# Use of a Billing Simulation Tool for Performance Measurement and Verification

*David Robison, Stellar Processes*
*Howard Reichmuth, Stellar Processes*
*M. Sami Khawaja , Quantec*

## ABSTRACT

USDOE has published the International Performance Measurement and Verification Protocol (IPMVP) (USDOE, 1997). As a result, government and international agencies are requiring that verifications be conducted in accordance with these protocols. This paper describes IPMVP verification conducted in Jamaica for a World Bank-funded project. Since the measures included HVAC improvements, it was necessary to use calibrated engineering models as the method of verification (Option D of IPMVP). Although one might expect the tropical climate to be uniform, in fact there are seasonal changes in humidity and latent loads. During the study period, El Niňo weather effects caused atypical consumption changes of about 10% -- a change large enough to prevent direct comparison of pre/post data. The models normalized consumption to a consistent annual basis. The modeling method utilized a simplified engineering simulation designed to be easily calibrated against billing data and to avoid laborious programming. Overall precision of the results was quite good. Monthly results matched with a relative mean Standard Error (SE) of about 5%; normalized annual results with a relative mean Standard Error (SE) of about 2%. Savings estimates were provided with a SE value that is less than 5% of annual consumption. This value defines the level of resolution that can be expected from monthly whole-building analysis. Since the measures were expected to save about 15%, the precision of this approach is adequate for statistical significance. The project demonstrates that sufficiently precise simulations can be developed from whole-facility billing data at a greatly reduced cost compared to traditional engineering models.

## INTRODUCTION

### Performance Verification

The U.S. Department of Energy initiated development of a consensus approach to verifying energy investments. This effort stemmed form the realization that, without guidelines, performance measurement seemed like a daunting task to energy service providers. Reluctance to assure performance presents a market barrier by discouraging investment. Since release of the first standard in 1996, the International Performance Measurement & Verification Protocol (IPMVP) has become a requirement for many Federal and international agencies. This paper describes performance verification following IPMVP requirements for a World Bank-funded project.

The IPMVP protocol presents four Measurement and Verification (M&V) options:

- Option A: Engineering calculations based on spot measurements
- Option B: Engineering calculations based on short-term monitoring
- Option C: Billing analysis at the whole-building level using statistical techniques
- Option D: Calibrated engineering simulation models

The protocol generally presents these methods as common sense general guidelines. The generality has been a source of confusion for some. Critics felt confused by the lack of specificity or even worse, misinterpreted the protocols as requiring an onerous degree of effort. This confusion seems to have derived from an expectation that complex projects require Option B. This option becomes quite expensive when there are large numbers of measures to sample and a large amount of data to interpret. In these cases, Option D is easier to implement.

The most important requirement of the IPMVP has not been emphasized in the documentation -- that is, the recognition that a savings estimate is meaningless without also stating the precision and accuracy. The IPMVP makes it clear the accuracy of the savings estimates must be included in the verification report. It is recognized that accuracy requirements affect the cost of the verification and that part of the M&V plan is establishing the tradeoff between cost and precision.

With this paper, the authors hope to demonstrate that utility bill analysis at the whole-building level provides a low cost method of verification. There are no metering costs because the utility data already have been collected. These data are used to calibrate an engineering simulation model that provides disaggregation of the savings and normalization for extraneous factors.

## METHODOLOGY

### Precision of the Savings Estimates

When estimating savings, it is important to report precision. For small savings, it is possible that the estimate falls within the range of measurement error. In that case, the estimate is really meaningless -- one has no way of knowing if random "noise" caused whatever change was observed. Statistics provides a way of checking the results to see if the estimate is "significant" or likely to be a real effect of the efficiency measures.

Statistical inference making is based on collecting data on a sample and making estimates that can be extrapolated to the population. Measured values ($Y_m$) from the sample are expected to represent the true values ($Y_i$) in the population with a certain amount of random error. Thus:

$$y_m = y_i + \varepsilon \qquad \textbf{Equation 1}$$

where ε is a random term representing measurement error or random uncertainty.

When a causal model is used to measure y, the model becomes $\hat{y}_i = f(x_i)$ indicating some relationship between x and the modeled value for y, y(hat). In the case of energy consumption models, x includes such variables as weather.

Of course, the relative size of the errors affects one's ability to draw conclusions about the true values of Y. The usual statistical approach is to compute some estimate of the variance of the error term that can be used to assess the uncertainty. Such methods often have restrictive assumptions regarding the statistical distribution of the error term.

For this project we chose to model the uncertainty around the error term using the mean Standard Error computed for monthly observations:

$$SE(M) = \sqrt{(y_i - \hat{y}_i)^2 / n} \qquad \textbf{Equation 2}$$

where

*SE(M)* is the standard error associated with the monthly observations.

*n* is the number of observations.

SE provides a measure of typical amount of measurement error associated with a typical observation. In this application, the SE describes the error associated with the monthly values of energy consumption. However, the goal is to examine consumption over a year comprised of twelve of these monthly observations. The standard error associated with a year can be described:

$$SE(A) = \sqrt{SE(1)^2 + SE(2)^2 + ... + SE(12)^2} = \sqrt{12 * SE(M)^2} = \sqrt{12} * SE(M) \quad \textbf{Equation 3}$$

where

*SE(A)* is the standard error of the annual sum.

*SE(i)* represents the standard error of the individual months, 1 through 12.

Assuming that the SE(M) represents the typical or average monthly standard error, one computes the annual error as being SE(M) times the square root of twelve or roughly three times larger. (As a simplification, we are weighting the twelve months equally although the number of days in each of the periods may vary). Notice that the annual sum is twelve times the average monthly consumption. Yet the SE of the sum is only larger by the square root of twelve. If we consider the SE as a fraction of the average observation, the annual SE expressed as a relative fraction of the annual sum is about one third the SE of the monthly observations. As one expects, since the monthly errors are randomly distributed, they will tend to cancel out when computing the sum. So the relative precision of the annual sum will be better than that of the individual months.

$$\frac{SE(A)}{A} = \frac{\sqrt{12} * SE(M)}{12 * SE(M)} = \frac{1}{\sqrt{12}} * \frac{SE(M)}{M} \qquad \textbf{Equation 4}$$

where

A is the annual consumption, sum of 12 months.

M is the average monthly consumption.

In order to normalize for variations in climate that may occur from year to year, the model extrapolates monthly consumption for a series of months with typical (long-term average) weather. Since there is no way of measuring error for such hypothetical constructions, we will assume that the SE for the observed period is equivalent to the amount of random error that would be associated with any other weather period.

Energy savings are computed as the difference between baseline estimate and post installation normalized consumption. For this operation, the standard error is once again computed as:

$$SE(diff) = \sqrt{SE(baseline)^2 + SE(postconsumption)^2} \qquad \textbf{Equation 5}$$

where

SE(diff) is the standard error of the difference or of the savings estimate.

With the standard error defined, one can draw conclusions about the relative amount of uncertainty inherent in the estimate. Drawing analogies with standard statistics, one can compute confidence limits. Useful parameters to consider are the following:

- Probable Error (PE), defined as the 50% confidence range. This represents the most likely amount of error. That is, it is equally likely that errors are either larger than or less than the PE. (ASHRAE, 1997)
- The 90% Confidence Limit (CL), defined as the range where we are 90% certain that random noise did not produce the observed difference.

These parameters are computed as:

$$PE = 0.675 * SE(Diff) \quad \textbf{Equation 6}$$

$$CL = t * SE(diff) = 1.96 * SE(Diff) \quad \textbf{Equation 7}$$

Note that 0.675 is the value obtained from standard normal distribution (t value) for 50% probability.

**Simulation Modeling**

Conventional engineering models require complex, time-consuming inputs. Furthermore, since they typically run from average weather, the results are difficult to reconcile against local weather conditions. For this project, we utilized a monthly simulation model (White and Reichmuth, 1996) that was specifically designed for ease of calibration to utility bills. (Stellar Processes, 1999) Hourly modeling represents an unnecessary level of overkill when only monthly data are available for calibration anyway. The modeling tool has been shown to provide similar results to the more complicated DOE2 simulation model (Robison and Reichmuth, 1999a) and is well suited to performance verification (Robison and Reichmuth, 1999b).

It is important to recognize that the most significant modeling inputs are usually available without extensive site measurements. Even cursory site observations have significant value in understanding how a building uses energy. With the correct calculation structure, assumed defaults for equipment lead to good annual energy estimates. More detailed quantitative information contributes toward refining the model, but simple qualitative answers provide enough information to establish the basic modeling parameters.

To develop a physical model of the building, one includes measurements derived from "snapshot" measurements, audits and short-term observations. The most important drivers of total building energy use are (1) internal loads (2) occupancy duration (3) outdoor weather, and (4) ventilation. Often a lighting survey is available and serves to quantify much of the internal loads. Occupancy duration can be inferred from occupant interviews or from a simple 24 hour load profile. Ventilation can be assessed by a general site inspection and by measurement of airflow and temperatures at the air handlers. However, flow measurements are not highly accurate -- assumed ventilation values are often adequate for the model. The mean outdoor temperature is an explicit model input and serves as a proxy for other weather drivers. Thus, the modeler is fortunate in that the most important inputs are readily available from the existing utility meters, simple weather (daily temperature) data, and a reasonably competent walk-through audit.

For this study, we utilized lighting surveys previously prepared by utility staff and 24-hour consumption profiles derived from short-term monitoring at the whole-building level.

For one participant, we had submetered consumption for lighting circuits. For two participants, we verified schedules with short-term lighting loggers. The profile information served to confirm the operating schedule and the level of internal usage during occupied and unoccupied periods.

Engineering simulation models operating from average weather present another difficulty for precision estimates. Equation 2 assumes that the modeled value, y(hat), is computed under the same conditions as the observed value, y. If the modeling tool computes under a different set of input conditions, this assumption is not valid. For that reason, estimating precision requires a modeling methodology that uses actual, local values for climate variables, rather than long-term average values.

This project illustrates the necessity for simulation modeling. One might think that in the benign tropical climate of Jamaica, energy consumption would not vary seasonally. In fact, there can be significant changes in the latent loads caused by humidity. Furthermore, during the study period, participants were exposed to unusual conditions due to El Niňo weather effects. Figure 1 shows the range of daily temperatures experienced during the pre- and post-retrofit years. Conditions during the pre-retrofit period of 1996 were typical of average weather conditions. However, 1998 was atypically warm.

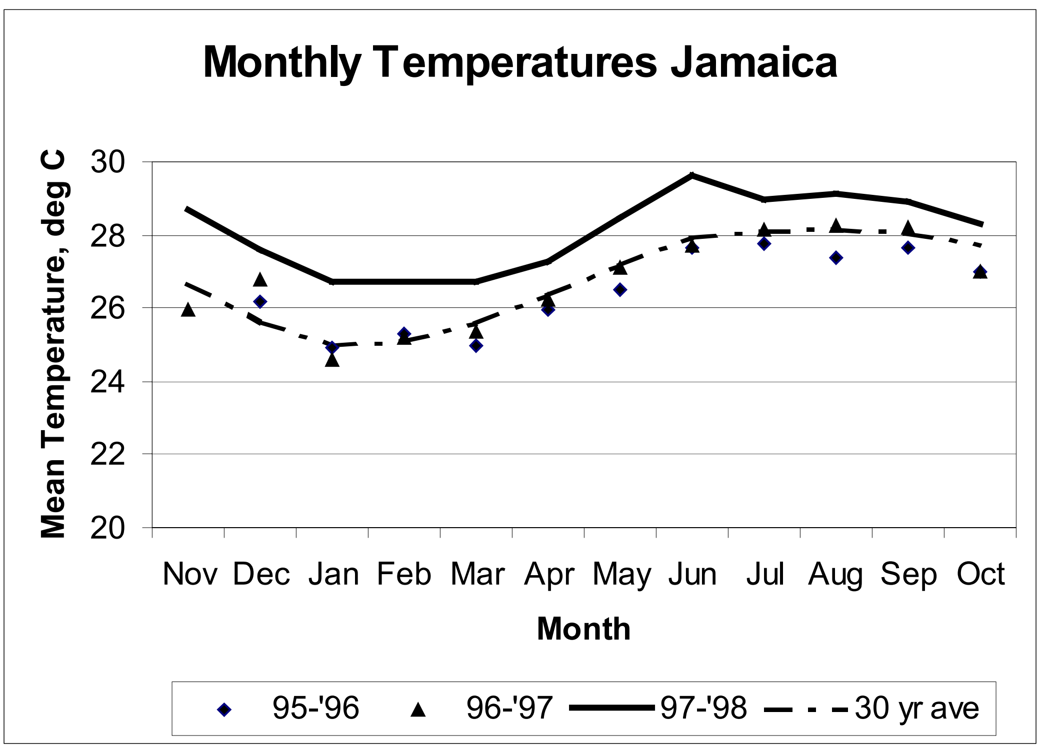


**Figure 1. Deviation from Average Weather**

Figure 1 shows that El Niňo weather influences traveled to the Caribbean. The two years prior to the retrofit proved to be quite consistent with the long-term average weather. However, 1997-8, the post-retrofit year departed significantly. Both temperature and humidity increased about 5%. Although the change may seem small, the climate-induced increase in cooling load represents as much as 10% increase in energy consumption. Savings were expected to be about 15%. Thus, the climate effect is about the same magnitude as the expected impact. A direct comparison of pre/post bills would not be able to distinguish savings. Thus, some method of normalizing for weather must be applied in order to accurately measure savings.

A statistical method (IPMVP Option C) could be applied but would be less satisfactory than an engineering model. Lacking specific data inputs, the statistical method is not able to account for other physical changes, such as increased operations, that might have taken place at the same time. Nor is it able to estimate the relative contributions of different energy conservation measures.

## RESULTS

Figure 2 shows a typical example of the curve fit comparing actual billing records to consumption predicted by the simulation model.

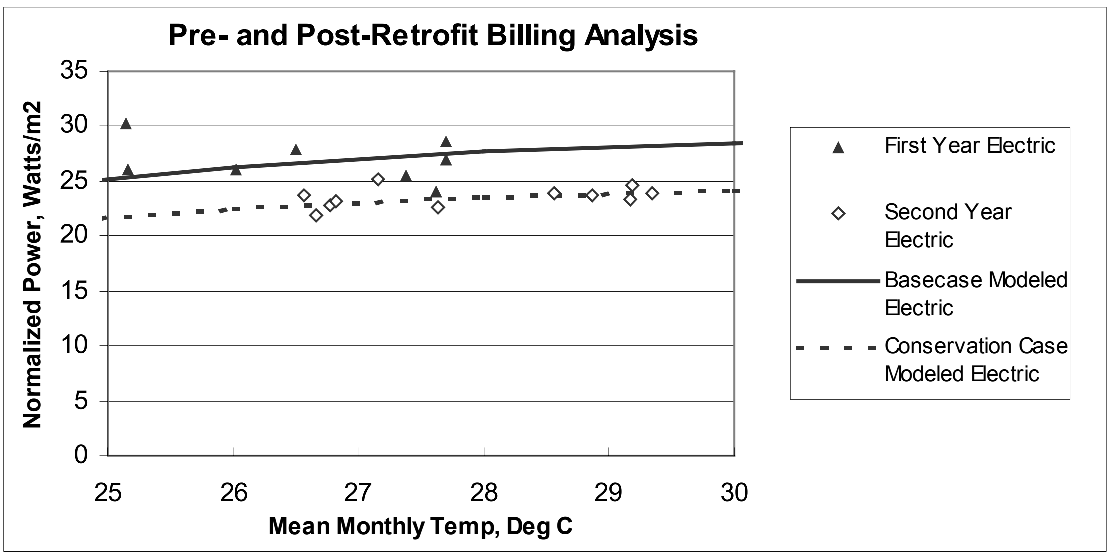


**Figure 2. Example Billing Analysis**

For this participant, pre-retrofit conditions were rather noisy. The SE of pre-retrofit monthly consumption model is 35,195 kWh or 7.49% of mean monthly consumption. The post-retrofit SE is 7,018 kWh or 4.6% of mean monthly consumption. The Standard Errors of the annual consumption estimate are 121,975 kWh or 2.1% pre- and 52,648 kWh or 1.1% post-retrofit. This illustrates the observation that the relative precision on an annual basis is improved by a factor of about three over the monthly. The Standard Error of the savings is based on the difference of the two annual estimates. In this case, the SE of the difference is 133,676 kWh or 2.4% of the annual pre-retrofit whole-building consumption. The annual savings are estimated at 15%, with 90% confidence limits of 260,399 kWh or 4.6% of total consumption. Thus, the precision of this method is clearly quite sufficient to provide a reliable savings estimate.

In the other cases, we observed that the models also matched monthly consumption well. That is, the SE error was relatively small compared to the total amount of monthly consumption. In general, the monthly SE was about 3% of monthly consumption. The standard error for annual consumption increased but not as rapidly as the sum of twelve months. Thus, we observed annual consumption with an error of about 2% of total annual consumption. Comparing two years to estimate savings produces an error of about 5%. Thus, we can expect to distinguish savings that are larger than 5% of total annual consumption.

The predicted savings for this program were expected to be about 15%. For such participants, the billing analysis method is sufficient to distinguish "real" savings from random noise. However, one participant did not implement the measures. For this participant, savings were close to zero and not statistically significant.

In general, this study demonstrates that billing analysis carries sufficient precision to be able to isolate actual savings for measures that save a reasonable amount. This is a welcome finding because it utilizes a relatively low-cost procedure to meet requirements of the USDOE International Performance Measurement and Verification Protocols (IMPMVP). Precision results from this study are summarized in Table 1. To simplify comparisons, we have reported standard errors as a percent relative to the average annual pre-retrofit consumption. Savings, in annual kWh, are considered significant if they exceed the 90% confidence limit.

For those participants with both pre- and post-retrofit billing data, reliable estimates of savings are obtained. For the first participant, savings were slightly negative and not statistically significant. Investigation determined that the conservation measures were not appropriately installed in this case. For the other participants, the savings estimates were strongly positive and significantly different from zero. The precision of the modeling technique is sufficient to develop creditable estimates for verified savings. These estimates, however, are far from the "90/10" rule often discussed as an accuracy goal. (This rule, with a goal of being 90% confident that the "true" savings are within 10% of the estimate, is often applied during evaluation design to determine the sample size.)

The relative precision of the savings estimate depends on the magnitude of the savings. As shown in Table 2 and Figure 3, confidence limits of about $\pm$5% of annual consumption may be about 30% of the savings estimate. However, this level of accuracy is quite sufficient to eliminate the null hypothesis and provide creditability to the estimates. Thus, an expectation for "90/10" precision of estimates is not necessary. In this study we observed about 90/30 precision (that is, we are 90% confident that savings are within 30% of the estimate). Figure 3 shows that this level of precision is clearly sufficient to reject the null hypothesis or the possibility that random noise produced the observed effects. This level of resolution is about as good as could be expected for any sort of whole building modeling. More complicated simulation tools are unlikely to provide better resolution.

**Table 1. Impact Results**

| Site | Pre | | Post | | Savings |
|---|---|---|---|---|---|
| | Monthly Standard Error, SE(M) | Annual Standard Error, SE(A) | Monthly Standard Error, SE(M) | Annual Standard Error, SE(A) | Standard Error SE(Diff)/ Annual |
| Hotel1 | 4.5% | 1.3% | NA | NA | NA |
| Hotel2 | 5.6% | 1.7% | 6.0% | 1.8% | 2.4% |
| Lg Office1 | 3.8% | 1.1% | 4.1% | 1.2% | 1.4% |
| Sm Retail1 | 5.3% | 1.7% | NA | NA | NA |
| Sm Retail2 | 4.8% | 1.4% | 5.3% | 1.6% | 1.9% |
| Sm Retail3 | 8.8% | 2.7% | 11.5% | 3.8% | 3.5% |
| Lg Office2 | 7.5% | 2.1% | 3.7% | 1.1% | 2.4% |
| Hotel3 | 4.0% | 1.1% | 3.3% | 1.0% | 1.5% |
| Sm Retail4 | 6.6% | 2.0% | 7.1% | 2.2% | 3.0% |
| Lg Office3 | 7.4% | 2.1% | 6.2% | 1.9% | 2.6% |

| Site | Savings Amount, kWh/yr | Probable Error, kWh/yr | 90% C.L. kWh/yr | Significant? |
|---|---|---|---|---|
| Hotel1 | 596,805 | NA | NA | NA |
| Hotel2 | -122,777 | 111,979 | 325,155 | No |
| Lg Office1 | 532,696 | 23,248 | 67,506 | Yes |
| Sm Retail1 | 34,836 | NA | NA | NA |
| Sm Retail2 | 7,606 | 719 | 2,086 | Yes |
| Sm Retail3 | 33,005 | 2,376 | 6,899 | Yes |
| Lg Office2 | 841,972 | 91,973 | 267,063 | Yes |
| Hotel3 | 372,033 | 50,193 | 145,745 | Yes |
| Sm Retail4 | 10,437 | 1,085 | 3,152 | Yes |
| Lg Office3 | 216,965 | 22,115 | 64,217 | Yes |

**Table 2. Confidence Limits of Savings Estimate for Sites with Pre/Post Billing Data**

| Project | Savings as Percent of Consumption | 90% C.L. as Percent of Savings |
|---|---|---|
| Lg Office1 | 22% | 13% |
| Sm Retail2 | 14% | 27% |
| Sm Retail3 | 33% | 21% |
| Lg Office2 | 15% | 32% |
| Hotel3 | 8% | 39% |
| Sm Retail4 | 19% | 30% |
| Lg Office3 | 17% | 30% |

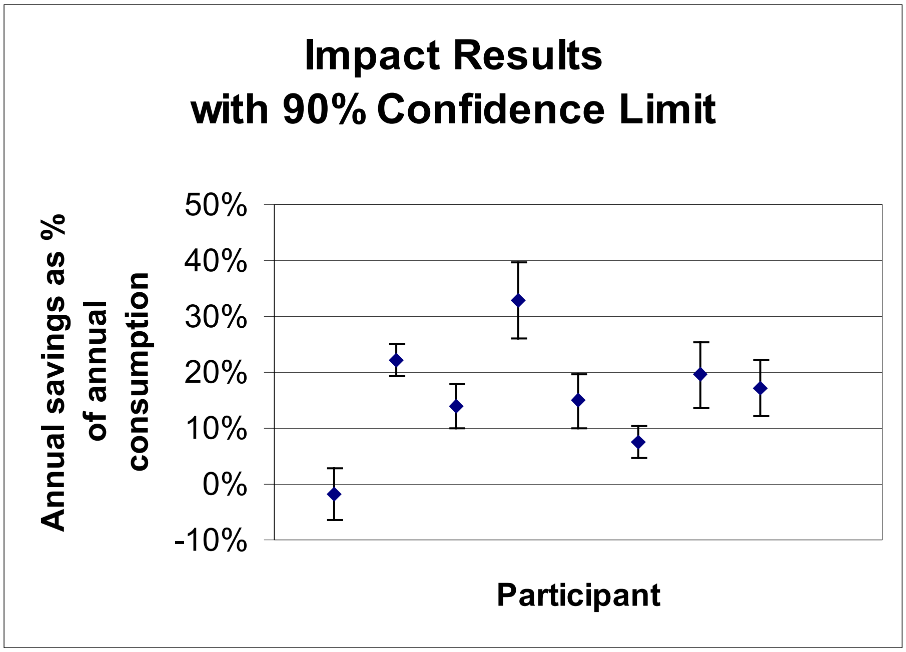


**Figure 3. Impact Results and Precision**

The IPMVP suggests as reasonable expectations of accuracy and cost.

- Option C: Statistical analysis of monthly data, accuracy ±20% of savings, cost 1-3% of retrofit project cost
  Statistical analysis of hourly data, accuracy ±5-10% of savings, cost 3-10% of retrofit project cost
- Option D: Simulation with monthly data, accuracy ±20% of monthly consumption, cost 5-10% of retrofit project
  Simulation with hourly data, accuracy ±1-5% of monthly consumption, cost 100% of annual bill

This study demonstrated that accuracy of simulation modeling is much higher and costs can be much less. The calibrated simulation method using monthly data resulted in a high level of accuracy (±4% monthly, ±1% annually) at a cost of about 1% of the retrofit cost.

## Transferability of Results

Looking at Figure 2, one observes a curve fit similar to what might be expected from a statistical regression fit. To an extent, the results shown here would be similar for a statistical analysis. There is, however, an important difference. With the statistical treatment, it is necessary to adjust the number of observations for the degrees of freedom. One might expect to use a multiple-variable model, including, for example, heating and cooling degree-days, relative humidity or an occupancy variable. Since there are only 12 monthly observations, adding variables could lead to a serious decrease in precision. To compensate, one would want to extend the period of observation to include more billing months.

However, additional billing data may not be readily available or may reflect a different set of operation conditions. Thus, while the overall precision may be similar to that which could be derived from regression modeling, statistical modeling (Option C of IPMVP) may not be practical.

The applicability of traditional engineering simulation models is even more in doubt. These models would be computing building performance under average weather or a different set of conditions. It is not clear how one would derive a Standard Error of modeled performance because the independent variable would be specified as a different value than occurred during the billing interval. For this reason, it may not be possible to compute the precision of a traditional simulation model as required for Option D.

## CONCLUSIONS

- Some sort of weather normalization was required for this study due to extreme climatic changes that interfered with simply comparing pre-post-retrofit energy bills.
- The monthly simulation method provided sufficiently precise estimates of savings. In this study, the standard error of the savings estimate was 2-3% of annual consumption. This relative error defines the resolution of the technique. However, the precision was more than adequate for savings that were typically 15-20% of annual consumption.
- Relative accuracy of the savings estimate depends on the size of the savings, since the relative error is fixed by the amount of noise in the monthly observations. In this study, the relative error of savings corresponded to about 90/30 precision.
- Meeting the verification requirements of Option D of the IPMVP can be accomplished without additional monitoring expense using data already available -- namely whole-building utility bills, supplemented with facility audits or other existing site information.
- Use of statistical billing analysis (Option C of IPMVP) is possible but suffers serious limitations.
- Other simulation tools would be expected to have a similar level of accuracy although the costs to implement them may be too great for their consideration.
- Use of traditional engineering simulation tools may not be capable of generating the type of precision reporting required by Option D of the IPMVP.

## REFERENCES


USDOE, *International Performance Measurement & Verification Protocol*, http://www.ipmvp.org, 1997.

ASHRAE, *Fundamentals Handbook*, 13.4, 1977.

White, James and Reichmuth, Howard, "Simplified Method Of Predicting Building Energy Use In Complex Buildings", *Proceedings of the International Conference on Energy Conversion*, Washington DC, August 1996.

Stellar Processes, *EZ Sim Manual*, http://www.ezsim.com, 1999.

Robison, David and Reichmuth, Howard, "Performance Evaluation of Commercial Construction Using a Billing Simulation Tool*", International Energy Program Evaluation Conference Proceedings*, Denver, CO, 1999a.

Robison, David and Reichmuth, Howard, "Use of Billing Simulation Tool for Commissioning", *National Commissioning Conference Proceedings*, Portland, OR, 1999b.