccSum = WdOccSum + RawData(rowRaw, 2)
                If RawData(rowRaw, 2) < WdOccMin Then _
                    WdOccMin = RawData(rowRaw, 2)
                If RawData(rowRaw, 2) > WdOccMax Then _
                    WdOccMax = RawData(rowRaw, 2)
                WdOccSumSqr = _
                    WdOccSumSqr + RawData(rowRaw, 2) ^ 2
            Case 2          'Weekday Unoccupied
                WdUnoccCount = WdUnoccCount + 1
                WdUnoccSum = WdUnoccSum + RawData(rowRaw, 2)
                If RawData(rowRaw, 2) < WdUnoccMin Then _
                    WdUnoccMin = RawData(rowRaw, 2)
                If RawData(rowRaw, 2) > WdUnoccMax Then _
                    WdUnoccMax = RawData(rowRaw, 2)
                WdUnoccSumSqr = _
                    WdUnoccSumSqr + RawData(rowRaw, 2) ^ 2
            Case 1          'Weekend Occupied
                WeOccCount = WeOccCount + 1
                WeOccSum = WeOccSum + RawData(rowRaw, 2)
                If RawData(rowRaw, 2) < WeOccMin Then _
                    WeOccMin = RawData(rowRaw, 2)
                If RawData(rowRaw, 2) > WeOccMax Then _
                    WeOccMax = RawData(rowRaw, 2)
                WeOccSumSqr = _
                    WeOccSumSqr + RawData(rowRaw, 2) ^ 2
            Case 0          'Weekend Unoccupied
                WeUnoccCount = WeUnoccCount + 1
                WeUnoccSum = WeUnoccSum + RawData(rowRaw, 2)
                If RawData(rowRaw, 2) < WeUnoccMin Then _
                    WeUnoccMin = RawData(rowRaw, 2)
                If RawData(rowRaw, 2) > WeUnoccMax Then _
                    WeUnoccMax = RawData(rowRaw, 2)
                WeUnoccSumSqr = _
                    WeUnoccSumSqr + RawData(rowRaw, 2) ^ 2
        End Select
    End Select
End Select

Next rowRaw

'Weekday Occupied
If WdOccCount > 0 Then _
    WdOccAvg = WdOccSum / WdOccCount
If WdOccSumSqr <= WdOccCount * WdOccAvg ^ 2 Then _
    WdOccSumSqr = WdOccSumSqr +
        (WdOccSumSqr - WdOccCount * WdOccAvg ^ 2) * -1 + 1
If WdOccCount > 1 Then _
    WdOccStDev = Sqr((WdOccSumSqr - WdOccCount _
        * WdOccAvg ^ 2) / (WdOccCount - 1))
If WdOccCount > 0 Then DailyData(rowDaily, 3) = "Yes"
If WdOccCount > 0 Then DailyData(rowDaily, 4) = WdOccCount
If WdOccSum > 0 Then DailyData(rowDaily, 5) = WdOccSum
If WdOccAvg > 0 Then DailyData(rowDaily, 6) = WdOccAvg
If WdOccMin < 1000000 Then DailyData(rowDaily, 7) = WdOccMin
If WdOccMax > 0 Then DailyData(rowDaily, 8) = WdOccMax
If WdOccStDev > 0 Then DailyData(rowDaily, 9) = WdOccStDev

'Weekday Unoccupied
If WdUnoccCount > 0 Then
    WdUnoccAvg = WdUnoccSum / WdUnoccCount
If WdUnoccSumSqr <= WdUnoccCount * WdUnoccAvg ^ 2 Then
    WdUnoccSumSqr = WdUnoccSumSqr +
    (WdUnoccSumSqr - WdUnoccCount * WdUnoccAvg ^ 2) * -1 + 1
If WdUnoccCount > 1 Then
    WdUnoccStDev = Sqr((WdUnoccSumSqr - WdUnoccCount *
                        * WdUnoccAvg ^ 2) / (WdUnoccCount - 1))
If WdUnoccCount > 0 Then DailyData(rowDaily, 10) = "Yes"
If WdUnoccCount > 0 Then DailyData(rowDaily, 11) = WdUnoccCount
If WdUnoccSum > 0 Then DailyData(rowDaily, 12) = WdUnoccSum
If WdUnoccAvg > 0 Then DailyData(rowDaily, 13) = WdUnoccAvg
If WdUnoccMin < 1000000 Then
    DailyData(rowDaily, 14) = WdUnoccMin
If WdUnoccMax > 0 Then DailyData(rowDaily, 15) = WdUnoccMax
If WdUnoccStDev > 0 Then DailyData(rowDaily, 16) = WdUnoccStDev

'Weekend Occupied
If WeOccCount > 0 Then
    WeOccAvg = WeOccSum / WeOccCount
If WeOccSumSqr <= WeOccCount * WeOccAvg ^ 2 Then
    WeOccSumSqr = WeOccSumSqr +
    (WeOccSumSqr - WeOccCount * WeOccAvg ^ 2) * -1 + 1
If WeOccCount > 1 Then
    WeOccStDev = Sqr((WeOccSumSqr - WeOccCount *
                        * WeOccAvg ^ 2) / (WeOccCount - 1))
If WeOccCount > 0 Then DailyData(rowDaily, 17) = "Yes"
If WeOccCount > 0 Then DailyData(rowDaily, 18) = WeOccCount
If WeOccSum > 0 Then DailyData(rowDaily, 19) = WeOccSum
If WeOccAvg > 0 Then DailyData(rowDaily, 20) = WeOccAvg
If WeOccMin < 1000000 Then
    DailyData(rowDaily, 21) = WeOccMin
If WeOccMax > 0 Then DailyData(rowDaily, 22) = WeOccMax
If WeOccStDev > 0 Then DailyData(rowDaily, 23) = WeOccStDev

'Weekend Unoccupied
If WeUnoccCount > 0 Then
    WeUnoccAvg = WeUnoccSum / WeUnoccCount
If WeUnoccSumSqr <= WeUnoccCount * WeUnoccAvg ^ 2 Then
    WeUnoccSumSqr = WeUnoccSumSqr +
    (WeUnoccSumSqr - WeUnoccCount * WeUnoccAvg ^ 2) * -1 + 1
If WeUnoccCount > 1 Then
    WeUnoccStDev = Sqr((WeUnoccSumSqr - WeUnoccCount *
                        * WeUnoccAvg ^ 2) / (WeUnoccCount - 1))
If WeUnoccCount > 0 Then DailyData(rowDaily, 24) = "Yes"
If WeUnoccCount > 0 Then DailyData(rowDaily, 25) = WeUnoccCount
If WeUnoccSum > 0 Then DailyData(rowDaily, 26) = WeUnoccSum
If WeUnoccAvg > 0 Then DailyData(rowDaily, 27) = WeUnoccAvg
If WeUnoccMin < 1000000 Then _
DailyData(rowDaily, 28) = WeUnoccMin
If WeUnoccMax > 0 Then DailyData(rowDaily, 29) = WeUnoccMax
If WeUnoccStDev > 0 Then DailyData(rowDaily, 30) = WeUnoccStDev

Next rowDaily

'***Copies DailyData array back to workbook

Dim SumString As String
Sheets("DailyPeriodStats").Select

For row = 1 To NumberOfDays
    For col = 1 To 4
        Cells(row + 3, col).Value = DailyData(row, col)
    Next col
    'col=5
    If DailyData(row, col - 1) > 0 Then
        Cells(row + 3, col) = ="IF(C" & (row + 3) & "," & DailyData(row, col) & ",""
    End If
    For col = 6 To 11
        Cells(row + 3, col).Value = DailyData(row, col)
    Next col
    'col=12
    If DailyData(row, col - 1) > 0 Then
        Cells(row + 3, col) = ="IF(J" & (row + 3) & "," & DailyData(row, col) & ",""
    End If
    For col = 13 To 18
        Cells(row + 3, col).Value = DailyData(row, col)
    Next col
    'col=19
    If DailyData(row, col - 1) > 0 Then
        Cells(row + 3, col) = ="IF(Q" & (row + 3) & "," & DailyData(row, col) & ",""
    End If
    For col = 20 To 25
        Cells(row + 3, col).Value = DailyData(row, col)
    Next col
    'col=26
    If DailyData(row, col - 1) > 0 Then
        Cells(row + 3, col) = ="IF(X" & (row + 3) & "," & DailyData(row, col) & ",""
    End If
    For col = 27 To 30
        Cells(row + 3, col).Value = DailyData(row, col)
    Next col
Next row

'***This calculates the basic MONTHLY stats For the 4 period types
Dim MonthlyData() As Variant
ReDim MonthlyData(1 To 16, 1 To 31) As Variant
Dim rowMonthly As Byte
Dim NumDaysInMonth(1 To 16) As Integer
Dim StartRow As Long, EndRow As Long
Dim MonthlyTempSum As Integer
StartRow = 0
EndRow = 1

'Loads Months into MonthlyData array
Sheets("MonthlyPeriodStats").Select
For rowMonthly = 1 To 16
    MonthlyData(rowMonthly, 1) = Cells(rowMonthly + 3, 1)
    NumDaysInMonth(rowMonthly) = Cells(rowMonthly + 4, 1) - Cells(rowMonthly + 3, 1)

    'Calculate the monthly average temp
    MonthlyTempSum = 0
    For i = 1 To NumberOfDays
        If (Month(DailyData(i, 1)) = Month(MonthlyData(rowMonthly, 1)) And Year(DailyData(i, 1)) = Year(MonthlyData(rowMonthly, 1)))
            MonthlyTempSum = MonthlyTempSum + DailyData(i, 2)
    Next i
    MonthlyData(rowMonthly, 2) = Round(MonthlyTempSum / NumDaysInMonth(rowMonthly), 1)
Next rowMonthly

For rowMonthly = 1 To 16
    'Resets variables For the Next month
    WdOccCount = 0
    WdOccSum = 0
    WdOccMin = 1000000
    WdOccMax = 0
    WdOccSumSqr = 0
    WdOccAvg = 0
    WdOccStDev = 0
    WdUnoccCount = 0
    WdUnoccSum = 0
    WdUnoccMin = 1000000
    WdUnoccMax = 0
    WdUnoccSumSqr = 0
    WdUnoccAvg = 0
    WdUnoccStDev = 0
    WeOccCount = 0
    WeOccSum = 0
    WeOccMin = 1000000
    WeOccMax = 0
    WeOccSumSqr = 0
WeOccAvg = 0
WeOccStDev = 0

WeUnoccCount = 0
WeUnoccSum = 0
WeUnoccMin = 1000000
WeUnoccMax = 0
WeUnoccSumSqr = 0
WeUnoccAvg = 0
WeUnoccStDev = 0

StartRow = EndRow + 1
EndRow = StartRow + (NumDaysInMonth(rowMonthly) * 96 - 1)

For rowRaw = StartRow To EndRow
Select Case RawData(rowRaw, 2)
Case Is > 0
Select Case RawData(rowRaw, 3)
Case 3          'Weekday Occupied
    WdOccCount = WdOccCount + 1
    WdOccSum = WdOccSum + RawData(rowRaw, 2)
    If RawData(rowRaw, 2) < WdOccMin Then _
        WdOccMin = RawData(rowRaw, 2)
    If RawData(rowRaw, 2) > WdOccMax Then _
        WdOccMax = RawData(rowRaw, 2)
    WdOccSumSqr = _
        WdOccSumSqr + RawData(rowRaw, 2) ^ 2
Case 2          'Weekday Unoccupied
    WdUnoccCount = WdUnoccCount + 1
    WdUnoccSum = WdUnoccSum + RawData(rowRaw, 2)
    If RawData(rowRaw, 2) < WdUnoccMin Then _
        WdUnoccMin = RawData(rowRaw, 2)
    If RawData(rowRaw, 2) > WdUnoccMax Then _
        WdUnoccMax = RawData(rowRaw, 2)
    WdUnoccSumSqr = _
        WdUnoccSumSqr + RawData(rowRaw, 2) ^ 2
Case 1          'Weekend Occupied
    WeOccCount = WeOccCount + 1
    WeOccSum = WeOccSum + RawData(rowRaw, 2)
    If RawData(rowRaw, 2) < WeOccMin Then _
        WeOccMin = RawData(rowRaw, 2)
    If RawData(rowRaw, 2) > WeOccMax Then _
        WeOccMax = RawData(rowRaw, 2)
    WeOccSumSqr = _
        WeOccSumSqr + RawData(rowRaw, 2) ^ 2
Case 0          'Weekend Unoccupied
    WeUnoccCount = WeUnoccCount + 1
    WeUnoccSum = WeUnoccSum + RawData(rowRaw, 2)
    If RawData(rowRaw, 2) < WeUnoccMin Then _
        WeUnoccMin = RawData(rowRaw, 2)
    If RawData(rowRaw, 2) > WeUnoccMax Then _
        WeUnoccMax = RawData(rowRaw, 2)
    WeUnoccSumSqr = _
        WeUnoccSumSqr + RawData(rowRaw, 2) ^ 2
End Select
End Select

End For
Next rowRaw

'Weekday Occupied
If WdOccCount > 0 Then
    WdOccAvg = WdOccSum / WdOccCount
If WdOccSumSqr <= WdOccCount * WdOccAvg ^ 2 Then
    WdOccSumSqr = WdOccSumSqr +
        (WdOccSumSqr - WdOccCount * WdOccAvg ^ 2) * -1 + 1
If WdOccCount > 1 Then
    WdOccStDev = Sqr((WdOccSumSqr - WdOccCount
        * WdOccAvg ^ 2) / (WdOccCount - 1))
If WdOccCount > 0 Then MonthlyData(rowMonthly, 3) = "Yes"
If WdOccCount > 0 Then
    MonthlyData(rowMonthly, 4) = WdOccCount
If WdOccSum > 0 Then MonthlyData(rowMonthly, 5) = WdOccSum
If WdOccAvg > 0 Then MonthlyData(rowMonthly, 6) = WdOccAvg
If WdOccMin < 1000000 Then
    MonthlyData(rowMonthly, 7) = WdOccMin
If WdOccMax > 0 Then MonthlyData(rowMonthly, 8) = WdOccMax
If WdOccStDev > 0 Then
    MonthlyData(rowMonthly, 9) = WdOccStDev

'Weekday Unoccupied
If WdUnoccCount > 0 Then
    WdUnoccAvg = WdUnoccSum / WdUnoccCount
If WdUnoccSumSqr <= WdUnoccCount * WdUnoccAvg ^ 2 Then
    WdUnoccSumSqr = WdUnoccSumSqr +
        (WdUnoccSumSqr - WdUnoccCount * WdUnoccAvg ^ 2) * -1 + 1
If WdUnoccCount > 1 Then
    WdUnoccStDev = Sqr((WdUnoccSumSqr - WdUnoccCount
        * WdUnoccAvg ^ 2) / (WdUnoccCount - 1))
If WdUnoccCount > 0 Then MonthlyData(rowMonthly, 10) = "Yes"
If WdUnoccCount > 0 Then
    MonthlyData(rowMonthly, 11) = WdUnoccCount
If WdUnoccSum > 0 Then MonthlyData(rowMonthly, 12) = WdUnoccSum
If WdUnoccAvg > 0 Then MonthlyData(rowMonthly, 13) = WdUnoccAvg
If WdUnoccMin < 1000000 Then
    MonthlyData(rowMonthly, 14) = WdUnoccMin
If WdUnoccMax > 0 Then MonthlyData(rowMonthly, 15) = WdUnoccMax
If WdUnoccStDev > 0 Then
    MonthlyData(rowMonthly, 16) = WdUnoccStDev

'Weekend Occupied
If WeOccCount > 0 Then
    WeOccAvg = WeOccSum / WeOccCount
If WeOccSumSqr <= WeOccCount * WeOccAvg ^ 2 Then
    WeOccSumSqr = WeOccSumSqr +
        (WeOccSumSqr - WeOccCount * WeOccAvg ^ 2) * -1 + 1
If WeOccCount > 1 Then
    WeOccStDev = Sqr((WeOccSumSqr - WeOccCount
        * WeOccAvg ^ 2) / (WeOccCount - 1))
If WeOccCount > 0 Then MonthlyData(rowMonthly, 17) = "Yes"
If WeOccCount > 0 Then
    MonthlyData(rowMonthly, 18) = WeOccCount
If WeOccSum > 0 Then MonthlyData(rowMonthly, 19) = WeOccSum
If WeOccAvg > 0 Then MonthlyData(rowMonthly, 20) = WeOccAvg
If WeOccMin < 1000000 Then  
    MonthlyData(rowMonthly, 21) = WeOccMin
If WeOccMax > 0 Then MonthlyData(rowMonthly, 22) = WeOccMax
If WeOccStDev > 0 Then  
    MonthlyData(rowMonthly, 23) = WeOccStDev

'Weekend Unoccupied
If WeUnoccCount > 0 Then  
    WeUnoccAvg = WeUnoccSum / WeUnoccCount
If WeUnoccSumSqr <= WeUnoccCount * WeUnoccAvg ^ 2 Then  
    WeUnoccSumSqr = WeUnoccSumSqr +  
        (WeUnoccSumSqr - WeUnoccCount * WeUnoccAvg ^ 2) * -1 + 1
If WeUnoccCount > 1 Then  
    WeUnoccStDev = Sqr((WeUnoccSumSqr - WeUnoccCount *  
        WeUnoccAvg ^ 2) / (WeUnoccCount - 1))
If WeUnoccCount > 0 Then MonthlyData(rowMonthly, 24) = "Yes"
If WeUnoccCount > 0 Then  
    MonthlyData(rowMonthly, 25) = WeUnoccCount
If WeUnoccSum > 0 Then MonthlyData(rowMonthly, 26) = WeUnoccSum
If WeUnoccAvg > 0 Then MonthlyData(rowMonthly, 27) = WeUnoccAvg
If WeUnoccMin < 1000000 Then  
    MonthlyData(rowMonthly, 28) = WeUnoccMin
If WeUnoccMax > 0 Then MonthlyData(rowMonthly, 29) = WeUnoccMax
If WeUnoccStDev > 0 Then  
    MonthlyData(rowMonthly, 30) = WeUnoccStDev

Next rowMonthly

'***Copies MonthlyData array back to workbook

Sheets("MonthlyPeriodStats").Select

For row = 1 To 16
    For col = 2 To 4
        Cells(row + 3, col).Value = MonthlyData(row, col)
    Next col

    'col 5
    If MonthlyData(row, col - 1) > 0 Then  
        Cells(row + 3, col) =  
            "=IF(C" & (row + 3) & "]"="Yes"," &  
                MonthlyData(row, col) & "]",""")"
    For col = 6 To 11
        Cells(row + 3, col).Value = MonthlyData(row, col)
    Next col

    'col 12
    If MonthlyData(row, col - 1) > 0 Then  
        Cells(row + 3, col) =  
            "=IF(J" & (row + 3) & "]"="Yes"," &  
                MonthlyData(row, col) & "]",""")"
    For col = 13 To 18
        Cells(row + 3, col).Value = MonthlyData(row, col)
    Next col
'col 19
If MonthlyData(row, col - 1) > 0 Then _
    Cells(row + 3, col) = 
        "=IF(Q" & (row + 3) & ")=""Yes""," & _
        MonthlyData(row, col) & ",""""
For col = 20 To 25
    Cells(row + 3, col).Value = MonthlyData(row, col)
Next col

'col26
If MonthlyData(row, col - 1) > 0 Then _
    Cells(row + 3, col) = 
        "=IF(X" & (row + 3) & ")=""Yes""," & _
        MonthlyData(row, col) & ",""""
For col = 27 To 30
    Cells(row + 3, col).Value = MonthlyData(row, col)
Next col

Next row

'***Wrap up
EndTime = Time
RunTime = EndTime - StartTime
Sheets("inputs").Select
Sheets("inputs").Range("e13").Value = RunTime

Application.ScreenUpdating = True
Application.Calculation = xlCalculationAutomatic

End Sub
Appendix C:
VBA for Decision Tool

Sub Trials()

Application.ScreenUpdating = False

Application.StatusBar = "Ok, let's rock this thing..."

Dim StartTime As Single, EndTime As Single, RunTime As Single
StartTime = Time

Dim i As Integer, Trial As Integer, RM As Byte
Dim NumTrials As Integer
Dim WindowLoop As Byte, WindowSize As Byte
Dim TrialNum As Byte, ThresholdLoop As Single

Dim Months() ReDim Months(1 To 12)
For i = 1 To 12
  Months(i) = Sheets("DailyData").Cells(49 + i, 1)
Next i

Dim Window() As Byte
ReDim Window(1 To 4)
For i = 1 To 4
  Window(i) = Sheets("Inputs").Cells(1 + i, 8)
Next i

Dim Prob() As Single
ReDim Prob(1 To 7)
For i = 1 To 7
  Prob(i) = Sheets("Inputs").Cells(2 + i, 11)
Next i

For ThresholdLoop = 1 To 7
  Sheets("DailyData").Range("B16") = Prob(ThresholdLoop)
  Select Case ThresholdLoop
  Case 1
    Sheets("T1").Select
    Range("B11:IK1010").Select
    Selection.ClearContents
  Case 2
    Sheets("T2").Select
    Range("B11:IK1010").Select
    Selection.ClearContents
  Case 3
    Sheets("T3").Select
    Range("B11:IK1010").Select
  End Select
Next ThresholdLoop
Selection.ClearContents
Case 4
    Sheets("T4").Select
    Range("B11:IK1010").Select
    Selection.ClearContents
Case 5
    Sheets("T5").Select
    Range("B11:IK1010").Select
    Selection.ClearContents
Case 6
    Sheets("T6").Select
    Range("B11:IK1010").Select
    Selection.ClearContents
Case 7
    Sheets("T7").Select
    Range("B11:IK1010").Select
    Selection.ClearContents
End Select

NumTrials = Sheets("Inputs").Range("NumberOfTrials")
Dim TrialsArray() As Long
ReDim TrialsArray(1 To NumTrials, 1 To 244) As Long
Rnd (-100871) 'Random number generator resets at beginning 'of each threshold loop

For Trial = 1 To NumTrials
    StaticRandomNumberGenerator
    For WindowLoop = 1 To 4
        WindowSize = Window(WindowLoop)
        Sheets("DailyData").Range("B4") = WindowSize
        Dim Mo As Byte
        For Mo = 1 To 12
            Sheets("DailyData").Range("B3") = Months(Mo)
            Application.Calculation = xlCalculationAutomatic
            For RM = 1 To 5
                TrialsArray(Trial, Mo + (WindowLoop - 1) * 12 + (RM - 1) * 49) = Sheets("DailyData").Cells(14, RM)
            Next RM
        Next Mo
    Next WindowLoop
EndTime = Time
RunTime = EndTime - StartTime
Application.StatusBar = "Complete: " & Format(((ThresholdLoop - 1) * NumTrials + Trial) / (NumTrials * 7), "0%") & " (Threshold " & ThresholdLoop & ", Trial " & Trial & ") Elapsed Run Time: " & Format(RunTime, "h:mm:ss")

Next Trial
Select Case ThresholdLoop
Case 1
    Sheets("T1").Range(Cells(11, 2), _
Cells(10 + NumTrials, 245) = TrialsArray()
Sheets("T1").Range("B16").Select
Case 2
Sheets("T2").Range(Cells(11, 2),
    Cells(10 + NumTrials, 245)) = TrialsArray()
Sheets("T2").Range("B16").Select
Case 3
Sheets("T3").Range(Cells(11, 2),
    Cells(10 + NumTrials, 245)) = TrialsArray()
Sheets("T3").Range("B16").Select
Case 4
Sheets("T4").Range(Cells(11, 2),
    Cells(10 + NumTrials, 245)) = TrialsArray()
Sheets("T4").Range("B16").Select
Case 5
Sheets("T5").Range(Cells(11, 2),
    Cells(10 + NumTrials, 245)) = TrialsArray()
Sheets("T5").Range("B16").Select
Case 6
Sheets("T6").Range(Cells(11, 2),
    Cells(10 + NumTrials, 245)) = TrialsArray()
Sheets("T6").Range("B16").Select
Case 7
Sheets("T7").Range(Cells(11, 2),
    Cells(10 + NumTrials, 245)) = TrialsArray()
Sheets("T7").Range("B16").Select
End Select
Next ThresholdLoop
Application.StatusBar = False
Application.ScreenUpdating = True
Application.Calculation = xlCalculationAutomatic
EndTime = Time
RunTime = EndTime - StartTime
Sheets("T1").Range("A9") = RunTime
End Sub

Sub StaticRandomNumberGenerator()
    Dim Col As Integer, Row As Integer
    Dim NumRows As Integer
    Dim StartDate As Long, EndDate As Long
    StartDate = Sheets("Inputs").Range("FirstTestMo").Value2
    EndDate = Sheets("Inputs").Range("LastAvailDate").Value2
    NumRows = EndDate - StartDate + 1
    Dim RandomArray() As Single
    ReDim RandomArray(1 To NumRows, 1 To 32)
    '*** Response Length
    For Col = 1 To 32
        For Row = 1 To NumRows
            RandomArray(Row, Col) = Rnd()
        Next Row
    Next Col
End Sub
Sub GenerateDynamicRandomNumbers()

    Application.ScreenUpdating = False

    Range("ProbResponseLength").Formula = "=RAND()"
    Range("ProbResponseType").Formula = "=RAND()"
    Range("ProbBaseAdjMult").Formula = "=RAND()"

    Application.ScreenUpdating = True
    Application.Calculation = xlCalculationAutomatic

End Sub
Appendix D: Building Daily Usage and TTM

**Building 1, Meter 1**

\[ y = 0.7738x - 25527 \]

- Oct-06
- Oct-07
- Oct-08
- Oct-09
- Oct-10

**Building 1, Meter 2**

\[ y = -1.6061x + 72793 \]

- Oct-06
- Oct-07
- Oct-08
- Oct-09
- Oct-10
Building 3, Meter 1

\[ y = -0.2853x + 31622 \]

Building 4, Meter 1

\[ y = -2.4455x + 105374 \]
Building 5, Meter 1

\[ y = 0.1456x - 831.5 \]

Building 5, Meter 2

\[ y = -0.1049x + 5188.7 \]
Appendix E: Normal Probability Plots

Figure 45: Building 1, Meter 1, normal probability plots by period type

- **Weekday Occupied, 16-mo**
- **Weekday Unoccupied, 16-mo**
- **Weekend Occupied, 16-mo**
- **Weekend Unoccupied, 16-mo**
Figure 46: Building 1, Meter 2, normal probability plots by period type

- **Weekday Occupied, 16-mo**

- **Weekday Unoccupied, 16-mo**

- **Weekend Occupied, 16-mo**

- **Weekend Unoccupied, 16-mo**
Figure 47: Building 2, Meter 1, normal probability plots by period type

**Weekday Occupied, 16-mo**

-4.0000
-2.0000
0.0000
2.0000
4.0000

0 2,000 4,000 6,000 8,000

**Weekday Unoccupied, 16-mo**

-4.0000
-2.0000
0.0000
2.0000
4.0000

0 1,000 2,000 3,000 4,000 5,000

**Weekend Occupied, 16-mo**

-4.0000
-3.0000
-2.0000
-1.0000
0.0000
1.0000
2.0000
3.0000
4.0000

0 1,000 2,000 3,000 4,000 5,000

**Weekend Unoccupied, 16-mo**

-4.0000
-3.0000
-2.0000
-1.0000
0.0000
1.0000
2.0000
3.0000
4.0000

0 1,000 2,000 3,000 4,000 5,000
Figure 48: Building 2, Meter 2, normal probability plots by period type

- **Weekday Occupied, 16-mo**

- **Weekday Unoccupied, 16-mo**

- **Weekend Occupied, 16-mo**

- **Weekend Unoccupied, 16-mo**
Figure 49: Building 3, Meter 1, normal probability plots by period type

Weekday Occupied, 16-mo

Weekday Unoccupied, 16-mo

Weekend Occupied, 16-mo

Weekend Unoccupied, 16-mo
Figure 50: Building 4, Meter 1, normal probability plots by period type

**Weekday Occupied, 16-mo**

**Weekday Unoccupied, 16-mo**

**Weekend Occupied, 16-mo**

**Weekend Unoccupied, 16-mo**
Figure 51: Building 5, Meter 1, normal probability plots by period type

Weekday Occupied, 16-mo

Weekday Unoccupied, 16-mo

Weekend Occupied, 16-mo

Weekend Unoccupied, 16-mo
Figure 52: Building 5, Meter 2, normal probability plots by period type

- **Weekday Occupied, 16-mo**
- **Weekday Unoccupied, 16-mo**
- **Weekend Occupied, 16-mo**
- **Weekend Unoccupied, 16-mo**
References


