Sustaining Optimum Building Performance Using Key Metrics

Robert Austin
E M C Engineers, Inc.

Synopsis

Sustainability of savings has become a major focus in the commissioning industry, especially in utility-funded retro-commissioning (RCx) projects. One key to sustaining optimum building performance is comparing predicted energy performance of a building with key measurement and verification (M&V) metrics to benchmark performance. M&V data often is used by utility companies to verify the effectiveness of large scale programs, but it is seldom used by building engineers to ensure that buildings perform as intended.

We will identify the most cost-effective and time-effective ways that M&V data can be used to ensure optimal building operation. We will also present the most essential metrics for tracking post-installation performance as well as outline a methodology for establishing these metrics using both hour energy model simulations and spreadsheet regression models. Equally as important as the establishment of these key metrics is a determination of what, if any, limitations they may have in identifying system level deficiencies should the facility revert back to "sub-optimal" performance. The findings are equally applicable in Commissioning (Cx) projects where determination of optimal performance parameters for future sustainability in high-performance buildings is desirable.

About the Author

Robert Austin is the Lead Project Manager for EMC Engineers in Irvine, California. He has a B.S. in Mechanical and Aerospace Engineering from Oklahoma State University. Mr. Austin is responsible for managing EMC's commissioning and retro-commissioning projects throughout the SoCal Region. Over the past five years, Robert has been directly responsible for the execution and implementation of retro-commissioning projects in over six million square feet of space including government, institutional, K-12, office, and hotel facilities.

What Is a Metric?

A metric is an equation that relates measurements of performance characteristics with the independent variables that have the greatest effect on them. It establishes a reference point against which future measurements can be compared in order to evaluate potential deviation from what was expected. In other words, it establishes a "baseline" performance characteristic showing how the equipment is expected to perform.

In order for a metric to be usable for performance tracking, it should be easily quantifiable and calculated, and it should take into account all of the most important independent variables likely to affect the measured values. It should clearly establish the expected level of performance in such a way that it is obvious when the performance measurement is not consistent with the value that was expected.

Which Metrics *REALLY* Matter?

The final phase in most Cx and RCx projects is typically referred to as the "Training" or "Hand-Off" Phase. It is at this point that the project has been completed, and it is now up to the facility engineer to maintain the building moving forward. In order to give both facility engineers and building owners the most value from the process, it is important to educate them on how to effectively track the performance of the facility moving forward. Performance metrics provide a means of doing this.

The International Performance Measurement and Verification Protocol (IPMVP) outlines commonly recognized guidelines for properly validating and documenting energy savings in energy efficiency retrofits and new construction projects with a goal of "energy use avoidance" in mind. Companies that specialize in this field follow these guidelines stringently in order to promote consistency in how energy savings are claimed in all types of projects. Energy programs sponsored through a Public Utilities Commission (PUC) are often required to utilize a third party provider to independently verify the savings claimed for a given project using IPMVP guidelines as a standard. This helps ensure that the public goods money invested in energy efficiency projects produces quantifiable and real results.

In most RCx and Cx projects, the commissioning provider is generally not involved in this process (often due to the requirement for third party validation). As such, performance monitoring and tracking is not covered by most providers during the training given to the facility staff and owners. However, increasingly, both owners and operators have a desire to track the energy performance of their facilities over time after the commissioning provider has completed the project, but they generally lack the tools to do so. With energy-driven RCx projects, the data is almost always available to develop these tools for the operator if the commissioning provider is willing to employ M&V methodologies.

After implementation is completed in most RCx projects, the commissioning provider is responsible for verifying that the newly installed systems operate correctly through Functional

Performance Testing (FPT). Once the measures have been verified, a final report is compiled and training is completed. The energy savings estimates given to the owner in the final report were likely calculated using either spreadsheet regression models or detailed hourly energy simulations. In either case, it is important to note that the savings reported at this time are still considered "*estimated* savings."

In order to provide the owner with predictive performance metrics that can be utilized to track performance over time, it is critical that the commissioning provider have the ability to go beyond "estimated annual savings" to something an owner can reliably count on moving forward. The challenge, however, is in understanding and establishing what the most essential metrics are for a given building. There is no one set of metrics that can be universally applied across the board for every facility and every scenario. Rather, in establishing the most effective metric types, a number of factors must be taken into consideration.

Whole Building or System-Level Metrics?

A few of the whole building energy usage metrics that are commonly employed in tracking building performance on an annualized basis:

- kWh/Square Foot
- kBtu/Square Foot (often used to combine multiple fuel types into a single metric)
- Therms/Square Foot

Certainly the simplest and easiest way to track energy performance savings is to track utility bill invoices over time and compare them to the predicted usage levels. After all, an owner or operator would know what baseline annual energy usage was prior to the project, so simply subtracting the annual energy usage from the baseline after the project was complete and comparing it to the estimated savings should be adequate, right?

In reality this method has a number of practical limitations. The most obvious limitation is due to the amount of savings that RCx actually achieves. For example, in a recent RCx program implemented by EMC for 10 large facilities (approximately 1.4 million square feet), the overall estimated electric usage (kWh) savings was 18% of the baseline usage totals. However, for the individual buildings in the project, the savings ranged from 1% at the low end to 34% at the high end.

Assuming that the facility occupancy patterns, hours of operation, and space utilization remain relatively constant such that the ambient condition is the only independent variable affecting the building load, the variance due to outside air temperature variations from one year to the next alone could account for as much as 5% to 10% variance of the overall electric and gas consumption. If whole building utility data alone were utilized as a performance metric with no correction for ambient conditions, then it would be nearly impossible to track expected performance over time when the savings estimates are 10% or less of the whole building usage.

In addition to the uncertainty caused by savings that are lower than the expected variance in the normal utility usage, there is also the issue of "modified" baselines. The IPMVP defines savings as “*avoided* energy use.” Suppose, for example, that a facility has an enclosed parking garage where the exhaust and ventilation fans are not operating in the manner required by code (which is 24/7 in systems where there is no CO monitoring). Restoring the fan operation to meet code requirements would use more energy than what has been recorded up to that point through the utility invoices.

If a CO monitoring system is then installed to allow the fans to shut down during off peak usage periods, what is the baseline from which savings can be calculated? IPMVP guidelines would allow savings to be calculated from a baseline assuming 24/7 operation, even though this would not be evident in the utility invoices. However, in this instance, the actual savings that the owner will see on the monthly utility invoice would be lower than the actual “avoided energy use” savings estimated by the commissioning provider. Thus, overall utility data (without adjustment) would again be insufficient to adequately predict future performance (though an adjustment could be made in this case to make a whole building metric work).

When Whole Building Metrics Are Inadequate, Which System Level Metrics Are Most Effective?

A few of the whole building energy usage metrics that are commonly employed in tracking building performance:

- Common Aggregate Metrics
  - Daily or Monthly Fan kWh/Average Degree OAT or Cooling Degree-Day
  - Daily or Monthly Chiller kWh/Average Degree OAT or Cooling Degree-Day
  - Daily or Monthly Therm/Average Degree OAT or Heating Degree-Day
- Common Instantaneous Metrics
  - Chiller kW/Degree Outside Air Temperature
  - Fan kW/Degree Outside Air Temperature

The ideal choice of system-level metrics will depend to a large degree on the system type, as well as which systems are affected by the measures. We find it is best to work from the top, down, so to speak. When whole building data isn’t adequate, the next logical choice would be total HVAC energy (including all plant equipment as well as fans together as a single metric). This would include all sub-systems that were affected by the RCx measures. For most RCx projects, this would include fans, pumps, chillers, and cooling towers, at a minimum.

The next level down would be to break the sub-systems shown above into individual metrics. The independent variable affecting each sub-system is a factor in determining the proper metric, as well. For example, if the fans are constant speed, the only independent variable would be the runtime. Thus, the proper metric in this instance would simply be daily kWh consumption compared to a predicted value based on the expected runtime. However, if the fans were serving a variable volume system, then fan energy would likely vary as a function of outside air temperature or cooling degree-days (as a minimum). In this case, a proper performance metric could be daily kWh versus cooling degree-days.

The choice of whether to utilize system-level metrics at the main sub-system level (total HVAC energy) or at a lower level (fans energy, chiller energy, pump energy, etc.) depends upon a number of factors. Monitoring and tracking performance at the higher level is much more cost-effective and time-effective in some respects. For example, if a building possesses a motor control center that houses all the HVAC equipment in a single location, all that would be required is a single kWh pulse meter (from which kW could be extrapolated). This is more cost-effective since individual electrical monitoring sensors would not need to be installed on every piece of equipment. The tradeoff, however, is diminished troubleshooting capability.

Suppose a facility is tracking the performance of the total HVAC weekly kWh as a function of cooling degree-days and comparing with predicted values. Over a given month, the HVAC system uses 25% more energy than expected. Would the operator then have any further insight into what the specific cause could be? Additional time would then have to be spent tracking down the problem and correcting it.

Now assume that the facility tracks performance of fan, chiller, pump, and cooling tower energy separately using similar metrics of total kWh as a function of cooling degree-days. In this scenario, if any one of the metrics drifts out of the allowable range, the building operator will already be able to pinpoint the specific system to address. But then there is the additional time and expense of checking multiple metrics and installing the additional metering.

All of these factors are important to consider when determining how much data can be too much and how much is not enough. There are personnel factors to consider as well. In a facility with multiple operating staff it is more reasonable to expect that consistently checking multiple system level metrics on a regular basis can become a normal part of the PM routine. However, at a site where there is perhaps only a single mechanic who is also simultaneously responsible for maintaining other facilities it may only be practical to track a single metric on a monthly basis.

When Does Whole Building Utility Data Yield Reliable Metrics?

The most common instance where whole building utility data can be reliably used to track performance is in the case of metered natural gas usage. In most types of buildings where RCx is typically employed, the domestic hot water accounts for only a small portion of the overall natural gas usage (except in facilities such as hospitals and hotels). In systems where the remainder of natural gas energy consumption is used by HVAC, total natural gas consumption can be a very reliable metric for tracking performance.

For any metric that falls under the category of “system level” or “whole building level” it is important to define any independent variables that may affect the numbers. The reason that gas usage from the baseline utility data cannot simply be compared to the post-installation data is that the usage is a function of the heating load. In most facilities the heating load is a function of ambient conditions (at a minimum). If two years of data were compared to one another without any adjustment for ambient conditions it would not be a fair comparison. Therefore the appropriate methodology is to develop a correlation that can predict the energy consumption as a function of its independent variable(s).

In some building types there may be more usage of domestic hot water (such as in hospitals and hotels). In these cases, when an owner desires to track HVAC natural gas consumption, the best option is to install a gas sub-meter for the HVAC boilers and tie the meter into the existing DDC system (if available). Other indirect means of tracking the boiler load may be employed, but they are likely much less reliable as well (this is discussed further below).

Instantaneous Metrics or Aggregate Metrics?

In addition to determining whether to use whole building data or system-level data, a decision must also be made whether to use instantaneous metrics or aggregate metrics. Typically when an energy analysis is performed for an RCx project, correlations for baseline operation are developed to determine system-level operation. These correlations are most often instantaneous power measurements, as opposed to aggregate energy usage. An example would be for a chiller analysis.

The most basic chiller analysis would consist of measuring the chiller power draw over a period of time (usually at least two weeks) while simultaneously measuring the ambient outside air temperature. After the data is collected, a correlation would be developed to present the chiller kW as a function of outside air temperature. This correlation would then be used in a bin data analysis, and energy efficiency measures would be analyzed against this baseline correlation.

Then, after the energy efficiency measures have been implemented and verified, a sample of chiller data could again be monitored and a new correlation developed. This is commonly done by M&V providers. However, this correlation could also be utilized from that point forward to predict and track performance and ensure optimal performance is sustained.

While both instantaneous metrics and aggregate metrics can be valuable resources for tracking performance and diagnosing potential issues, aggregate metrics seem to have fewer limitations and are generally more sensitive to deviant operation. An example of the limitations that can sometimes be posed by instantaneous metrics is that they ignore time. Thus, an instantaneous metric would not be capable of detecting any deviation in runtime patterns. This is important since most RCx projects have at least some degree of schedule optimization involved.

End Use Energy Metrics or Capacity/Efficiency Metrics?

End use metrics are those that only take into account the electrical or natural gas energy used by a single system or sub-system. Efficiency metrics factor both the energy required to operate the system and the cooling or heating capacity delivered from that system. For example, Chiller kW/Degree OAT is an end use metric. Chiller kW/ton/Degree OAT is an efficiency metric. Both can be valuable tools when monitoring and assessing building performance.

The primary limitation with efficiency metrics is related to how they are derived. For example, chiller tonnage delivered can be derived in several ways. Most commonly, it is arrived at by multiplying the flow through the chiller (in GPM) by the temperature drop across the chiller.

This number is usually then multiplied by a constant based on the fluid properties to calculate the chiller tonnage.

The inherent uncertainty is due to the fact that the flow is typically assumed constant (and sometimes not even spot measured), and the temperature drop is calculated from the difference between two sensors (chilled water supply and chilled water return). Assume that we have a flow meter in a system with +/- 5% accuracy, and the water temperature sensors are accurate to within 0.5 degrees. If the actual flow through the system is 500 GPM with a temperature difference of 10 degrees, the actual tonnage in this case would be approximately 208.3 tons. However, with sensor inaccuracy, the calculated tonnage we would likely see would be +/-15% of the actual value. This uncertainty is further compounded in the absence of a flow meter.

In the end, after the uncertainty inherent to the measurement is factored into the confidence interval of the predicted performance, the resulting metric would likely not have the desired level of sensitivity to deviant performance. In spite of this limitation, the added benefit of utilizing an efficiency metric is increased diagnostic capability.

For example, when the energy consumption of the chiller drifts outside of the acceptable range, it may not be indicative of a problem with the chiller or chiller plant. It may be that there is simply a problem with the economizers. If only electrical consumption were being tracked, there would be no way to distinguish from the data alone whether the issue resided with the chilled water plant or the air-side systems. However, if chiller tonnage were also being tracked, there would be a tonnage increase at the same time the kW increased. This would allow the user to narrow out the chilled water system as a likely source of the issue (since the kW per ton would remain within its tolerable range).

What Is the Best Way of Generating Performance Metrics?

Another issue that should be considered at the outset of any RCx project with energy sustainability in mind is how to best generate the predictive metrics for performance tracking. In most RCx projects, the energy savings are generated using spreadsheet regression analysis tools. This section will present general guidelines and methodologies on how energy performance metrics can be generated through both hourly simulation energy models and spreadsheet tools. Additional discussion is also included to illustrate some of the limitations of using each methodology.

Generating Metrics Using Hourly Simulation Energy Models

When energy savings in RCx projects are calculated using energy simulation models, it is a simple and natural second step to utilize the model to define predictive energy performance metrics. The same general practices employed to generate a reliable model that accurately represents the building simulated are equally applicable when defining energy performance metrics. For example, there should be a good statistical match between the monthly energy usage and demand predicted through the baseline energy model and the actual usage shown through the facility's energy invoices. Additionally, the input performance and set point parameters put into the energy model should also match closely with how the building truly behaves.

Often due to budget and time constraints in RCx projects, when energy modeling is used the modeler may only employ the guidelines listed above to calibrate the energy model. However, in order to generate more reliable savings, we recommend going to the additional level of calibrating the model down to the system level. This becomes an even more critical step whenever predictive energy performance metrics are a desired outcome of the energy model.

In general, the main systems that utilize electrical energy within a typical facility include lighting, plug loads, chillers, pumps, cooling towers, and fans. If an energy model is calibrated to monthly utility data alone, it could be possible that while the summation of all end uses is within tolerance there is still a significant imbalance among the end use energy of the different subsystems. In other words, the modeled chiller and fan energy could be higher than the actual consumption while the lighting energy is low enough to compensate. The end result would indicate good calibration, but this deficiency within the model would likely bias the energy savings calculations and would most certainly make any system level energy metric generated from the model unreliable. The figure below illustrates an example of a system-level calibration comparison of baseline modeled chiller energy to actual chiller usage.

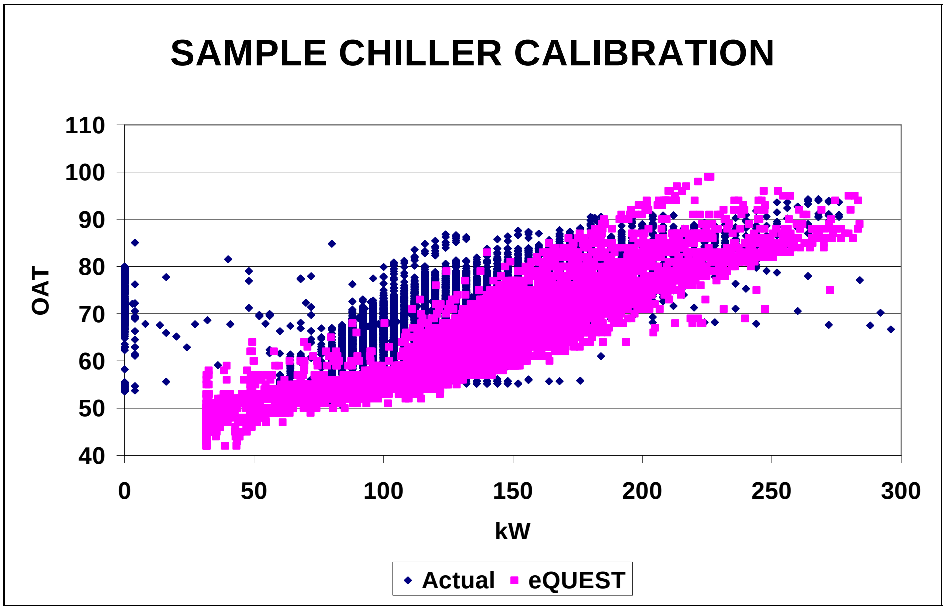

Figure 1: Chiller Calibration Example

Energy modeling offers multiple strong advantages over spreadsheet approaches when used to generate predictive performance metrics. One advantage is that energy models inherently account for interactive effects between each sub-system for each Energy Efficiency Measure. Rule-of-thumb estimates of the interactive effects of multiple measures are often used in spreadsheet calculations to prevent over-estimating the savings. While this may be adequate when calculating energy savings, it is not sufficient when system-level performance metrics are needed post-implementation.

Another advantage to using energy models to develop predictive metrics is that a model simulates the entire year, including all potential modes of operation. When spreadsheet regression tools are utilized, they are typically based only on two weeks to one month of operation. This may not capture enough of the data to reliably define predictive performance metrics under all potential modes of operation.

Generating Metrics Using Spreadsheet/Regression Models

Using spreadsheet/regression models to calculate energy savings does not necessarily render a project unable to generate reliable energy performance metrics. However, spreadsheet tools in and of themselves should not be used to generate these metrics. Rather, post-implementation data should be collected, and new regression correlations generated based on the optimized performance of the facility. Typically this would be the same sort of data and correlations that would be built through the M&V process after the completion of the project.

The following general guidelines should be employed when spreadsheet tools are utilized with the goal of creating performance metrics:

- As much data as possible should be collected, and the data should encompass all modes of operation. Two weeks is typically not enough to do this reliably.
- Aggregate metrics are much more likely to be reliable than instantaneous metrics.
- When other data points in the system are available that are relevant to the measures implemented they should be monitored as well over the same period. This ensures that the implemented measures are “holding” over the period of time that the post-implementation performance data is being collected.
- Where measures are affected by ambient conditions, the post-implementation data should be collected over the entire range of likely conditions.

How Should the Metrics Be Utilized for Sustainability?

As part of the M&V process, techniques and methodologies similar to those described in the preceding sections are often utilized to verify whether the projected savings for energy projects has actually been achieved. Data is collected to generate metrics similar to those described above in order to validate the savings claims made for the projects. In most cases with RCx projects, however, this data is never provided to the owner or the RCx provider. In addition to being very valuable for validating savings estimates, this post-project data can be used to establish performance tracking metrics to help sustain the savings for the project.

In order for the RCx provider to establish tools that are usable and reliable for the owner, it is important to consider two factors that are generally not emphasized in RCx projects. One of these factors is the tendency for some RCx measures to either be “tweaked” or adjusted by the building operator after the project has ended, or to be undone altogether. The metrics selected should address the most likely ways in which the system could potentially be restored to pre-RCx operation. The other factor that should be taken into consideration is that any metric established from savings estimates will usually need to be fine-tuned after implementation.

Which Measures Tend to Revert Back to Sub-Optimal Performance?

The most likely measures to not hold up as well over time are the measures that involve the adjustment of any set point parameters. This includes central plant lockout temperatures, scheduled start/stop times, discharge air temperature set points, etc. These measures may fail over time for any number of reasons, but performance tracking using the appropriate metrics can effectively help curb this tendency.

There is a range of causes, from temporary set point adjustments that are accidentally forgotten, to measures that simply don't work well under every condition. A measure that was installed and verified in the winter may cause issues during the summer, for example.

How Should the Process Be Modified to Address These Issues and Promote Better Sustainability?

There are two important steps that should be taken in order to reduce the potential for these issues to defeat the measures, and to improve the reliability of performance metrics established through RCx. The most important step is to include more comprehensive warranty follow-up in the RCx scope. Typically the provider is not called back to the site unless something significant fails. There needs to be a paradigm shift within the RCx industry that encourages providers to be less *reactive* and more *proactive* during the warranty and follow-up phase of work.

This would entail tasks such as a more comprehensive post-implementation data monitoring and verification plan that goes beyond simply monitoring the measures for two weeks beyond installation. Additional tasks should also include a six month warranty inspection (similar to those often employed for Cx projects) in order to interview the facility operators and help work through any potential issues that may have come up as a result of the measures. Sometimes when measures create other perceived issues within the building, the "knee-jerk" reaction is to restore the system to the way it operated before. This can be as simple is setting the discharge air temperature set point to a constant value to defeat a reset strategy, or as extreme as restoring 24/7 operation to the HVAC system to defeat a schedule optimization measure.

In most cases, the measures simply require further tuning rather than wholesale abandonment. For example, suppose a facility was operating 24/7 in the baseline, even though there was no need to operate on the weekends. Post-implementation, the schedule was operating 12 hours per day, but on warmer days in the summer the building did not stay cool. So the operator then decides to extend the schedule to start the building four hours earlier, thus operating 16 hours per day year-round. In this example, an automated optimum start strategy could accomplish the same goal, without running the building longer during the entire year.

In order for issues such as these to be resolved to satisfy the operator's concern without compromising any more of the energy savings than necessary, it is imperative that the operator and the RCx provider work *together*. But again, in most projects the warranty phase is more *reactive*, and thus the provider may never know that some of the measures have been defeated over time.

In addition to a more comprehensive follow-up plan, there is a need for the RCx provider to take on more of a role of “pseudo-M&V” provider after the measures are implemented. In order to promote better sustainability, the RCx provider’s expertise should be utilized to verify that the savings that were estimated were actually achieved. Once the savings have been verified, performance metrics should then be established with the RCx provider, the owner, and the building operator in order to provide a means of benchmarking and tracking how the facility ought to be performing into the future.

Where energy modeling was utilized to perform savings estimation, the provider should re-calibrate the model to match post-implementation data. Once this is completed, the performance metrics should then be established. The criteria used in the selection of the most appropriate metrics should take into consideration all the factors discussed in this paper, and there should be complete buy-in from all parties. Once the metrics are established, savings have been verified, and the system has had an opportunity to operate for a substantial period of time under a range of conditions, the project can be completely “handed off” to the owner and operator and post-project M&V can be performed.

Conclusion

The findings presented herein are also applicable in new construction projects with a goal of energy efficiency. The same principles will apply for determining the appropriate amount of metering, metrics, and protocols for systematically checking and verifying system performance. Where the goal of any project is sustainable energy savings, the generation of performance metrics is always an important tool for owners and utility programs alike. RCx providers have the data and the expertise to develop these metrics at the completion of a project after savings have been verified. It is important that the metrics presented to building owners and operators are appropriate for detecting deviant operation relating to all the energy efficiency measures that were implemented for the facility. The metrics should be as simple and straightforward as possible, with due consideration given to the amount of time, expense, and effort that the facility has available to monitor performance.