Evaluating Impact of Retrofit Programs on Commercial Buildings: Results from the Energize Phoenix Project

Karthik Thalappully, Agami Reddy, Oscar Nishizaki,
Marcus Myers, Michael Dalrymple and Patrick Phelan
Arizona State University, Tempe, Arizona.

ABSTRACT

Energize Phoenix (EP) is a three year energy efficiency program conducted by a joint collaboration between the City of Phoenix, Arizona State University and a large electricity provider. The intent was to improve energy efficiency in residential, multi-family, commercial/industrial buildings located in a portion of the city of Phoenix around the light rail corridor. There are several facets to the EP program with engineering-based verification of the energy savings due to commercial buildings upgrades being one of them, and the focus of this paper. Various issues had to be addressed such as incomplete data, spurious data behavior, multiple upgrade projects in the same facility, weather normalization using well-accepted change point models applied to utility bill data, and baseline model uncertainty. Considering the nature and characteristics of utility bill data, the evaluation methodology initially adopted was labor intensive and involved a manual data screening procedure of projects on an individual basis and then one-by-one baseline modeling and savings assessment. An automated process was developed in order to reduce the labor required to analyze and update energy savings in hundreds of buildings on a periodic basis. This paper presents results of over 200 completed upgrade projects comparing measured savings with those predicted by the contractors prior to the upgrades. Reasons for these differences are discussed, and follow-up investigations into this discrepancy are also described. Preliminary savings uncertainty is also reported. This paper ends with conclusions and suggestions for further investigation needed to improve the accuracy and reliability of determining energy savings in allied large scale energy efficiency programs.

Introduction

Energize Phoenix (EP) is a three year energy efficiency program led by a joint collaboration of three major institutions –the City of Phoenix, Arizona State University and Arizona Public Service (APS) – the state's largest electricity provider. The main goal of the program was to improve the energy efficiency in the buildings located around the Phoenix light rail corridor and to create jobs. The participating buildings included residential, multi-family, commercial/industrial etc. The program is one of 41 from across the United States which are supported by the U.S Department of Energy's Better Buildings Neighborhood Program and the American Recovery and Reinvestment Act of 2009 in order to test new models for scaling energy efficiency and to create jobs. Historically, energy efficiency programs have faced a trio of interconnected forces — technical, economic, and socio-behavioral — which continue to hinder mass-market scaling. Inter-disciplinary research underway for EP is aimed at understanding and helping resolve these barriers. Research projects cover numerous facets (such as behavioral and attitudinal differences between participating and non-participating homeowners and business owners, contractor marketing methods, the effects of energy feedback devices coupled with other education or budgeting information, spatio-temporal trends in participation rates, econometric modeling of savings, and economic impact analysis), all of which are meant to study the above influences and aimed at helping energy efficiency programs realize their full potential.

EP is a contractor-driven program in which participants receive incentives based upon the amount of kilowatt-hours estimated to be saved in the first year by the energy conservation measures installed. Thus, the prospective buildings were not chosen by us, instead, the task of initiating contact and convincing customers to join the EP program was with the contractor. Incentives can range up to 100% of incremental or project costs. This paper is narrowly focused in its scope on research work performed to better understand the accuracy of energy savings estimates in commercial building upgrades within the framework of the EP program.

Objectives

The primary objective of the EP commercial team was to analyze the data and to quantify the energy savings achieved in the commercial buildings which underwent upgrades incentivized by the EP program. These savings were then compared to the stipulated savings or savings predicted by the energy contractors during the project sales process. Note that the energy contractor estimated savings using either custom audits or prescriptive guidelines that rely on equipment count (such as lights).In addition, the "solutions for business" approach, meant for small businesses only, was based on standard software and other proprietary tools supplied by a 3rd party contractor. This comparison helps to assess the overall effectiveness of the upgrades and the accuracy of the savings estimates of the program as a whole. Other secondary issues were also investigated; for example, whether certain contractors tended to consistently over-estimate savings as compared to others, and reason for doing so.

There were hundreds of commercial buildings which underwent energy upgrades through the program. These projects varied vastly in characteristics like business type, size, etc. The monthly utility bill analysis approach was deemed to be the only realistic method to determine savings in a program this big with limited personnel involved in the measurement and verification (M&V) process. Further, since EP was an ongoing program and contractors often specialized in certain types of retrofits, buildings underwent upgrades on a continuing basis, and so savings calculations had to be redone at frequent intervals. We are sent utility bill data for all projects (old as well as recent) on a quarterly basis, and it was logical for us to recalculate and update the savings for all buildings every 3 months. This prompted us to define an additional objective, namely to simplify and automate the savings analysis methodology as far as possible so that future energy conservation programs similar to EP could reduce M&V analysis costs.

Overall Approach

Because of time constraints, the baseline electricity consumption prior to the implementation of energy upgrades could not be determined by in-situ measurement. Hence, the whole building analysis approach, which is one of the four general M&V approaches widely followed by the professional M&V community (see for example, ASHRAE Guideline 14, 2002 or IPMVP, 2010) was adopted. The approach involves relying on a whole year of utility bill data prior to the upgrade to establish a baseline model of energy use against monthly mean outdoor temperature. Such monthly utility bill data was made available from the APS customer billing database. The model is then applied to measured outdoor temperature during the post-upgrade period and the sum of the monthly differences between these model predictions and the actual measured utility bills during the post-upgrade period constitutes the upgrade energy savings. The entire process is described in more detail in the Appendix.

Because of the error introduced in such a general approach (called Level 1 analysis) which does not involve inspecting the buildings individually, it was decided to conduct a limited number of in-depth analyses in buildings where large discrepancies were found between measured and contractor-predicted savings. This approach (referred to as Level 2) would provide some degree of credibility in our speculation as to the observed differences, and allow us to correct the data as appropriate. Due to the large number of projects, our approach was to sample a sub-set of the completed upgrade projects and verify the savings estimated by the contractor through follow-up field visits, installing in-situ equipment and monitoring for a relatively short period of time. The degree of over- or under- prediction of the savings could then be determined more accurately, and the causes for any such discrepancies identified. This would provide useful feedback to APS and to the contractor, and suggest ways by which future upgrade savings estimations can be improved.

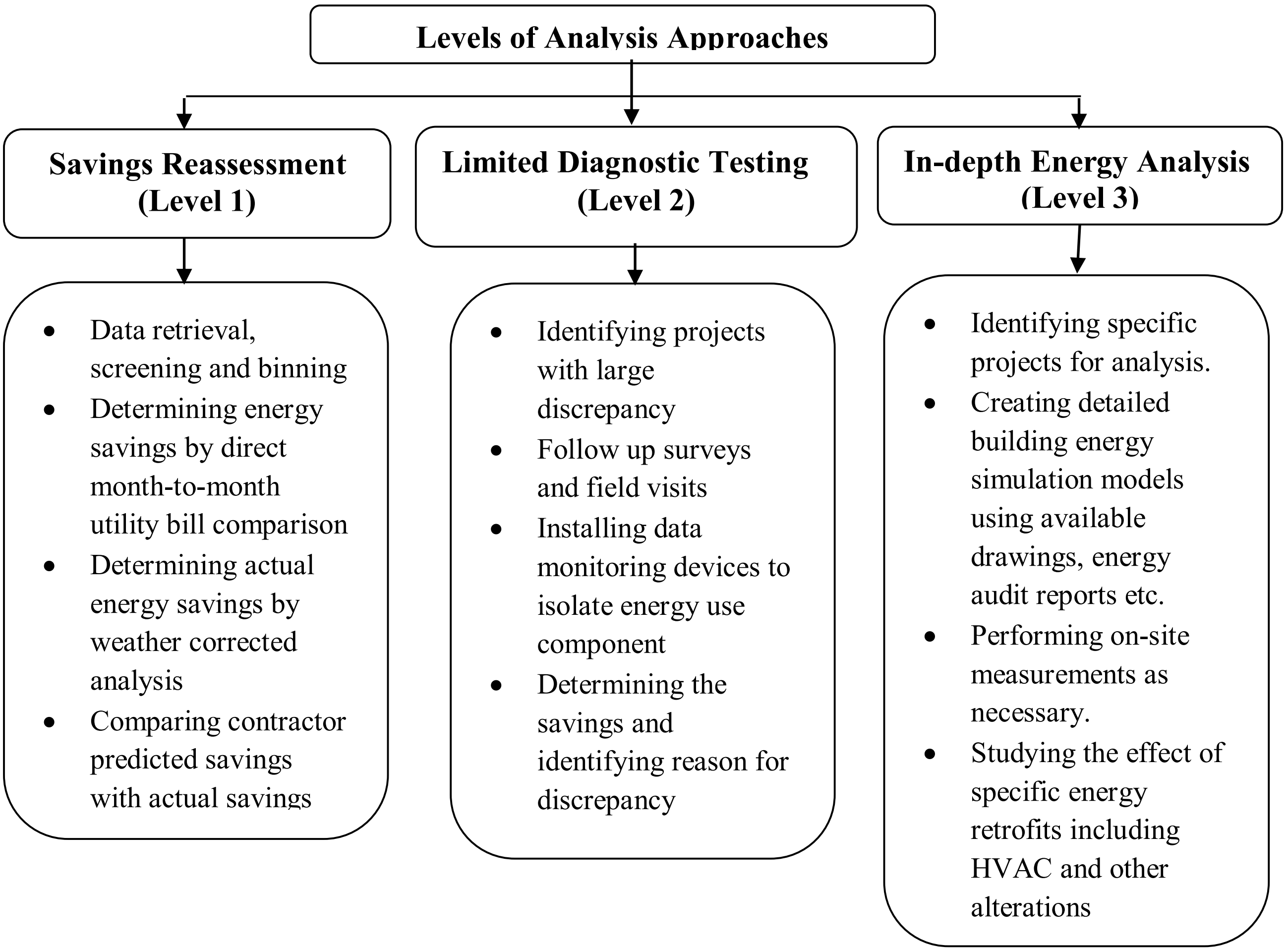

Figure 1. The three levels of analysis

Finally, Level 3 involved an in-depth energy analysis of a few selected projects so as to evaluate energy savings and to provide recommendations for additional potential energy conservation measures. Level 3 analysis involved developing a calibrated detailed simulation model of the energy use in the building based on owner-provided architectural drawings, energy audit reports, usage data and project applications. This is consistent with another standard M&V approach described in such documents as ASHRAE Guideline 14 (2002) and IPMVP (2010). Sub-monitoring the energy use and indoor environment was also done to calibrate the model. The primary objective was to determine quantitatively the effect of individual energy efficiency upgrades on overall energy consumption and to identify other possible energy conservation measures (ECMs). Figure 1 depicts these three levels of analysis in a succinct manner. This paper primarily presents the results of our Level 1 analysis, with follow-up papers anticipated to report on the results of the other two levels.

Methodology

Data screening and binning

The first step involved ascertaining consistency of energy use over the years. This was conveniently done by simply generating time series plots (see Figure 2) of historic utility bills, and looking at them visually. Some of the projects showed considerable variation in usage pattern which made it necessary to manually screen all individual projects for data quality. This also led to the decision of using only one year of data immediately prior to the upgrade as the baseline period since, as is well known, energy use patterns in commercial buildings tend to change over time.

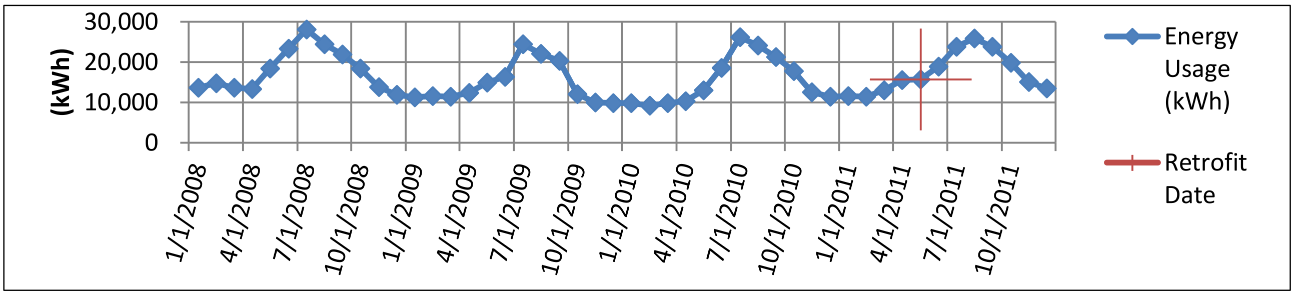

Figure 2. Time series plot of monthly utility data for over three years before the upgrade and one year after upgrade for a specific EP building

The data visualization step allowed identification of anomalous behavior and grouping of buildings into bins, as illustrated in Figure 3. Bin A consisted of buildings where there were missing or inadequate pre-upgrade data (i.e., less than twelve utility bills). Buildings with abnormal data patterns were placed in Bin B. Three of the common generic cases encountered are illustrated in Figure 3.Some buildings exhibited an increase in energy use after the upgrade, some had abnormal spikes, and others had markedly different seasonal variation patterns. Bin C consisted of buildings which did not have at least six months of post-upgrade data, in which case the calculation of energy savings was deferred until more utility bill data was forthcoming. Finally, those buildings which did not fall in any of the above three bins, were placed in Bin D for which the savings were determined. Additional manual screening criteria for data quality had to be empirically framed as shown in Table 1. For example, if predicted savings were less than 1% of the energy use, our analysis procedure was deemed to be unsuitable. An example of anomalous behavior which warranted placing a project in Bin B was a case in which the audit estimated savings exceeded the total energy use of the building. This screening process made the whole analysis labor intensive, but it needed to be done only once per project.

Since there were several buildings which fell into Bin B, phone calls to the facility managers or owners of several of these buildings were also undertaken in order to identify possible reasons and to reconcile the odd behavior. If the behavior could be explained convincingly, these buildings were moved to Bin D, otherwise they were moved to Bin A. A possible factor causing some of the anomalous behavior could be attributed to the fact that we were unable to perform account matching with the master meter of the facility. Such data was not made available to us due to privacy reasons.

Table1. Data screening and binning criteria employed for screening

Bin A (Excluded projects)	**Bin B (Projects requiring further analysis)**	**Bin C (Projects awaiting more data)**	**Bin D (Projects which analysis were done)**
1) When savings estimated are less than 1% of the pre retrofit utility consumption 2) Less than twelve months of pre-retrofit data available 3) Observed discrepancy in time series data could not be resolved	1) Unexplained increase in the pre or post retrofit consumption 2) Contractor claimed savings are greater than 100 % of energy use 3) The post retrofit energy use has gone up 4) Pre and Post retrofit patterns are different	Less than 6 months of post retrofit data	If the project does not fall in any of the other categories

Automation of savings calculation

Savings were estimated by comparing the energy consumption between corresponding months of pre and post upgrade periods. The billing cycle was assumed to match calendar months due to lack of meter read dates, which introduces some error in our analysis. A large number of projects showed weather dependency where the total energy consumption was influenced by cooling and heating loads of the building. Thus, the influence of the weather had to be taken into account for these projects in order to properly estimate the upgrade savings.

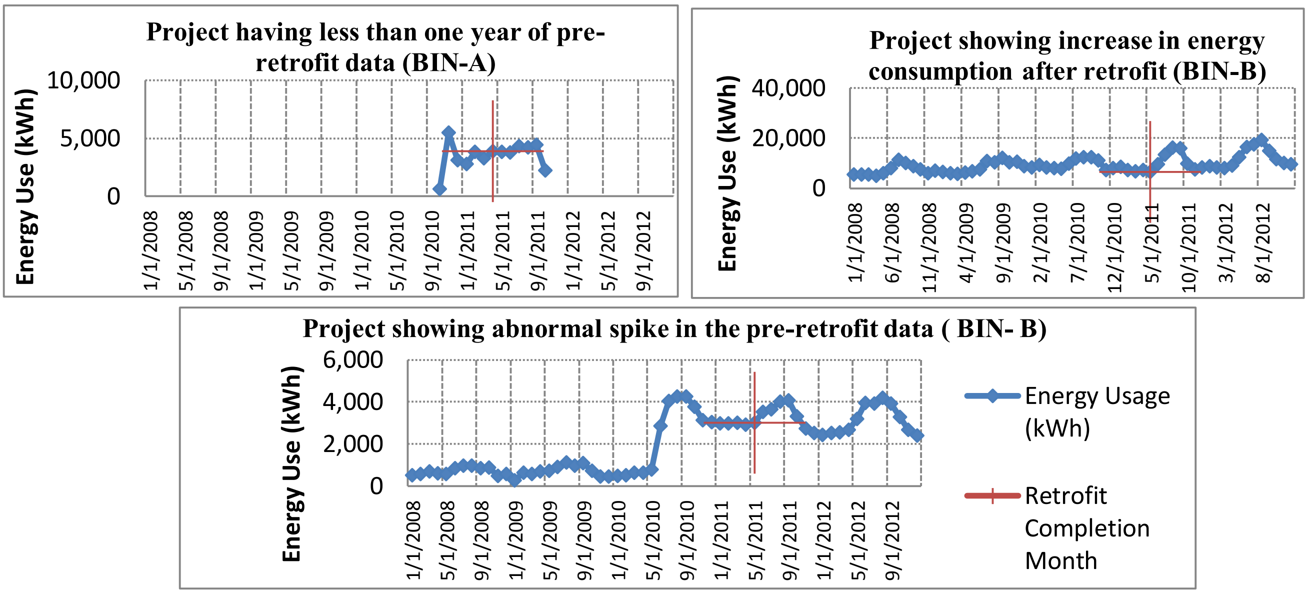

Figure 3. Illustrative examples of abnormal data. The data screening behavior pertinent to Bin A and Bin B were identified during the process.

There were several buildings which qualified for EP incentives involving multiple energy upgrades. These were treated as single projects using the simple approach illustrated in Figure 4. The data period in between the first and the last upgrades was simply excluded from the analysis since in most cases these multiple upgrades were done within a few months of each other. All the upgrades were treated as one single upgrade with the post upgrade period assumed to start after the last upgrade was completed. The sum total of all the contractors' savings estimates for the building was taken to be the overall predicted savings.

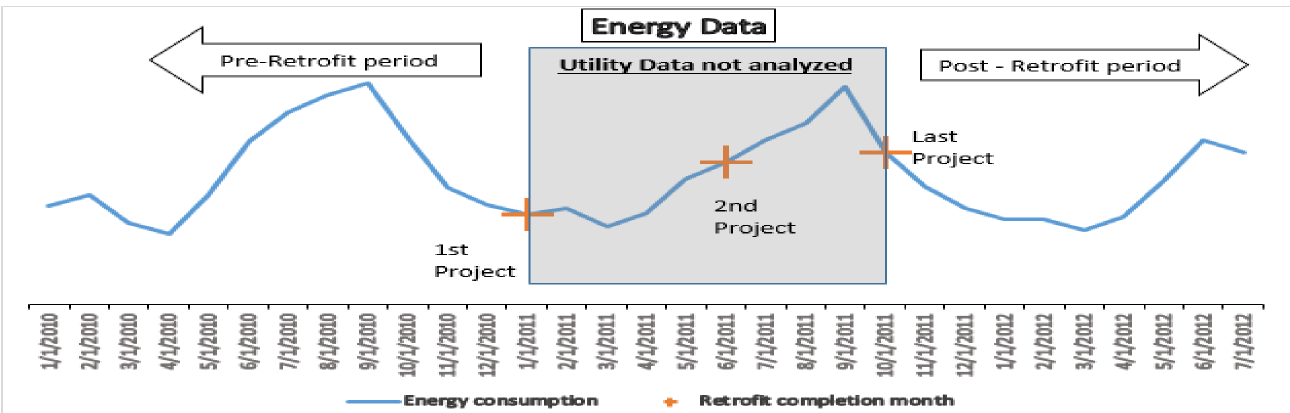

Figure 4. A hypothetical building showing multiple retrofit projects. Energy savings were simply determined using pre retrofit and post retrofit periods as shown.

As the number of projects increased and since savings had to be recalculated at quarterly intervals as more data was forthcoming, it was critical to automate the process as much as possible. The automation scheme which evolved is shown in Figure 5. Note that there are still two steps which require manual screening.

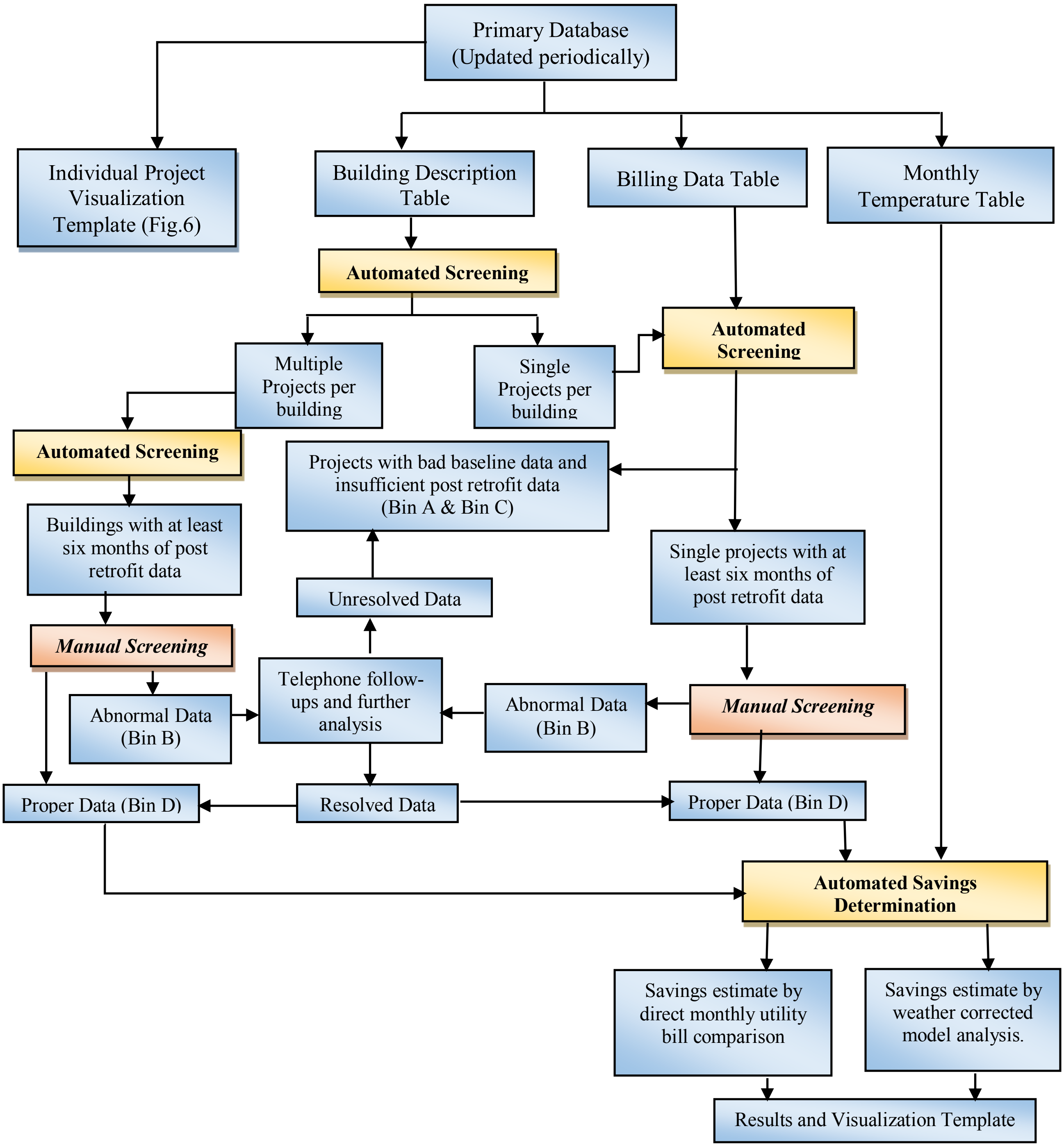

Figure 5. Flowchart of the automated routine developed to determine energy savings from numerous retrofitted buildings in the framework of EP program

To facilitate the manual screening in the automation process, a visual template (shown in Figure 6) was developed. This involved generating scatter plots of energy use versus outdoor temperature and annual time series plots superimposed on each other.

The methodology for developing the baseline model is described in the Appendix. It is consistent with the modeling procedures advocated in the engineering literature involving identifying the best change point regression model among several different model formulations with outdoor temperature as the independent variable. A FORTRAN program was developed specifically for the purpose of the EP commercial building analysis effort which incorporated the widely used Inverse Modeling Toolkit (IMT) computer code (Kissock, Haberl and Claridge 2002) as a subroutine. The program reads the utility bill data for a specific building along with outdoor temperature, and assigns it to the pertinent bin. If the building falls into Bin D, the program then identifies the best change point model among several possible functional forms, calculates savings for that building, and does this for all the buildings in the database. The total savings are then determined along with the contractor estimated savings. Finally, the automated routine generates pertinent summary statistics and graphics of the entire program savings.

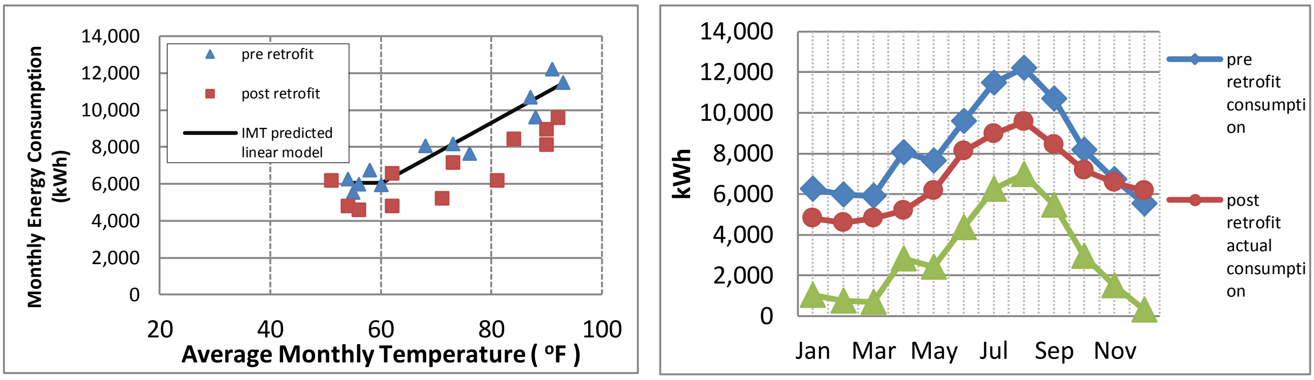

Figure 6. Plots from the visualization template meant as an aid to perform manual screening of data

Analysis Results

As of March 2013, 557 retrofit projects were completed. The distribution of the adopted energy conservation measures (ECM) is shown in the Figure 7 .Of all the projects, 141 projects fell into Bin D (Table 2). The energy savings determined are summarized in Table 2 and also plotted in Figure 8. Note that:

(i) There is a major discrepancy between the total savings predicted by the contractors and those determined from our weather normalized savings calculation often referred to as "measured" savings in the M&V literature. While the former is found to be 8.1% of the baseline energy use, the 'measured' savings fraction was 5.2%, a significant difference.

(ii) There is a large difference between weather normalized savings and those savings determined by direct pre-post utility bill comparison (5.2% versus 3.2%). The difference in outdoor temperatures between the years 2010 and 2011 was not large, but that between 2011 and 2012 was significant and could explain this difference.

(iii) Figure 9 shows the measured savings percentage (i.e. energy savings divided by baseline energy use) on an annual basis for all individual projects along with the associated fractional uncertainty (i.e. energy savings uncertainty divided by energy savings). The uncertainties of the change point models, characterized by their coefficient of variation – root mean squared error (CV-RMSE), are generally large. However, the baseline model is used to predict energy use each month for the 12 months of the year and so the uncertainty of the summed values are lower. The relevant formulae are given in various publications (Reddy and Claridge, 2000 or ASHRAE 14, 2002). A follow up paper will report uncertainty in more detail as well as on the whole portfolio of buildings of the EP program.

Table 2. Summary of analyzed commercial projects

Number of Completed Projects (As on 3/30/2013)	**557**
Sum of contractor estimated annual energy savings (kWh)	**45.2 x 10^6**
Total floor area of all completed projects (sq. ft.)	**20.4 x 10^6**
Number of projects in Bin A	**6**
Number of projects in Bin B	**73**
Number of projects in Bin C	**315**
Number of projects in Bin D	**141**

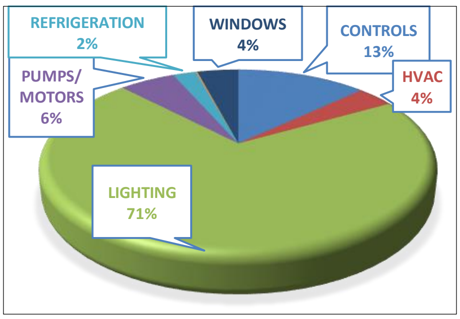

Figure 7. Distribution of ECM types for 557 projects

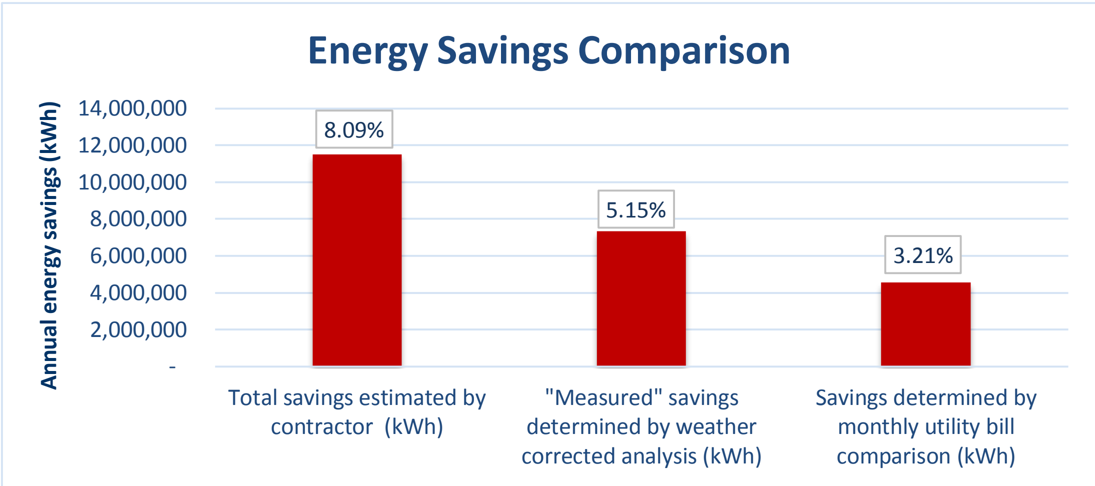

Figure 8. Energy savings in 141 retrofit projects estimated by different methods.

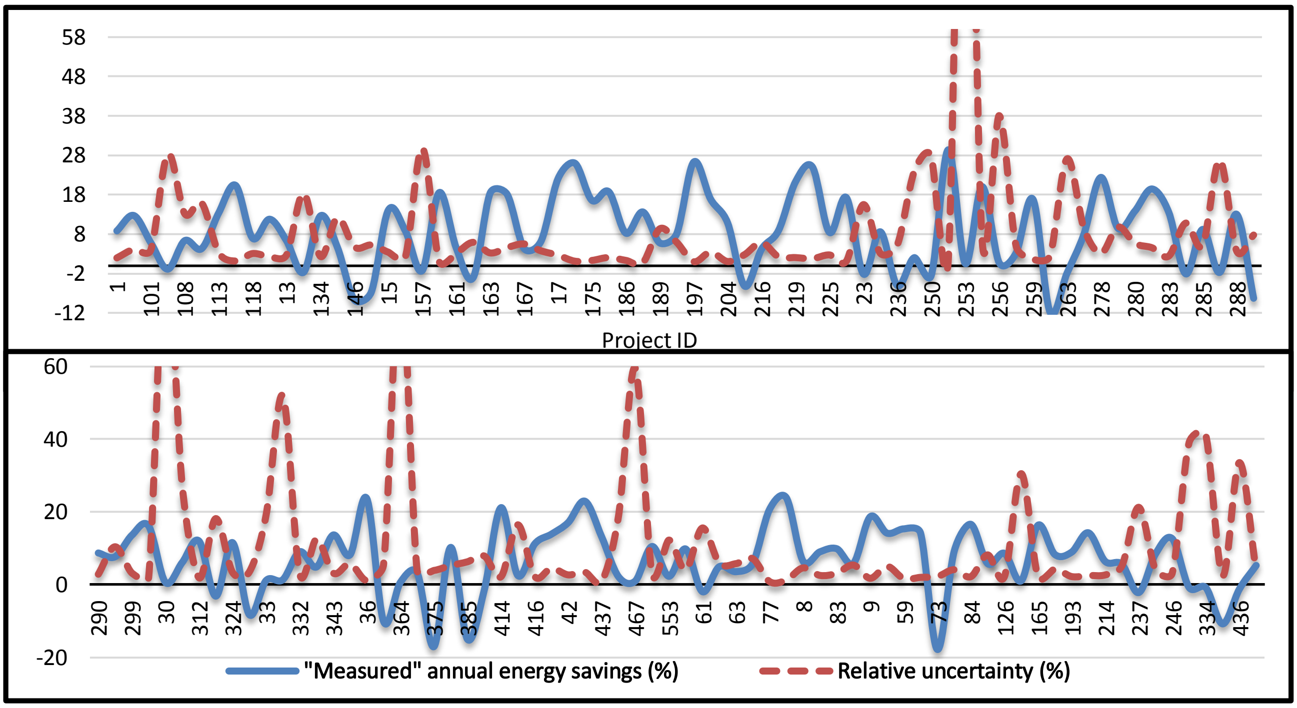

Figure 9. Plot depicting differences in "measured" energy savings percentage and the associated relative uncertainty for individual projects. The relative uncertainties are generally small though a number of projects have large uncertainties which need further investigation.

Follow-up Investigations

Estimated and measured savings distributions

Possible causes for the discrepancy stated under (i) above were investigated. Figure 10 is a plot of the distributions in annual energy savings fraction for the 141 projects estimated by the contractors and those determined by the weather normalized approach. The estimated savings fraction have a noticeable wide distribution across the various projects, exhibiting a long positive tail. On the other hand, the measured savings have a tighter distribution, and peak around 9% savings. However, many of the projects show negative savings which is probably the reason why the total "measured" energy savings fraction turns out to be only around 5%. A possible reason for this discrepancy could be due to bias that has been introduced in our analysis due to the visual manner by which buildings are sorted to the various bins described above. We repeated the savings analysis without such a binning for 220 projects (which had complete baseline data and more than six months of post retrofit data) and found that the measured savings fraction varied very little (from 5.2% to 4.9 %) while the contractor predicted savings showed a larger increase (from 8.1% to 9.6%). Thus, our analysis procedure does not seem to introduce any bias. This preliminary result is an important one since it could reduce /eliminate much of the effort expended in the manual screening process in future energy conservation programs.

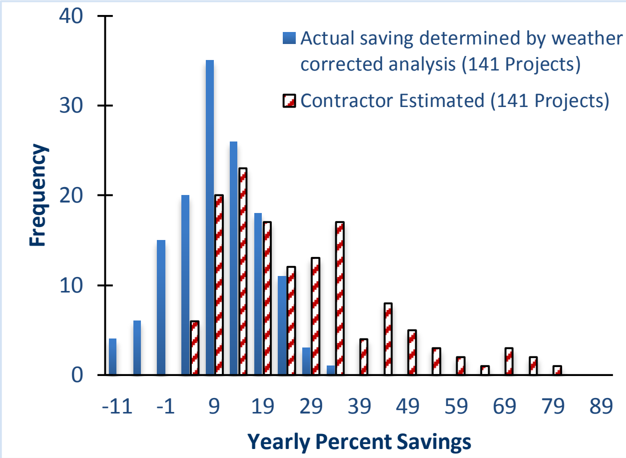

Figure 10. Comparison of frequency distribution of annual savings as a percentage of baseline consumption. Data is from 141 projects with

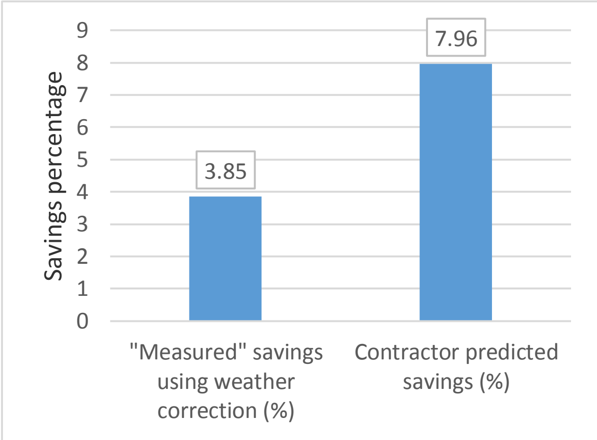

Figure 11. Comparison of contractor predicted savings and actual weather corrected savings fraction for lighting only projects (total 64 projects)

Possible causes of differences

In an effort to isolate the cause for the discrepancy between contractor predicted and "measured" savings, lighting-only upgrade projects were studied because they were simpler to analyze. Figure 11 provides a direct comparison of the estimated and measured savings percentages. While the former is close to 8%, the latter is close to 4%, a 50% discrepancy, which reflects our analysis results for the larger data set as well. The root causes for this discrepancy are discussed below.

In the case of a lighting project, the contractor-predicted savings were calculated as the kW reduction multiplied by the number of hours of operation. The kW reduction appears fairly straightforward since it entails counting the number of fixtures and using engineering formulae to account for ballast and other effects. The number of hours, on the other hand, is an estimate, often supplied by the building owner. A study is ongoing to investigate this source of error using data loggers at several facilities. The results of this study will be reported in a subsequent paper.

A second potential source of estimation error would arise from inaccurate assessment of pre-upgrade equipment conditions. Field measurements on one project determined that, while the owner and/or contractor had assumed that all existing ballasts consisted of older, inefficient magnetic technology, at least some of the ballasts had been replaced with newer electronic versions during regular maintenance as ballasts had burned out. While a 100% pre-upgrade audit is not a cost-effective solution, appropriate sampling could improve outcomes.

Another cause for savings discrepancies could be due to the quality of the utility bill data itself and how it was designated in the database. A field visit was made to another facility where estimated savings were over 100% of the baseline energy use, and it was found that the utility bills provided were only from one electric meter while the facility had four electric meters. Such discrepancies would greatly skew any analysis results, and so procedures must be put in place to ensure proper quality control in future energy conservation programs. Yet another reason for this discrepancy could be due to both snap back effect and due to energy creep (gradual increase in installed plug loads) in a facility after the retrofits were completed. This issue is also under investigation and will be reported in a subsequent paper.

Contractor bias in savings estimation

Another investigation involved determining whether certain contractors tended to consistently over-estimate energy savings. While there were 24 different contractors in total, there were nine who undertook numerous projects or projects with large energy savings. The results of this study are summarized in Figure 12. Notice, for example, that contractor #2 and #5 performed 34 and 36 projects respectively, and consistently over-estimated savings to a large degree. The ratios of measured savings to predicted savings for the 34 individual projects attributed to contractor #2 are shown in Figure 13. The causes of this result are also being investigated.

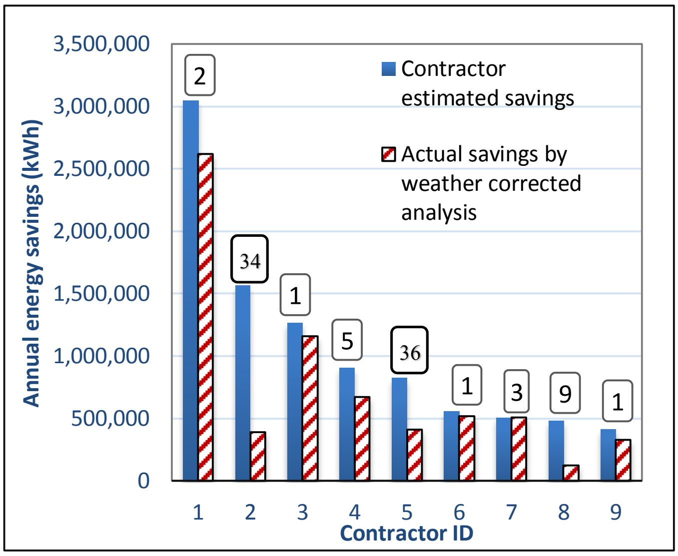

Figure 12. Performance comparison for nine contractors. The number of projects are indicated in the graph.

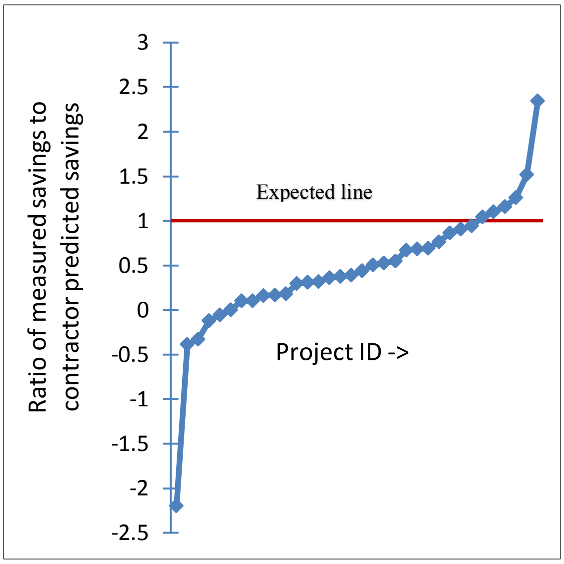

Figure 13. Savings ratio breakup for the 34 projects completed by contractor # 2 (Figure 12)

Summary and Suggestions for Future Projects

In summary, the major conclusion of the EP commercial analysis effort is that while contractor estimated savings fraction were around 8% of the baseline energy use for the entire program till March 2013, the measured savings fraction were only 5% (as of February 2013). Though these numbers may change as additional buildings are analyzed, this trend merits further investigation. If energy efficiency programs are to be scaled substantially, large portfolio financing is one logical path to reach scale. Financing sources need predictable returns in order to invest without requirements for the high risk premiums warranted by uncertainty. We have suggested possible means of reconciling estimated versus actual performance, either by installing data loggers or by field visit surveys. The proliferation of interval data from smart meters also opens up new possibilities for increasing estimation accuracy at the individual building level and through analysis of "Big Data" at the program level. We have also found that contractor bias accounts for some, if not much of the observed differences in savings. One suggestion is that contractors be provided with utility bills of the facility at the time of estimating savings since it would eliminate the very high estimated savings fractions found during our analysis. We would also advice that each and every building data be screened in order to identify and remove spurious data spikes and patterns even though we have reported in this paper that the difference seems to be small at the program level if such individual screening is not done. Further, this study suggests that it is imperative to perform weather normalization in order to predict savings in a more realistic manner. Finally, we conclude that change point models used for weather normalization generally have relative uncertainties lower than the savings themselves at the individual building level though this is not true for certain number of buildings. The uncertainty would still be even lower at the portfolio level, and this aspect will be reported in a subsequent paper.

References

ASHRAE Guideline 14 (2002), ASHRAE Guideline 12-2002. Measurement of Energy and Demand Savings. American Society of Heating, Refrigeration and Air-Conditioning Engineers, Atlanta, GA.

Haberl J.S. and Culp, C. (2007). "Measurement and Verification of Energy Savings," in Energy Management Handbook (Edited by W.C. Turner), Fairmont Press.

IPMVP (2010). International Performance Measurement and Verification Protocol, Volume 1. Efficiency Valuation Organization, U.S. Department of Commerce, Springfield, VA.

Kissock, J. K.,Haberl, J. S. and Claridge, D. E (2002) "Development of a Toolkit for Calculating Linear, Change-Point Linear and Multiple-Linear Inverse Building Energy Analysis Models," ASHRAE RP 1050, American Society of Heating, Refrigeration and Air-Conditioning Engineers, Atlanta, GA.

Kissock, J.K., Reddy, T.A. and Claridge, D.E. (1998). "Ambient-Temperature Regression Analysis for Estimating Retrofit Savings in Commercial Buildings," ASME Journal of Solar Energy Engineering, vol. 120, no. 3, p. 168.

Reddy, T.A., Claridge, D.E. (2000), Uncertainty of "Measured" Energy Savings from Statistical Baseline Models. HVAC&R Research, Journal Vol. 6, no. 1, pp. 3-20, American Society of Heating, Refrigeration and Air-Conditioning Engineers, Atlanta, GA.

Appendix: Baseline Model Development and Uncertainty

The savings methodology adopted is consistent with the one suggested in the professional literature (see for example, Haberl and Culp 2007).The process includes the following steps:

1. Acquire monthly energy use data (from utility bills) and data on influential variables (limited in this study to outdoor dry-bulb temperature) during the pre-upgrade period.
2. Develop a regression model of pre- upgrade energy use as a function of influential variables- this is the "baseline model".
3. Acquire date of energy use (from utility bills) and influential variables during post upgrade period.
4. Use the values of influential variables from the post upgrade period (from step 3) in the pre upgrade model (from step 2) to predict how much energy the building would have consumed on a monthly if it had not been upgraded.
5. Subtract measured post upgrade energy use (step 3) from the predicted pre-upgrade energy use (step 4) to estimate savings on a monthly basis.
6. Sum the individual monthly savings to determine cumulative (or annual) savings and percentage savings.
7. Compare the model goodness-of-fit (using the coefficient of variation of the root mean square error or CV-RMSE) with the percentage savings determined.

The model approach is statistical in nature, involving identifying a regression model of monthly energy use against monthly mean outdoor temperature using the monthly mean temperature model (Kissock, Reddy and Claridge 1998). The ambient temperature is chosen as the only independent variable because of the easy availability of the data, the difficulty in acquiring other data, and to avoid statistical difficulty arising from a small data set (only 12 data points) and multi-collinearity with environmental indices such as ambient humidity and solar radiation.

Another significant parameter to be considered is the uncertainty in the baseline model for a specific site characterized by the CV-RMSE (coefficient of variation of the root mean square error) of the model. This allows direct insights into the statistical soundness of the associated savings deduced. The CV-RMSE is a rough measure of the fractional (or percentage) uncertainty in the baseline model compared to the mean baseline energy use. A 10% CV-RMSE would imply that model uncertainty is 10% of the mean annual pre-upgrade energy use. If the savings fraction is less than the CV-RMSE then one is unjustified statistically in placing too much confidence in the associated savings estimated at that site. Adding this filter criterion to the analysis would have further reduced the total number eligible projects within the EP program. So for a single project all models were evaluated as shown in Figure A1, and the model with the least CV-RMSE was chosen as the best fit baseline model.

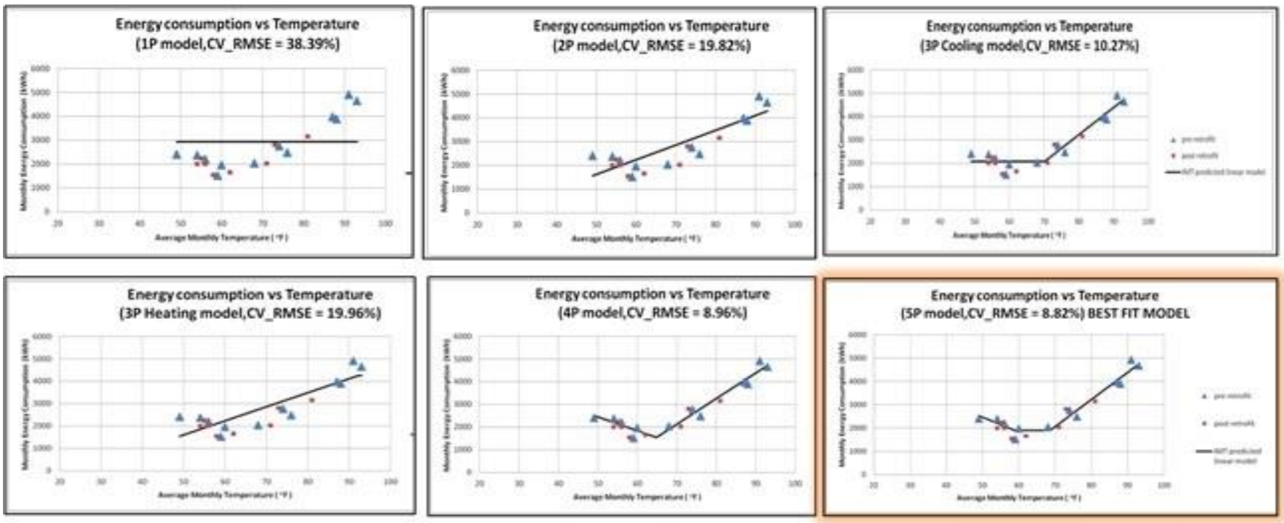

Figure A1. Process of determining the best fit regression model for a specific project involves fitting all forms of change point models and identifying the one with the least root mean square error

Acknowledgements

This material is based upon work supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy under Award Number DE-EE0003563.We would like to thank the following individuals and organizations for their support in this project: Lani Macrae. U.S. Dept. of Energy – Better Buildings Regional Account Manager; Rob Melnick, ASU Global Institute of Sustainability – Executive Dean and Energize Phoenix Principle Investigator; Carolyn Bristo, City of Phoenix Assistant Public Works Director and Phoenix Sustainability Officer; Dimitrios Laloudakis, Energize Phoenix Project Director and City of Phoenix Energy Manager and Deputy Public Works Director; Valerie Wynia, APS Non Residential Programs; and James Holbrook, Leader APS Residential DSM Programs.

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.