FEMP Federal Energy Management Program

# Supplement to M&V Guidelines: Measurement and Verification for Performance-Based Contracts Version 4.0

September 2023

# Disclaimer

# Acknowledgments

Contributors to this document include Christine Walker of Pacific Northwest National Laboratory, Phil Coleman of Lawrence Berkeley National Laboratory, and Phil Voss of National Renewable Energy Laboratory. Other materials cited or used here were developed by various industry-government working groups and the authors of FEMP Guidelines v3.0 (Lia Webster, James Bradford, Dale Sartor, John Shonder, Erica Atkin, Steve Dunnivant, David Frank, Ellen Franconi, David Jump, Steve Schiller, Mark Stetz, and Bob Slattery).

# Nomenclature or List of Abbreviations and Acronyms

| | |
|---|---|
| AHU | air handling unit |
| ASHRAE | American Society of Heating, Refrigerating, and Air Conditioning Engineers |
| $C_v$ | coefficient of variation |
| CEA | President's Council of Economic Advisers |
| DOE | U.S. Department of Energy |
| ECM | energy conservation measure (generally inclusive of water conservation measures, also) |
| EERC | Energy Escalation Rate Calculator |
| EIA | Energy Information Administration |
| EMCS | energy management control system (also referred to as BAS, building automation system) |
| ESCO | energy service company |
| ESPC | energy savings performance contract |
| FEMP | Federal Energy Management Program |
| HDD | heating degree day |
| HRU | heat recovery unit |
| HVAC | heating, ventilation, and air conditioning |
| IDIQ | indefinite delivery-indefinite quantity |
| IGA | investment grade audit (also referred to as feasibility study) |
| IPMVP | International Performance Measurement and Verification Protocol |
| M&V | measurement and verification |
| MATOC | multiple award task order contract (USACE) |
| MBE | mean bias error |
| NEMA | National Electrical Manufacturers Association |
| NIST | National Institute of Standards and Technology |
| O&M | operations and maintenance |
| RFP | Request for Proposal |
| RMS | root mean square |
| RMSE | root mean square error |
| TES | thermal energy storage |
| TMY | typical meteorological year |
| UESC | utility energy service contract |
| USACE | US Army Corps of Engineers |
| VAV | variable air volume |
| VSD | variable speed drive |

# Executive Summary

Measurement and verification (M&V) are key to determining and confirming the continued operation of, and savings associated with, performance-based contracts. The Federal Energy Management Program (FEMP) updates these guidelines to reflect the evolution of M&V to incorporate industry best practices and ensure that savings from these performance-based projects are properly planned, communicated, and realized.

FEMP released *M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0* in 2008. This document was developed with input from industry-government working groups, previous M&V guideline authors, and national laboratory subject matter experts. The document provided the methods and guidelines, with a focus on federal energy savings performance contracts (ESPCs). When FEMP released *M&V Guidelines: Measurement and Verification for Performance-Based Contracts Version 4.0* in 2015, FEMP expanded the intended audience to include anyone (energy managers, procurement officers, and contractors) involved in implementing performance-based contracts. The streamlined M&V Guidelines Version 4.0 provided the procedures and guidelines needed to quantify savings resulting from installing energy-efficient equipment, water conservation, improved operation and maintenance (O&M), renewable energy, and cogeneration projects with a performance-based contract; however, it removed some of the associated details on topics such as developing regression models and sampling plans. The details and information provided in M&V Guidelines v. 3.0 are still pertinent to M&V for performance-based contracts.

The objective is to keep FEMP M&V version 4.0 condensed, providing the reference document for specifying M&V methods and procedures and a resource for developing project-specific M&V plans; this supplement provides additional details. The content of this supplement has been updated to include pertinent revisions between M&V version 3.0 and M&V version 4.0 by providing details and updating the content to ensure relevancy.

The references have been updated to include the most current versions of existing policy documents, along with new references tied to new information in the guide.

# Table of Contents


Executive Summary ..... v

Introduction ..... 1

Intent of this Document ..... 1

1 Measurement & Verification Plan Details ..... 2

1.1 Defining the Baseline ..... 3

1.2 Adjustments ..... 4

1.2.1 Routine Adjustments ..... 4

1.2.2 Non-Routine Adjustments ..... 5

1.3 Interactive Effects ..... 5

1.4 Metering ..... 6

1.4.1 Equipment ..... 7

1.4.2 Sensor and Meter Accuracy and Calibration ..... 8

1.4.3 Metering Duration ..... 8

1.4.4 Sampling ..... 9

1.5 Energy Costs ..... 10

1.5.1 Utility Escalation Rates and EERC ..... 11

1.6 O&M and other Energy-Related Savings ..... 12

1.7 Measurement & Verification Reporting ..... 13

2 Developing Regression Models ..... 14

2.1 Independent Variables ..... 14

2.2 Choosing a Model ..... 15

2.3 Weather Data ..... 16

2.4 Documentation ..... 17

2.5 Savings Determination ..... 17

3 Sampling Guidelines ..... 18

3.1 Sampling ..... 18

3.1.1 Mathematical Methods for Sampling ..... 19

3.1.2 Sample Size Calculation ..... 20

3.2 Application of Sampling to Projects ..... 22

3.2.1 Compile Project/ECM and M&V Plan Information ..... 23

3.2.2 Designate Sampling Groups ..... 23
3.2.3 Select Samples ..... 24
3.3 Collect and Analyze Sample Data ..... 25
3.4 Extrapolate the Result from the Sample Over the Entire Population ..... 26
3.5 Final Note on Sampling ..... 27
4 Review and Oversight of M&V Activities ..... 28
4.1 Witnessing of M&V Activities ..... 28
4.2 Using the Review Checklists and Report Templates ..... 29
4.2.1 Project Documentation Needed ..... 29
4.2.2 Using the Checklists ..... 29
4.2.3 Summarize Findings in Evaluation Report ..... 30
4.3 Reviewing M&V Plans ..... 30
4.3.1 Prescriptive and Qualitative Evaluation Items ..... 30
4.4 Reviewing Post-Installation and Annual Reports ..... 32
4.4.1 Prescriptive and Qualitative Evaluation Items ..... 33
References ..... 0

# List of Figures

Equation 1-1: General Equation Used to Calculate Savings ........................................................ 4
Equation 1-2: Expanded Equation Used to Calculate Savings ..................................................... 4
Figure 3-1 Population and Sample.......................................................................................... 18
Figure 3-2 Normal distribution with confidence interval .......................................................... 20

# List of Tables

Table 2-1 Statistical Validations Guidelines........................................................................ 16
Table 3-1 First Year ($C_v$=0.5) Sample Size Table based on Usage Group Sampling .................. 24
Table 3-2 Example Inputs for Calculation of Monitoring Sample for Complex A ...................... 26
Table 3-3 Monitoring Results Based on Usage Group Sampling in the First Performance Period ........................................................................................................................... 27
Table 3-4 Revised Sample Requirements Using Usage Group Sampling ................................... 27

# Introduction

This document assumes that the reader is familiar with the M&V methods, risk and responsibility, and general considerations when selecting an M&V approach, which is covered in *M&V Guidelines: Measurement and Verification for Performance-Based Contracts Version 4.0* (M&V 4.0). The chapters included in this supplement cover additional details originally provided in *M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0* (M&V 3.0) and that are not specifically covered in M&V 4.0. Topics for each chapter are described below:

Chapter 1. Measurement & Verification Plan Details

Chapter 2. Developing Regression Models

Chapter 3. Sampling Guidelines

Chapter 4. Review and Oversight of M&V Activities

Appendix A: Additional Resources

M&V 3.0 was released in 2008 as an update of Version 2.2, which was published in 2000. Version 2.0 was published in 1996. M&V 3.0 incorporated significant updates to the 2000 version, with much of the guidance developed by a variety of industry-government working groups, facilitated by the Department of Energy (DOE).

## Intent of this Document

This document is not intended as a duplication of M&V 4.0, but rather a supplement containing additional details that were available in M&V 3.0. Some of these were modified slightly to reflect updates captured in M&V 4.0. It provides additional details on regression models, sampling methods, and specific activities associated with witnessing and M&V activities. This document is intended to supplement M&V 4.0.

# 1 Measurement & Verification Plan Details

The measurement and verification (M&V) plan is a document that defines project-specific M&V methods and techniques that will be used to determine savings resulting from a specific performance contracting project. The plan may include 1) a single option that addresses all the measures installed at a single facility, or 2) several M&V options to address each individual measure installed at the facility.

In addition to providing accurate and conservative methods to calculate energy savings, a good M&V plan is clear, consistent, and repeatable. In a long-term contract, it is very important to ensure that all assumptions, procedures, and data are recorded properly so they may be easily referenced and verified by others. The data included should be sufficient for a third party to implement or verify the M&V procedures.

M&V activities include site surveys, energy measurements, metering of key variables, data analyses, calculations, quality assurance procedures, and reporting. All these key components need to be adequately detailed in the M&V plan.

A project-specific M&V plan must be submitted and approved by the customer before M&V activities begin. In some cases, the customer will specify an approach in the task order request for proposal (RFP), while in other cases the energy service company (ESCO) will propose a site-specific plan for approval. Final resolution of M&V and other project issues are left to the discretion of the customer, although the details of the M&V plan can be a highly negotiated item. The project M&V plan includes project-wide items as well as details for each energy conservation measure (ECM). A graded approach to M&V is recommended, with the greatest rigor applied to the ECMs that provide the most savings.

In general, the contents of a project-specific M&V plan should:

- Provide an overview of the ECM and verification activities, including:
  - State the goals and objectives of the verification activities
  - Define the M&V "option" and techniques to be used for each measure
  - Identify the key physical characteristics of the facility, system, and ECM to be installed
  - Define the critical factors that affect energy consumption of the system or ECM
- Adequately define the baseline conditions, including:
  - Identify the key baseline performance characteristics of the system or ECM, such as lighting intensities and temperatures
  - Define baseline operating conditions, such as loads and hours of operation
  - Define operations and maintenance (O&M) costs

- Detail all measurements, data analyses procedures, algorithms, and assumptions
- Define all performance period verification activities, including:
  - Specify the parameters to be measured, period of metering, accuracy requirements, calibration procedures, metering protocols, sampling protocols, and archiving requirements
  - Explain requirements for customer witnessing of M&V activities
- Detail the schedule for periodic M&V reports and procedures
- Describe procedures and details for annual inspections
- Describe O&M reporting requirements by customer and ESCO
- Detail how savings will be calculated, including:
  - Provide rationale and procedures for any baseline or reporting period energy adjustments anticipated
  - Detail how interactive effects will be handled

The first step in defining a project-specific M&V plan involves selecting an appropriate M&V approach or approaches. This process is discussed in detail in M&V 4.0 and includes evaluating project-wide and ECM-specific objectives and constraints, assessing the viability of various M&V options, ascertaining savings risks, and evaluating implementation costs.

The following sections discuss and provide insight into the key areas covered by an M&V plan. Procedures for reviewing M&V plans for content and quality are detailed Section 4 of this document, and links to resources for review checklists can be found in Appendix A: Additional Resources.

## 1.1 Defining the Baseline

Since energy savings must be determined by comparing energy use before and after a retrofit, the characterization of the pre-retrofit or baseline conditions is critical. Defining the baseline consists of identifying the performance and operating factors that influence energy consumption and determining their values through observations and measurements.

Regardless of the M&V option or method used, the baseline conditions for all projects and ECMs must be adequately defined. Typically, the ESCO will define the baseline conditions during the investment grade audit, but the customer may define baseline conditions or certain parameters of the baseline conditions.

The purpose of establishing the baseline conditions is to:

- define the baseline sufficiently for purposes of calculating savings; and

- document the prevailing environment in case operational changes occur after ECM installation that mandate adjustments to the baseline energy use to make it comparable to performance period conditions.

Baseline conditions include physical, operational, and energy use data on the facility and systems. Baseline conditions are typically determined through surveys, inspections, and spot and short-term metering activities. Typically, pre-installation metering is conducted for a period of time required to capture all operating conditions of affected systems and/or processes.

Physical conditions that should be documented include equipment inventories, locations, nameplate data, system design features, and building occupancy. The key operational conditions include control strategies, set points, operating schedules, condition of equipment, loads, maintenance procedures used, peripheral equipment conditions, and weather. Energy use data that constitute the baseline may include utility billing data, sub-metered system data, and utility rate structures.

Although only a portion of a facility's systems may be included in the ESPC project, it may be appropriate to document the site conditions for other key energy-using systems. This is especially true if a whole-building M&V approach (Option C or D) is being used. Often, changes outside the scope of the ESPC project at a large facility can affect the overall energy consumption at a site and may warrant an adjustment, as discussed in Section 1.2.

## 1.2 Adjustments

Changes made to the baseline and/or the performance period energy usage can be used to account for changes at a facility, within a building, or for a specific system or equipment. Adjustments are sometimes required to account for changes unrelated to the ECM that affect energy use. Such adjustments may account for changes in weather, occupancy, or other factors between the baseline and performance periods.

**Equation 1-1 General Equation Used to Calculate Savings**

$$\text{Savings} = (\text{Baseline Energy} - \text{Post Installation Energy}) \pm \text{Adjustments}$$

The purpose of adjustments is to express both baseline and post-installation energy under the same set of conditions. The modifications to the savings can be further distinguished as routine and non-routine adjustments, as shown in Equation 1-2.

**Equation 1-2 Expanded Equation Used to Calculate Savings**

$$\text{Savings} = (\text{Baseline Energy} - \text{Post Installation Energy}) \pm \text{Routine Adjustments} \pm \text{Non-routine Adjustments}$$

### 1.2.1 Routine Adjustments

Routine adjustments are used to account for expected variations in independent variables and energy use. These adjustments often use regression analysis to correlate and adjust energy use to independent variables such as weather, but simple comparisons may also be employed. Routine adjustments are used to normalize energy use as a function of one or more independent parameters such as temperature, occupancy, or meals served.

Normalizing energy savings to a prescribed set of conditions is a very important technique used in ESPC projects. Using the same set of conditions for both the baseline and performance period cases, such as characteristic weather conditions (e.g., TMY weather data) and the corresponding cooling load profile, allows the risks associated with these operational factors to be reduced.

One of the key assumptions made when normalizing savings is that the performance period energy use will have a predictable relationship to the independent variables. The baseline model will be completely defined in the contract but will need to be populated by measured data collected during the performance period. Typically, a valid baseline model assumes that performance period data can be successfully plugged in, without disruption of the key relationships among independent variables. If performance period conditions are used to adjust the baseline case (as is often practiced in actual project M&V), the savings calculated will estimate the actual avoided energy use for that period.

Once the baseline and performance period models of the equipment's energy consumption and parameter(s) are established and validated, the standardized values of the independent parameters can be used to drive both models and calculate savings.

Therefore, a project M&V plan should identify critical independent variables, explain how these variables will be measured or documented, and discuss how they will be used in the empirical models. Additionally, assumptions and mathematical formulas used in the M&V plan must be clearly stated, and the validity of any mathematical model used should be verified. The verification strategies discussed in M&V 4.0 Section 4 can be applied to any mathematical model.

### 1.2.2 Non-Routine Adjustments

*Non*-routine adjustments are used to compensate for unexpected changes in energy-driving factors, such as facility size, operating hours, and facility use. These variables – referred to as "static" factors – are assumed to stay constant over the performance period but should be monitored to ensure that they are not changing, and thus affecting the performance of the ECM(s). Tracking these factors is primarily a concern for projects using whole-building options (Options C & D). Option A (and to a lesser extent B) approaches typically avoid these types of adjustments 1) the ECM is being evaluated in isolation, and b) many of these static factors (such as hours of operation) are stipulated by the customer and ESCO and documented as part of the M&V plan. If future changes (e.g., of facility size) are expected, the M&V plan should incorporate methods for making these non-routine adjustments.

## 1.3 Interactive Effects

It is commonly understood that energy systems interact with one another. Reduced lighting loads, for example, can reduce air conditioning energy consumption (a cooling bonus), but increase heating consumption (a heating penalty). Whole-building M&V approaches such as building simulation or utility billing analysis account for these types of interactive effects, whereas retrofit isolation M&V approaches do not.

When using retrofit isolation M&V (Options A and B), careful consideration must be given to dealing with interaction between ECMs. One must properly account for interactive effects and avoid double-counting of savings, which can occur inadvertently if interactions are not carefully considered.

For example, if the lighting retrofit mentioned above is accompanied by a chiller replacement, care must be taken to account for the reduced cooling loads on both the new and existing chillers due to the change in lighting.

In general, the possibility of double-counting energy savings can be reduced by considering one ECM at a time. The later ECMs should start from (i.e., their baseline condition should represent) the performance period condition of the previous ECMs. For related ECMs, such as lighting efficiency and lighting controls, double-counting can sometimes be avoided by using a single equation to determine savings from both measures.

Methodologies for determining some of the more common interactions, such as lighting and HVAC, have been developed (see M&V 4.0 Section 6: Guidance for Specific ECMs). However, detailed relationships between many dissimilar but interactive ECMs are not known, and the methods for measuring interactive effects are not cost-effective for many applications. For projects using retrofit isolation approaches (Options A or B), one of three approaches can be taken to account for savings associated with interactive effects between ECMs:

- Ignore interactive effects. If they are trivial and hard to decipher, this can sometimes make sense.
- Use mutually agreed-upon values that are based on the specifics of the building and HVAC equipment types. The values can be developed on the basis of computer model simulations for typical building conditions or assigned on the basis of available information for typical buildings. For instance, the heat generated from lighting is usually a simple multiplication exercise whose factors have often already been determined for the purposes of calculating the lighting savings.
- Develop a site-specific method to measure and estimate interactive effects. The customer and/or ESCO will need to agree on the merit and reasonableness of the proposed approach, which may include directly measuring the effects.

## 1.4 Metering[1]

To determine energy savings, some measurement processes need to be conducted to identify the pre- and post-retrofit conditions. These measurements typically include energy consumption and

[1] More information on metering is available in *Metering Best Practices: A Guide to Achieving Utility Resource Efficiency, Release 3.0*, Federal Energy Management Program, March 2015.

energy-related variables. Metering issues that should be considered in preparing a project-specific M&V plan are discussed below.

A project M&V plan should demonstrate that any metering and analysis will be done in a consistent and logical manner and with a level of accuracy acceptable to all parties. Metering and monitoring reports must specify exactly what was measured, how and when the measurements were made, what meter or meters were used, and who conducted these measurements. Any metering protocols that will be followed must be specified.[2]

Issues covered below include types of meters, meter accuracy and calibration, metering protocols, duration of metering, and the use of sampling.

### 1.4.1 Equipment

Many tools are available that help collect and analyze system-wide HVAC, controls, and lighting data. Data may include power (kW), energy (kWh), and operating parameters such as temperature, humidity, pressure, flow rates, status, and lighting levels. Data can be collected through one-time measurements or can be recorded in user-defined intervals. Prices, applications, and complexity of these tools vary.

For data collection, storage, and reporting, there are two general categories of metering equipment for M&V activities: data loggers and energy management control systems (EMCSs).

Data loggers range from simple battery-powered portable devices to more complex tools that can collect inputs from up to 30 transducers. The simplest portable data loggers collect information about a single variable (such as light fixture on/off status or amp draw from a motor). Others can capture multiple inputs (such as voltage, power factor, and amperage) and perform some calculations. Portable data loggers tend to be inexpensive per unit but are more limited in applications than permanent ones (or additional sensors for an EMCS). The downloading of data is usually done manually off site through a connection to a personal computer, although modem connections are sometimes used. Battery-powered portable loggers can offer non-intrusive monitoring within an occupied area, are relatively simple to use, and are inexpensive. More complex data loggers can collect information from a range of different inputs, conduct some analyses, prepare reports, and, typically through modems, download information for remote data collection. Permanently installed data loggers tend to be relatively expensive (when transducer and installation costs are included) and, if hard-wired, not very portable, which is an issue when only short-term measurements are required.

EMCSs are used for controlling energy systems. These would logically be an excellent option since such systems are often already in place and have data collection, trending, and computing

[2] Metering protocols are standardized procedures developed for measuring physical characteristics and metering specific types of equipment. For example, ASHRAE Guideline 14 Annex E describes standard procedures for measuring physical characteristics, including power, temperature, flow, pressure, and thermal energy and describes standards for measuring the performance of chillers, fans, pumps, motors, boilers/furnaces, and thermal storage.

capability; however, caution should be exercised, as many systems are not designed for data storage and reporting, and many operators are not familiar with M&V requirements.

### 1.4.2 Sensor and Meter Accuracy and Calibration

Before any data are collected, all sensors and meters should be reviewed to ensure that they are appropriate for the application. The accuracy of the device used to collect data can increase the level of error that is introduced in any calculations. Often, measurement error will be the primary source of uncertainty in a savings value. Using high-quality sensors for gathering key data can help increase the accuracy of savings estimates. Measurement uncertainty is discussed in detail in M&V 4.0 Section 5.3.

Equipment accuracies provided by the manufacturer are meaningful only if the equipment is in calibration. Sensors and meters used to collect M&V data should be calibrated to known standards (such as those of the National Institute of Standards and Technology). Forms indicating that calibration has been conducted are a required part of the M&V reports.

For the calibration to be valid, the equipment used to calibrate the sensors and meters must be of a greater accuracy than the sensors or meters themselves. Calibration methods for a variety of applications are included in ASHRAE Guideline 14.

### 1.4.3 Metering Duration

The duration of metering and monitoring must be sufficient to ensure an accurate representation of the amount of energy used by the affected equipment both before and after project installation. The appropriate measurements should be taken within representative time period. These measurements can then be used to determine time-of-use, seasonal and annual energy consumption. The time period of measurement must be representative of the long-term (e.g., annual) performance of the ECM or system. For example, lighting retrofits in a 24-hour warehouse that is operated every day of the year may require only a few days of metering. However, a chiller retrofit may require metering throughout the cooling season or perhaps for one month each season of the year.

The required length of the metering period depends on the type of ECM(s) or system. Some common scenarios are discussed below.

- For equipment that operates according to a well-defined schedule under a constant load, such as a constant-speed exhaust fan motor, the period required to determine annual savings could be quite short. In such a case, short-term energy savings can be extrapolated easily to the entire year.
- If the project's energy use varies across both day and season, as with most air-conditioning equipment, a much longer monitoring period may be required to characterize the system. In a case like that, long-term data are used to determine annual energy savings. When the metering is complete, the limits of the model used to characterize the system must be defined. For example, if data were taken on the chiller

system only when the outside air temperature ranged from 50°F to 70°F, then the resulting chiller model would probably be valid only at those temperatures.

- For some types of projects, metering time periods may be uncertain. For example, there remains a diversity of opinions over how long lighting operating hours must be measured in office buildings to obtain a representative indication of annual operating hours. In these situations, an agreement is required between the project parties to determine the appropriate measurement period and accuracy level for the ECM(s) or systems under consideration. For most lighting projects, three weeks of monitoring during non-holiday periods is typically effective.
- For some projects, the metering time period can be reduced by forcing a system to go through all of its operating modes in a short period of time. For example, a variable-speed drive ventilation system that is controlled by outside air temperature may require months of data collection to capture a full range of performance data. However, if the control system were overridden to force it to operate in various modes, the data collection might take only one day. This approach should be used with caution, as additional monitoring may be required to determine the system's relationship to independent variables.

### 1.4.4 Sampling

Sampling techniques should be used when it is unrealistic to monitor every piece of equipment affected by a retrofit. The sampling procedures outlined in Section 3 ("Sampling Guidelines") provide guidance on selecting a properly sized random sample of equipment for monitoring energy-related factors such as operating hours, loads, or power draw. The measurements taken from a sample of equipment can then be used to estimate the energy-related factors for the entire population.

A successful sample will be sufficiently representative of the population to enable one to draw reliable inferences about the population as a whole. The reliability with which the sample-based estimate reflects the true population is a function of specified statistical criteria, such as the confidence interval and precision level, used in the sample design. The reliability of a sample-based estimate can be computed only after the metered data have been collected. Before collecting the data, one cannot state the level of reliability that a given sample size will yield. However, one can compute the sample size that is expected to be sufficient to achieve a specified reliability level. This is done by using projections of certain values and criteria in the sample size calculations.

Based on the data gathered for a selected period of time, the sample size required may be reduced or increased. If the projections were too conservative, the estimate will exceed the reliability requirements. If these projections prove to be overly optimistic, then the reliability of the estimates will fall short of the requirements, necessitating additional data collection to achieve the specified reliability level. This method of using projections to calculate the necessary sample size is the one adopted for these guidelines.

## 1.5 Energy Costs

The purpose of an ESPC is to reduce energy and/or water costs, including related operations and maintenance (O&M) costs for facilities. The M&V plan should be designed to provide energy, water, and operating savings information in such a way that cost savings can be reasonably estimated.

For example, energy cost savings will be calculated using energy savings and the appropriate cost per unit of energy saved. In most cases, the unit cost of energy will be based on the servicing utility's energy rate schedules at the time the project is implemented. The unit cost of energy that will be used in calculating energy cost savings each year during the performance period must be defined in sufficient detail in the contract to allow savings to be calculated using each of the factors that affect cost savings. These factors include items such as (for electric bills) kWh saved, kW reduced, power factor, kW ratchets, and energy rate tiers. If the rate uses time-of-use periods, the energy and demand savings may be calculated separately for each time-of-use period. More complex rates, such as dynamic (e.g., real-time or day-ahead) pricing or rates with demand ratchets, will usually require additional calculations.

Demand savings is generally based on peak demand reduction, typically the reduction in the utility-metered maximum monthly demand. as measured in 15-minute intervals. some utilities and independent system operators/regional transmission organization (such as PJM or Electric Reliability Council of Texas, ERCOT) determine their demand charge based on a customer's power draw coincident with the *system* peak demand. Each site, working with its ESCO, must determine how the demand reduction will truly affect the utility bill (and thus payments to the ESCO). Permanent demand reductions, such as those from a lighting retrofit, are the simplest to measure. But even those require a determination of the proportion of lights that are actually likely to be turned on at the point that monthly peaks are set – it's usually not 100% and is best calculated with the help of light loggers. For seasonal equipment such as most air-conditioning systems, the calculations are more challenging, as the reductions will likely be much more pronounced in summers and trivial or non-existent during winter.

For performance contracts with cost savings based on peak or billing period load reductions, an M&V method should be selected that provides energy savings data by time-of-use periods corresponding to the facility's rate structure. For example, at a prison, the water heating peak load might be 252 kW over a 2-minute averaging period, 228 kW over 15 minutes, or 192 kW using 60-minute time periods of analysis. Considerable error in cost savings estimates is introduced by measured data that do not correspond to the rate structure (15 minutes, in this case). Similarly with time-of-use demand charges, if the utility's peak demand period is from 9 A.M. to 5 P.M., any demand savings realized outside of these hours will not result in reductions in peak period demand charges.

When determining the value of the energy, caution should be exercised to ensure that a conservative estimate is used that will not overvalue savings. The marginal cost of energy (i.e., the actual cost for the last portion of energy used for each month) should be used, rather than average values. The marginal costs can be determined by simulating reductions of 1 kWh or 1 kW, for electricity, and then recalculating the bill. Because there can be many non-volumetric bill components such as fixed fees, it is important that average values are not used, and only commodity-based charges are included.

Rates, especially for electricity, can be very complex. Increasing or decreasing block rates, "customer baseline load" rates, and kWh-per-kW rates are particularly complicated ones that are still in use. The ESCO may propose rates to use, but it is up to the customer to ensure that the correct rates are applied (and that the arithmetic is correct). At a minimum, the ESCO must provide the unit cost of fuel for each source of savings in the M&V plan.

### 1.5.1 Utility Escalation Rates and EERC

For each project, the ESCO and customer must mutually agree upon both the unit cost of energy for each fuel source and any escalation factors that may be applied during the performance period. Escalation rates are often employed in long-term contracts to estimate the future values of energy more accurately. Although higher values of energy will provide better cash-flow for the project, overvaluing savings is a serious concern that can cause budgetary problems for the customer.

A common source on which to base energy price escalation rates is the energy price projections of the DOE Energy Information Administration (EIA).

The National Institute of Standards and Technology (NIST) annually develops a tool, the Energy Escalation Rate Calculator (EERC), that can calculate a single appropriate escalation rate to use over the entire contract term for each energy utility. The EERC uses percentage of base-year cost savings attributable to each fuel in the project, commercial or industrial rate type, project location, and start and duration of the contract term. It then retrieves the matching EIA rates and calculates the weighted average escalation rate in real terms (excluding inflation) and nominal terms (including inflation). EERC's default inflation figure is the long-term inflation rate published annually by the President's Council of Economic Advisers (CEA) and is also used by FEMP in its life-cycle cost analysis tools for federal energy and water conservation and renewable energy projects.

The EERC-calculated average annual escalation rate, when applied to the base-year costs or savings of ESPC projects, results in the same future total amounts over the contract period as do the EIA-projected rates, which sometimes vary through the years, making the single EERC rate an ideal proxy for escalating contract payments. EERC is an on-line tool available from NIST's web site[3], and available through FEMP's Energy Management Tools web page.[4] It is generally updated in April of each year with the latest EIA energy price data and CEA long-term inflation rate.

[3] https://pages.nist.gov/eerc/
[4] https://www.energy.gov/femp/federal-energy-management-tools

The EERC tool focuses on energy prices. There is currently no equivalent governmental tool for forecasting water and sewer prices or O&M costs. Escalation rates for these costs should be determined separately from energy rates since their price paths may be substantially different from inflation in energy costs. Links to more specific guidance on how to determine utility rate estimations and escalations is located on the FEMP website and included in Appendix A: Additional Resources.

## 1.6 O&M and other Energy-Related Savings

Energy-related cost savings can result from avoided expenditures for operations, maintenance, equipment repair, or equipment replacement due to the ESPC project. This includes capital funds for work (e.g., equipment replacement) that, because of the ESPC project, will not be necessary. Sources of energy-related savings include:

- Avoided current or planned capital expense
- Transfer of responsibility for O&M and/or equipment repair and replacement (R&R) to the ESCO
- Avoided renovation, renewal, or repair costs as a result of replacing old and unreliable equipment

O&M and other energy-related cost savings are allowable in federal ESPCs and utility energy service contracts (UESCs) and are defined as reduction in expenses (other than energy cost savings) related to energy and water consuming equipment. Specific guidance on documenting and verifying O&M savings in federal ESPCs, and equally applicable to UESCs, was developed by FEMP with industry input. The link to the resulting document, *How To Determine and Verify Operating and Maintenance (O&M) Savings in Federal Energy Savings Performance Contracts*, is included in Appendix A: Additional Resources.

Any savings claimed from O&M activities must result in a real decrease in expenditures. O&M budget baselines cannot be based on what the customer should be spending for proper O&M; baseline expenditures must be based on what the customer is actually spending. The customer's O&M expenditures after implementation need to decrease for savings to be considered real, and should be documented (e.g., the continued reduction in an O&M contract can be checked and reported on by the ESCO as proof or savings persistence). "Savings" due to redirected labor or O&M efforts that do not reduce real expenditures cannot be claimed as savings under the DOE ESPC program. For example, labor reductions for agency staff that free them up to pursue other tasks, though valuable, do not qualify as savings for the purpose of an ESPC if labor expenditures do not decrease.

The approach for calculating energy-related cost savings mirrors the concepts used for determining energy savings — performance-period labor and equipment costs are subtracted from baseline values, plus or minus any adjustments required. Similarly, determining the appropriate level of effort to invest in the M&V of energy-related cost saving is the same as for energy cost savings — the level of M&V rigor will vary according to 1) the value of the project and its expected benefits, and 2) the risk in not achieving the benefits. Accordingly, a graded approach towards measuring and verifying O&M and R&R savings is advisable.

Baseline O&M and R&R costs should be based on actual budgets and expenditures to the greatest extent practical. This essentially “measures” and documents the baseline consumption of these parts or services. The use of estimated expenditures should be avoided if at all possible. Performance period or baseline adjustments are used to reflect any site-specific factors that would affect costs.

Some additional key points to keep in mind are as follows:

- A customer or site decision to commit ongoing funds from O&M budgets towards ESPC project payments has a long-term impact and must be documented adequately for future customer or site staff in both the M&V plan and the annual reports.
- Customers should maintain O&M cost records that will be needed to document baseline O&M costs. These records should be included in the performance-based contract proposal.
- ESCOs should include detailed information in annual reports to clearly convey the source of O&M savings as well as sufficient data to verify any savings calculations performed.
- Using an Option B or continuous measurement approach to track ongoing O&M savings can be cumbersome to the customer or site because of the required record keeping and accounting for ongoing changes at the site.

Links to more specific guidance on how to determine and verify energy related savings, is included in Appendix A: Additional Resources.

## 1.7 Measurement & Verification Reporting

The M&V submittals detailed in the M&V plan are the post-installation report, the annual reports, and any additional periodic reports required. Reporting formats for these reports are outlined in M&V 4.0 Appendix D (post-installation) and M&V 4.0 Appendix E (annual M&V reports). Specific content requirements, schedules, and approval procedures for each ESPC project are defined in the project M&V plan. All of the submittals, however, should adhere to the same general standards and procedures.

In general, all relevant documentation should be included with the M&V submittals, and these data should be provided in both electronic and hard-copy formats, or as specified by the customer. When submitting an M&V report, ESCOs should provide the data supporting the M&V activities. Data formats should be specified in the M&V plan, and both the original data and final data analyses should be submitted in support of surveys, savings estimates, and calculations. Metered data must be furnished in formats that are usable by the customer and based on products or software that are publicly available. For billing analysis (Option C) and computer simulation (Option B) M&V methods, electronic and hard-copy input and output files must be provided. If special software products are required for the reading or analysis of ESCO submittals, the customer may reject the data or request that the ESCO supply the software.

Facility personnel ensure that the report and verification documentation are complete and accurate and in compliance with the contract and approved M&V plan. If the customer believes that the conditions at the site are not accurately represented by the ESCO's submittals, the ESCO should be allowed to address the problem and re-submit.

# 2 Developing Regression Models

All M&V options utilize models to predict the baseline and performance period energy use of the project or ECM based on the behavior of the appropriate independent variables. An independent variable is a parameter that is expected to change regularly and has a measurable effect on the energy use of a system or building. The models used to predict energy use, with the exception of Option D, which utilizes simulation software, are often mathematical equations derived through regression analysis that incorporates key independent variables. Regression models involve an evaluation of the energy behavior of a facility or system to determine how it relates to one or more independent variables (e.g., weather, occupancy, or production rate). Regression models are a technique often used to adjust baseline or performance period energy use to account for changes in weather, occupancy, or other factors between the baseline and performance periods. Proper applications of these routine adjustments are discussed in Section 1.2.

A number of industry guidelines, including International Performance Measurement and Verification Protocol (IPMVP®): Core Concepts[5] Section 12, Uncertainty Assessment for IPMVP[6] and ASHRAE Guideline 14-2023[7] Annex D, have additional details on developing models by means of regression analyses as well as techniques for validating these models.

## 2.1 Independent Variables

Typical independent variables that drive energy consumption and can be incorporated in regression models include outdoor temperature, other weather parameters (e.g., heating or cooling degree days), occupancy, operating hours, and other variable site conditions.

Data on independent variables may be from a third party or may be tracked using onsite data collection, depending on their nature. Weather data are typically more reliable when supplied by an independent source but should be validated with site data to ensure applicability.

Once the data have been collected, the mathematical model that is used to predict the baseline (or performance period) energy use is developed. The model should make intuitive sense—the

[5] International Performance Measurement and Verification Protocol (IPMVP®): Core Concepts, EVO 10000 – 1:2022, March 2022, Efficiency Valuation Organization.

[6] Uncertainty Assessment for IPMVP: International Performance Measurement and Verification Protocol®, EVO 10100 – 1:2018, April 2018, Efficiency Valuation Organization.

[7] *ASHRAE Guideline 14-2023: Measurement of Energy, Demand and Water Savings*, American Society of Heating, Refrigerating and Air-Conditioning Engineers.

independent variables should be reasonable, and the coefficients should have the expected sign (positive or negative) and be within an expected range or magnitude.

## 2.2 Choosing a Model

There are various forms of models used in standard statistical practice. Examples of multi-variant regression models are included in *Uncertainty Assessment for IPMVP* and *ASHRAE Guideline 14 Annex D*.

An example of a linear multi-variant regression model for a weather-dependent ECM is shown in Equation 2-1 below. In models using weather data, it can be beneficial to define a custom temperature base for calculating heating degree day (HDD) and cooling degree day (CDD) data based on the actual behavior of the building.[8]

Equation 2-1 Multi-Variant Regression Model for a Weather-Dependent ECM

$$E = B_1 + (B_2 \times T_i - T_{i-1})) + (B_3 \times HDD_i) + (B_4 \times CDD_i) + (B_5 \times X_1) + (B_6 \times X_2) + (B_7 \times X_3)$$

where:

E = energy use

i = index for units of time for period

$B_{1\text{-}7}$ = coefficients

T = ambient temperature

HDD = heating degree days using a base temperature of 60ºF

CDD = cooling degree days using a base temperature of 65ºF

$X_n$ = independent steady-state variables

It is important that the best model type be used, which in turn will depend on the number of independent variables that affect energy use and the complexity of the relationships. Finding the best model often requires testing several models and comparing their statistical results. The number of coefficients should be appropriate for the number of observations. Similarly, the form of the polynomial should be suitable to number of independent variables. Additionally, the independent variables must be truly independent of one another. The model should be tested for possible statistical problems (e.g., auto-collinearity[9]) and corrected.

Validation steps should include checks to make sure that statistical results meet acceptable standards. The statistical requirements outlined in Table 2-1 are examples of validation standards for mathematical models using typical statistical indicators. An example application of these indices is included in the Standard M&V Plan for Chiller Replacements found in Appendix B. The statistical validity of models can be demonstrated using other published sources such as

[9] Auto-collinearity can result when one or more important independent variables were left out of the model.

ASHRAE Guideline 14. Specific goals should be set for validating mathematical models used in each project based on suitable levels of effort (see M&V 4.0, Chapter 5) and should be specified in the M&V plan. Many analysis tools provide some of these statistical results, while others will need to be calculated.

**Table 2-1 Statistical Validations Guidelines**

| Parameter Evaluated | Abbreviation | Suggested Acceptable Values | Purpose |
|---|---|---|---|
| Coefficient of determination | $R^2$ | > 0.75 | Indicates model's overall ability to account for variability in the dependent variable. Lower $R^2$ values may indicate independent variables are missing or additional data are needed. |
| Coefficient of variation of root-mean squared error | CV(RSME) | < 15% | Calculates the standard deviation of the errors, indicating overall uncertainty in the model |
| Mean Bias Error | MBE | +/- 7% | Overall indicator of bias in regression estimate. Positive values indicate higher than actual values; negative values indicate that regression under-predicts values. |
| t-statistic | t-stat | > 2.0 | Absolute value > 2 indicates independent variable is significant |

## 2.3 Weather Data

If the energy savings model incorporates weather data, several issues should be considered:

- The relationship between temperature and energy use may vary depending upon the time of year. For example, an ambient temperature of 55°F in January has a different implication for energy usage than the same temperature in August. Thus, seasons may need to be addressed in the model.
- The relationship between energy use and weather may be nonlinear. For example, a 10°F change in temperature may result in a very different energy use impact if that change is from 75°F to 85°F rather than 35°F to 45°F.
- Matching degree-day and other data with billing start and end dates.
- Typical meteorological year weather data are available from the National Renewable Energy Laboratory[10].

[10] https://nsrdb.nrel.gov/, Datasets, TMY

## 2.4 Documentation

All models should be thoroughly documented including specifying model limits. Ideally, the range of values of the independent variables used to create the model span the entire range of possible conditions. Models are generally good only for the range of independent variables used in creating the regression model.

The criteria used for identifying and eliminating any available data must be documented. Outliers are data beyond the expected range of values (or two to three standard deviations away from the average of the data). The elimination of outliers, however, should be justified by abnormal or specific mitigating factors. If a reason for the unexpected data cannot be found, the data should be included in the analysis. Outliers should be defined using common sense as well as common statistical practice.

## 2.5 Savings Determination

In general, the procedure for determining energy savings with a regression model is as follows:

- Develop and validate an appropriate baseline model relating the baseline energy use during normal operations to key independent variables.
- Install ECMs and continuously measure the independent variables used in the baseline model along with any additional variables that may be needed for performance period model development.
- Using the baseline model, estimate what the energy use would have been without the ECMs by driving the baseline model with the performance period weather or other independent variables.
- Calculate savings by comparing the *predicted* baseline energy use (i.e., that which the model predicts *would have* occurred given the performance period values of the independent variables) with the actual energy use of the performance period.

An alternative approach that is sometimes warranted (e.g., in new construction) includes creating a separate regression model to describe performance period energy use. In the case of new construction, the performance period model (reflecting the proposed design) is created, and then calibrated with actual utility data. Its predicted performance period energy usage is then subtracted from the predicted usage of a baseline model, for instance one that just meets the prevailing building energy code; the result is the savings estimate. This approach allows for calculation of normalized savings[11] based on a predefined set of parameters, such as typical weather. All independent variables and criteria for validating performance period models should be included in the M&V plan.

[11] See Section 1.2.

The best regression model is one that is simple and yet produces accurate and repeatable savings estimates. Determining the best model often requires testing several models to find one that is easy enough to use and meets statistical requirements for accuracy (see M&V 4.0, Section 4.1).

# 3 Sampling Guidelines

This section introduces the statistical background, theory and formulas used to select, analyze and validate samples for project monitoring and evaluation. It also provides guidelines and procedures for the design and implementation of sampling.

## 3.1 Sampling

The purpose of monitoring a sample, as opposed to an entire population, is to characterize a population with adequate accuracy while reducing costs and effort.

As shown in Figure 3-1, sampling involves selecting several members from a population for monitoring and evaluation. The measured characteristics of the sample group is then used to infer the characteristics of the entire population. As expected, the assumption is that the sample is representative of the population. To ensure that the sample is indeed representative, calculations must be performed to assess and quantify the statistical validity of the sampled data. These calculations are presented later in this section.

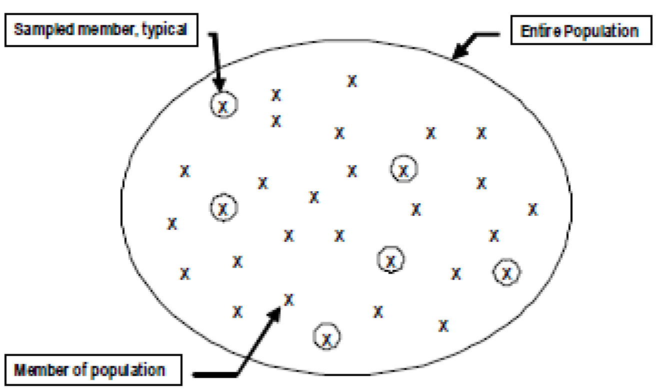


**Figure 3-1 Population and Sample**

Sampling is applicable to projects such as lighting retrofits, energy-efficient motor replacements, HVAC unit replacement, steam trap monitoring, or any other project in which a number of similar pieces of equipment are affected by the same ECM. In the most common applications, sampling strategies are used to characterize the hours of operation and the instantaneous power draw of a constant-load device.

When selecting a sample from a population to determine hours of operation, it is necessary to ensure that the load, or device being sampled, is monitored at or down-stream of its last point of control (LPC). The LPC is the portion of an electrical circuit (or other source of energy) that serves a set of equipment that is controlled on a single switch. As a result, all of the fixtures or pieces of equipment on that LPC are typically operated the same number of hours per year. For metering purposes, it is assumed that measurements taken of a single piece of equipment on an LPC capture the operating hours for all of the equipment served on the same circuit.

### 3.1.1 Mathematical Methods for Sampling

Sampling must be conducted using accepted methods and use an appropriate level of care to ensure that the M&V results that rely on the sampling and analysis are sufficiently accurate. This section provides a summary of the concepts, methods and equations to be used.

Although various assumptions regarding the distribution of sampled data can be made, the large majority of sampling statistical analysis assumes that the data are *normally distributed* about the *mean.* We also make that assumption in this section.

Statistical validity requires that samples be randomly selected. Use of a random number generator, such as that found in Microsoft Excel™, is convenient for ensuring the sample is randomly selected.

#### *3.1.1.1 Point Estimation – Confidence and Precision*

When we use sampling to estimate an average value of an entire population, we are performing an activity know as *point estimation.* A value or 'point' that is estimated based on a sample is not the actual average value, but rather a value that is "reasonably close" to the actual average. The question for the M&V practitioner is: "What do we mean by 'reasonably close'?" The question is answered using the following statistical terms:

- Confidence: Confidence is fundamentally the same as probability, except that confidence refers to data already obtained, while probability refers to a future value. A confidence of 90% is commonly used in M&V. So, using our 90% example, when we refer to a confidence level, we are saying "I am 90% *confident* that the measured value is within my stated *confidence interval*."
- Confidence interval (or precision): Because the value estimated by sampling is likely not the actual value, it is useful to state an interval in which we have confidence the true value lies. Confidence interval is also often referred to as *precision*. An M&V practitioner may state that they know an estimated value has a *precision* of 10%, which would mean that the true value is within 10% of the estimate.

Confidence and precision, then, are the values referred to when a 90/10 (or 80/20 or any other) criterion is specified.

**Example**

Imagine that we wish to measure the run-hours of a sample of equipment for one month. Imagine now that we measure 200 'on' hours. If we are hoping to meet a 90/10 criterion, we are hoping that we can say, with a 90% probability, that the actual value is within 10% of the estimated run hours – that is, we are 90% sure the actual runtime is between 180 and 220 hours.

To graphically illustrate the concepts of normal distribution, confidence, and precision/confidence interval, Figure 3-2 shows a normal distribution with a confidence interval. Note that the confidence interval in the figure is defined by the error (+/- E). This error figure is discussed further below and is defined in Equation 3-1.

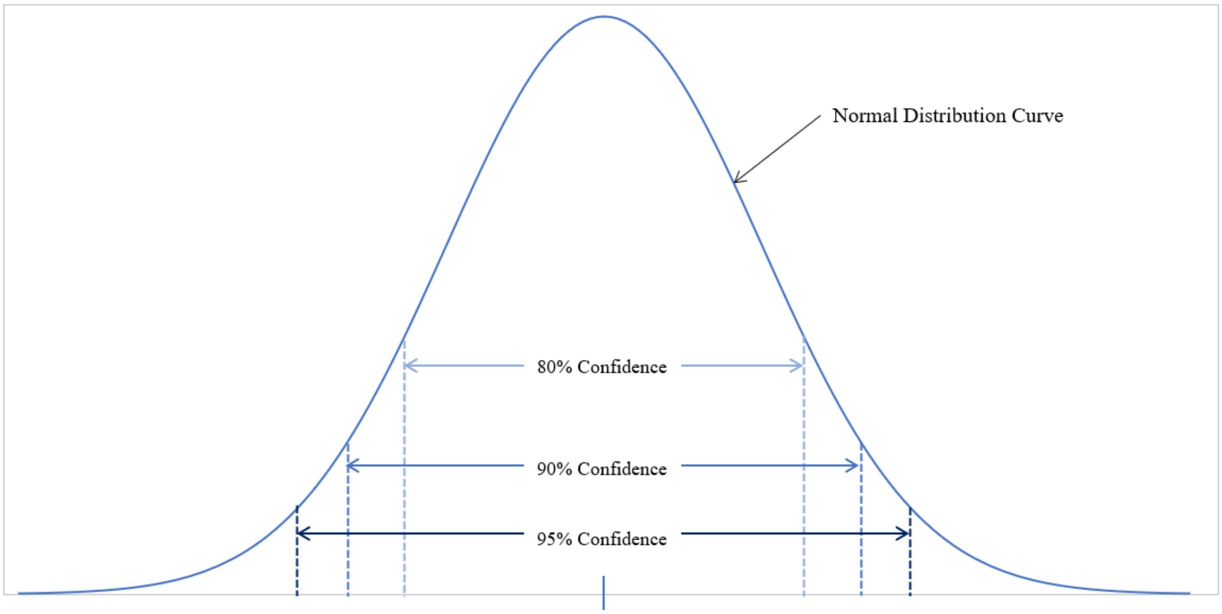


Figure 3-2 Normal distribution with confidence interval

The confidence interval (or precision) and the confidence level are positively linked; for any sample, as the confidence *interval* increases (that is, the precision is reduced, and the range of possible values of the true mean increases), the confidence *level* increases. Or, looking at it another way, as the confidence interval is reduced, the confidence level is also reduced.

### 3.1.2 Sample Size Calculation

When sampling, it is the M&V professional's job to meet certain levels of confidence and precision and to calculate the actual confidence and precision that result from a sampling exercise. In order to accomplish this analysis, it is helpful to start with Equation 3-1, the statistical equation for calculating the *maximum error* in the result. This value is also depicted as 'E' in Figure 3-2.

$$\mathrm{E} = z\frac{s}{\sqrt{n}}$$ **Equation 3-1**

Where:

E = maximum value of error

s = the standard deviation of the sample[12]

[12] Calculating the standard deviation is not defined herein

n = sample size

z = the z-statistic[13], typically denoted as $z_{\alpha/2}$.

Alpha (α), is equal to 1 – % confidence, and is used in most statistical references because it represents the top and bottom tails of the normal distribution, which together bound the confidence level.

The z statistic is a variable that is calculated such that the following equation is true: $z = \frac{\bar{\bar{\bar{x}}}-\mu}{s/\sqrt{n}}$

where μ is the population mean (unknown) and *x* is the sample mean. Although the true mean, μ, is unknown, values of z, at various levels of confidence (1 – alpha) are known and are tabulated in many statistics books.[14] The value of z -- assuming that the number of samples, n, is greater than 30 -- is 1.645 for 90% confidence and 1.282 for 80% confidence.

Rearranging Equation 3-1, we can solve for the number of samples needed to ensure we are within a certain confidence interval:

$$\boldsymbol{n = \frac{z^2 \times s^2}{E^2}}$$ Equation 3-2

Note in Equation 3-2 that the standard deviation of the sample, s, and the maximum allowable error, E, are in the relevant units of measurement (e.g., hours or kW). The standard deviation, s, of the sample can be expressed in terms of the *coefficient of variation* (or $C_v$), which is a fraction of the mean, as shown in Equation 3-3.

$$\boldsymbol{C_v = \frac{s}{\bar{x}}}$$ Equation 3-3

Where:

$\bar{x}$ is the sample mean

Similarly, the maximum error, E, can also be expressed as a fraction of the mean (precision), as shown in Equation 3-4.

$$P = \frac{E}{\bar{x}}$$ **Equation 3-4**

Substituting $C_v$ and P into Equation 3-2, we get a unit-less expression, as shown in Equation 3-5.

$$n = \frac{z^2 \times (s/\bar{x})}{(E/\bar{x})^2}$$ **Equation 3-5**

[13] A similar statistic, known as the t-statistic, assumes a 't-distribution' rather than a normal distribution, and is a function of the number of samples. This can be substituted for the z statistic for a slightly more accurate approach. Although the t-statistic is preferable for small populations that exhibit more 'spread', for samples large than 30, use of the normal distribution gives a good approximation of the t-distribution. Slightly larger samples than are prescribed using the normal distribution should be taken when the indicated samples are smaller (especially fewer than 30).

[14] For example: Statistics, 5th Edition, by Robert S. Witte and John S. Witte or Probability and Statistics for Engineers, by Iwrin Miller and John E. Freund

Or, expressed another way,

$$\boldsymbol{n} = \frac{z^2 \times (C_v)^2}{(P^2)} \qquad \text{Equation 3-6}$$

Where

z = Z-statistic (1.645 for 90% confidence, 1.282 for 80% confidence); and

P = Precision required, typically 10% or 20%.

Equation 3-6 is the basic equation used in sample group sizing. For small populations the sample size should be modified using the finite population correction shown in Equation 3-7. Typically, this correction is required when the population is less than 500. The finite population adjustment calculation gives $n^*$, which is the new sample size corrected for population size.

$$\boldsymbol{n}^* = \frac{Nn}{n+N} \qquad \text{Equation 3-7}$$

Where

$n^*$ = sample size corrected for population size

$n$ = sample size for infinite population

$N$ = population size

A critical step, which is often not completed, is the post-monitoring calculation of the **actual** *standard deviation*, *coefficient of variation*, and subsequent calculation of *precision* at various *levels of confidence* using the above equations. Ultimately, the maximum error (E) using Equation 3-1 should be calculated for various levels of confidence. The job is not complete until these post-monitoring calculations are completed and reported.

## 3.2 Application of Sampling to Projects

In the next sections, we explore considerations for the design and application of sampling. The analysis steps in conducting sampling are as follows:

1. Compile and analyze the project, ECM and M&V plan information
2. Designate sampling groups
3. Select samples
4. Collect and analyze sample data
5. Extrapolate the result from the sample over the entire population

These steps are discussed below.

### 3.2.1 Compile Project/ECM and M&V Plan Information

In this step, the goal is to fully understand several things, including: the measure scope, the savings calculations quantifying the intended performance, the M&V method to be used and the data to be collected. Once the project is understood, an M&V practitioner can identify the calculation method and select variables to be sampled.

In some energy conservation projects, it is worthwhile to conduct both pre- and post-installation sampling, for example in the case where fixture configurations may have changed. Regardless of whether the sampling is for evaluating the baseline or the post-retrofit conditions, the following information is typically required to properly assign usage groups and determine sample sizes.

- ***Number of circuits, devices or LPCs.*** Identify and document the LPCs that are affected by the installation of ECMs. This should be provided in the form of an equipment inventory survey in which each line in the survey represents an LPC and includes descriptions of affected and proposed ECM nameplate data and quantity, as well as location information.
- ***Actual load or wattage.*** Using the equipment inventory survey, the total change in load or wattage of the affected equipment by usage group can be computed.
- ***Hours of operation.*** Sampling can be used to estimate the average hours of operation of the equipment. After the first sampling period (whether it is a year, month or week), the sampling result (actual $C_v$, Equation 3-3) should be used to compute the sample size. If it is expected that the equipment will be used in a significantly different manner in the current period than it was in the previous period, the estimate may be adjusted.

### 3.2.2 Designate Sampling Groups

Each device or LPC should be assigned to a usage group based on similarities in the parameter being determined, such as operating hours or connected load. If usage groups are small because of considerable observed differences, the resulting variance of the data may result in unsatisfactory confidence and precision levels. So, while considering the tradeoffs, usage groups should be developed from criteria such as:

- Area type (for example, office, hallway, meeting room)
- Annual operating hours
- Timing / usage patterns
- Variability of operating hours, load, or another variable
- Similar functional use

Usage groups should be selected so that equipment or LPCs are similar in that the sampled value (for example, hours or kW) is clustered around a specific estimate. When possible, avoid designating usage groups with populations that will yield less than 10 sample points. Examples of standard usage groups for fan motors with similar operating hours are HVAC ventilation supply fans, return fans, and exhaust fans. Examples of standard usage groups to determine lighting operating hours are fixtures with similar operating characteristics in similar spaces (e.g., offices, laboratories, hallways, stairwells, common areas, etc.).

Usage groups may be defined for the population on a building-by-building basis or across a number of buildings with similar usage areas. Monitoring can be done for single or multiple buildings provided the usage groups are similar. Defining populations for multiple buildings is acceptable and usually results in fewer monitoring points than if each building were considered separately.

### 3.2.3 Select Samples

Select desired confidence and precision levels. A 90/10 confidence/precision level is commonly used in M&V, but others may be justified depending on customer desire.

*Establishing the Coefficient of Variation*. Prior to selecting a sample, an estimate of the coefficient of variation ($C_v$) must be made. A $C_v$ of 0.5 has been historically recommended, and numerous projects have shown this to be reasonable for most applications. After the first year of monitoring, the coefficient of variation for each usage group can be projected from the results of the metering in that year.

Having selected a confidence and precision level (90/10) and a $C_v$ (perhaps 0.5), use Equations 3-6 and 3-7 above to calculate a sample size for each sampling group. Then, randomly select that number of samples from the population. We strongly recommend over-sampling (at a 10% or greater level) in case of data collection device failure or unexpectedly high data scatter.

Table 3-1 illustrates the effect of confidence and precision on sample size.

**Table 3-1 First Year ($C_v$=0.5) Sample Size Table based on Usage Group Sampling[15]**

| Precision | 20% | 20% | 10% |
|---|---|---|---|
| Confidence | 80% | 90% | 90% |
| Z-Statistic | 1.282 | 1.645 | 1.645 |
| Population Size, N | Sample Size, n* | | |
| 4 | 3 | 4 | 4 |
| 8 | 5 | 6 | 8 |
| 12 | 6 | 8 | 11 |
| 16 | 7 | 9 | 13 |
| 20 | 8 | 10 | 16 |
| 25 | 8 | 11 | 19 |
| 30 | 9 | 11 | 21 |
| 35 | 9 | 12 | 24 |

[15] Table does not reflect over-sampling. However, because data collection problems are very common and also due to the departure from normal distribution for small samples (especially less than 30), over-sampling is a recommended practice.

| Precision | 20% | 20% | 10% |
|---|---|---|---|
| Confidence | 80% | 90% | 90% |
| Z-Statistic | 1.282 | 1.645 | 1.645 |
| Population Size, N | Sample Size, n* | | |
| 40 | 9 | 12 | 26 |
| 45 | 9 | 13 | 28 |
| 50 | 10 | 13 | 29 |
| 60 | 10 | 14 | 32 |
| 70 | 10 | 14 | 35 |
| 90 | 10 | 15 | 39 |
| 100 | 10 | 15 | 41 |
| 125 | 11 | 15 | 45 |
| 200 | 11 | 16 | 51 |
| 300 | 11 | 17 | 56 |
| 400 | 11 | 17 | 59 |
| 500 | 11 | 17 | 60 |
| infinite | 11 | 17 | 68 |

The samples in each usage group should be drawn at random[16], so that each member has an equal probability of being selected.

If there is reason to believe that there are significant seasonal variations in the operation of the equipment, sufficient monitoring will need to be conducted to capture these variations.

## 3.3 Collect and Analyze Sample Data

After metering has been completed, calculate mean, standard deviation and $C_v$ (Equation 3-3) of the collected data for each usage group. If the actual $C_v$ is equal to or less than the originally assumed $C_v$, then the desired confidence can be retained.

Using Equation 3-1, calculate the maximum error and precision/confidence interval at the selected confidence level. The confidence interval is then either accepted, or if it is too large, additional sampling (and possible sampling redesign) may be required. Once a sample has been selected and monitoring is done, the engineer has no say over the results, and can only report her findings and confidence in them.

[16] Random selection of monitoring points is critical to avoid bias in the sample. Spreadsheet or other computer software should be used to generate a list of random numbers that may be used to place loggers on a given LPC.

## 3.4 Extrapolate the Result from the Sample Over the Entire Population

Once the sample mean and standard deviation are known, the result can be applied to the entire population by assuming the mean of the sample is accurate for the entire population. For example, if the mean of the sample is x *kW* per unit, multiplying x by the number of units in the entire population gives the total kW.

**Example**

Usage group sampling can be applied to just one or numerous buildings that are similar in function, layout, and operation.

Suppose that an ESCO is retrofitting lighting fixtures in a large office complex containing six buildings that have similar floor plans, functions, and operating schedules. As shown in Table 3-2, usage group sampling is applied to each of the four usage groups that appear in the six buildings, and the sample size is 76 fixtures.

**Table 3-2 Example Inputs for Calculation of Monitoring Sample for Complex A**

| Usage Groups for Complex A | Number of Lighting LPCs (N) | | | | | | | Sample Size (90/20) n* + 10% (rounded) |
|---|---|---|---|---|---|---|---|---|
| BUILDING | A-1 | A-2 | A-3 | A-4 | A-5 | A-6 | All | |
| Offices | 400 | 350 | 450 | 440 | 350 | 450 | 2,440 | 19 |
| Hallways | 600 | 550 | 450 | 440 | 550 | 450 | 3,040 | 19 |
| Meeting Rooms | 150 | 200 | 200 | 160 | 200 | 200 | 1,110 | 19 |
| Other | 200 | 220 | 180 | 180 | 220 | 180 | 1,180 | 19 |
| Total | 1,350 | 1,320 | 1,280 | 1,220 | 1,320 | 1,280 | 7,770 | 76 |

*Note:* Sample (19 for each usage group, as shown above) should be distributed randomly across the sites.

The sampling procedure varies depending on whether it is the first monitoring period (no prior sampling data available) or a subsequent one:

- ***First Monitoring Period:*** Using Table 3-2 or Equations 3-6 and 3-7 and assuming $C_v$ of the hours = 0.5 to determine the sample size based on number of lighting areas (*N*) in each usage group, one obtains a total sample size of 76, as shown in Table 3-3.
- ***Subsequent Monitoring Periods:*** In the second and subsequent years, the same procedure is used to calculate the sample size, except that the actual value of $C_v$ will be determined from the data collected in the previous year's sample.

Suppose the ESCO obtains useful metered data for the required number of sample fixtures and computes the standard errors of the actual measured operating hours for each usage group, where the actual values are presented in Table 3-3. Using Equation 3-1, the standard error of the total estimated savings for each usage group can be calculated. The calculated values are shown in

Table 3-3. For two of the four usage groups, (hallways and meeting rooms), the actual metered standard error is greater than the allowable amount; thus, the reliability requirement is not met for these two groups.

**Table 3-3 Monitoring Results Based on Usage Group Sampling in the First Performance Period**

| Usage Groups for Complex A | Number of Samples Metered | Measured Annual Op Hours | Standard Deviation | Maximum Error | Allowable Error | Actual Precision at 90% Confidence | Reliability Requirement Met? |
|---|---|---|---|---|---|---|---|
| Offices | 19 | 5,256 | 1,314 | 495.9 | 1051.2 | 9.4% | Yes |
| Hallways | 19 | 7,008 | 5,605 | 2115.3 | 1401.6 | 30.2% | No |
| Meeting Rooms | 19 | 2,628 | 1,568 | 591.74 | 525.6 | 22.5% | No |
| Other | 19 | 1,752 | 701 | 264.5 | 350.4 | 15.1% | Yes |
| Total | 76 | | | | | | |

For the subsequent monitoring periods, a revised sample size is calculated from the metered data. The actual coefficients of variation (Equation 3-3) can be calculated from the standard deviation of the operating hours in each usage group divided by the average measured hours. These values for $C_v$ are used in Equations 3-6 and 3-7 to calculate a revised total sample size and allocation across usage groups. In this example, the revised rounded total sample size is 92. The allocation by usage group is presented in Table 3-4.

**Table 3-4 Revised Sample Requirements Using Usage Group Sampling**

| Usage Groups for Complex A | N | Original Sample Size | Measured hours | Actual $C_v$ | New Sample Size | New Sample Size n* + 10% |
|---|---|---|---|---|---|---|
| Offices | 2,440 | 19 | 5,256 | 0.25 | 4 | 5 |
| Hallways | 3,040 | 19 | 7,008 | 0.8 | 43 | 48 |
| Meeting Rooms | 1,110 | 19 | 2,628 | 0.6 | 24 | 27 |
| Other | 1,180 | 19 | 1,752 | 0.4 | 11 | 12 |
| Total | 7,770 | 76 | | | | 92 |

## 3.5 Final Note on Sampling

The purpose of sampling is to monitor a representative set of points rather than the entire population. The end result is to obtain reliable estimates within a specified precision with statistical confidence. Monitoring the specified number of points does not necessarily mean that compliance with project requirements has been obtained. Again, **the job is not done until post-monitoring calculations are completed and reported.**

Sample problems may include improperly designated usage groups, incorrect sample design assumptions, or selection of nonrandom points, all of which may lead to sample-based estimates that are biased or unreliable within specified levels. Data logger failure is common; therefore, over-sampling is usually warranted. It is critical to take care during the initial developmental stages to design a sample that truly reflects the project site. In any case, the M&V practitioner should use whatever reliable data are available.

# 4 Review and Oversight of M&V Activities

The purpose of customer oversight and review of M&V activities is to:

- Verify that all M&V activities are conducted in accordance with the M&V plan;
- Confirm that the reported results of inspections and measurements are accurate and represent actual operation of the equipment or systems involved;
- Confirm contractor payments are based on verified savings;
- Ensure that M&V activities are properly documented; and
- Follow up on any outstanding issues identified.
- Depending on available resources, the customer may seek outside assistance in the review of M&V reports, analyses, and results, which would need to be covered by savings generated from the implemented project. The customer representative who is responsible for oversight of M&V activities and review and acceptance of reported findings should have significant experience in the analysis, design, commissioning and/or measurement and verification of energy efficiency projects and be familiar with both the site and the details of the ESPC project. Detailed requirements for each project will be included in the project M&V plan.

Government agencies are expected to witness baseline, post-installation, first-year, and annual M&V activities and commissioning of installed ECMs and approve required submittals in writing. This requires that the agencies designate individual(s) to observe these activities, review the resulting M&V reports submitted by the ESCO, and certify in writing that those reports are acceptable to the agency.

## 4.1 Witnessing of M&V Activities

Witnessing of M&V activities by knowledgeable customer representative(s) is recommended primarily to:

- Assure that both customer and ESCO fully understand the results of the M&V activities,

- Provide increased confidence that savings expected under the ESPC are being achieved, and
- Make a direct link between payments made to the ESCO and the verification that savings are being achieved.

Active involvement by customer personnel in the verification of savings helps ensure that the statutory and regulatory requirements for ESPCs are met. The Federal Acquisition Regulations (FAR) provisions generally require federal agencies to verify in writing that applicable procurement terms and conditions have been met by a contractor prior to payment by the government.

Active participation in the M&V process by customer or site staff can reduce the number and intensity of disputes about performance. See Appendix A for links to FEMP resources on witnessing guidance for federal agencies.

## 4.2 Using the Review Checklists and Report Templates

Links to review checklists and templates for written reviews of M&V plans, and post-installation and annual reports are included in Appendix A ("Additional Resources"). These checklists and templates should be utilized to complete a thorough evaluation by a customer representative prior to accepting the submitted M&V plan or report.

### 4.2.1 Project Documentation Needed

Prior to conducting the review, ensure that all related project documentation is on hand. At a minimum for any ESPC project report or M&V plan review, the M&V plan, the final cost schedules, and the IGA results should be available. For review of annual reports, the post-installation report, any previous annual reports, and any contract modifications are needed. Savings calculations should be scrutinized and will often need to be reviewed in their electronic format. Missing documentation can cause confusion and lead to incorrect conclusions.

### 4.2.2 Using the Checklists

The layout of the checklists follows the prescribed outlines for M&V plans, Post-Installation and Annual Reports and each one has two parts - Project Level items and ECM Level items. Prior to beginning the review, determine the percent contribution of cost savings for each ECM in the project. The reviewer may customize the review checklists, which are available electronically through the links provided in Appendix A.

Read through the M&V submittal (Plan, Post-install, or Annual Report) while checking off topics and making notes in the customized checklists. Note the location of key items in the first column of the checklists (labeled "Reference Page") so they can be easily cross-referenced. The inability to comment on an item suggests that relevant information may be missing or not in complete form. Items in the checklist that require follow-up should be flagged by placing an "X" in the last column of the checklist (labeled "Follow-Up?") and noting the deficiency or issue identified in the adjacent column. Some of the items in the checklists are marked "Evaluation".

This indicates that additional qualitative assessment is necessary. These qualitative issues are discussed for each individual M&V submittal type in subsequent sections of this chapter.

#### 4.2.3 Summarize Findings in Evaluation Report

After reading the M&V submittal (Plan, Post-Installation or Annual Report), filling out the Review Checklists, and evaluating the qualitative issues, the findings from the review should be summarized in a written report. The written review should follow the format in the appropriate Review Template and include the completed Review Checklists. The format of the report can be modified as needed to meet the specific project needs. Complete all the sections and customize placeholder text included in the Review Template, and delete any instructions once completed.

The written review of the M&V submittal should be provided to the facility and/or customer staff; the customer should follow up on any questions or action-items identified through the review, including the ESCO as appropriate, and document any subsequent actions taken.

### 4.3 Reviewing M&V Plans

Evaluating M&V plans is an inexact science that requires technical expertise and experience. Ideally, the reviewer will have been involved in the project development phase and has an intimate understanding of the customer's goals, the agreed-upon allocation of project risks, site specific issues, as well as the objectives and constraints for each ECM.

The M&V Plan deserves careful evaluation as it defines the requirements for all future M&V activities. Discussion with the customer on the findings from the M&V plan review is usually warranted and may result in revisions to the M&V plan. Often, the review process is iterative. After an initial review, subsequent revisions of the M&V plan must be assessed to determine if adequate modifications have been made. Written evaluations of these subsequent M&V plans are needed to document follow-up actions taken.

#### 4.3.1 Prescriptive and Qualitative Evaluation Items

The first step in evaluating an M&V plan is to complete the M&V plan review checklists. The inclusion of all items on the checklists does not indicate the appropriateness of the M&V approach, only that the required information is included. Each measure requires extensive qualitative assessment, and tips for evaluating the M&V approach are included herein. All findings resulting from review, including completed checklists, should be included in the evaluation report, as discussed in Section 4.2.3.

Evaluate overall project-level items:

- Do all M&V strategies included in the plan support the concepts included in the Risk & Responsibility Matrix?
- Are contractual energy rates based on actual rates, including time-of-use provisions and peak demand ratchet charges? Are marginal (not blended) energy rates used?

- Are proposed escalation rates based on the latest Energy Escalation Rate Calculator (EERC)?
- Are M&V costs reasonable? Do costs align with planned activities?
- Is the level of savings predicted reasonable? Were project-level savings compared to overall site usage? (optional)
- Were all objectives and constraints of the project considered?

Assess each ECM:

- Review the agreed-upon Risk & Responsibility Matrix for the project. Ensure the M&V strategy for each measure conforms to the agreed-upon risk allocation.
- Note the source(s) of cost savings for each measure (electricity, demand, natural gas, water, O&M, etc.). Ensure M&V activities are adequate for all significant sources of savings.
- What is the likelihood of success for this measure? More rigorous M&V strategies are warranted for ECMs with substantial uncertainty or technical complexity.
- Is the level of savings predicted reasonable? Were ECM savings compared to system usage?
- Were key variables affecting energy use measured for each ECM (e.g., watts/fixture and hours/year)?
- Are assumptions or agreed-upon parameters reasonable, and do they include the source of data?
- Do the measurements include the parameters that are the source of the savings (e.g., reduction in watts/fixture or hours/year)?
- Are M&V costs reasonable? Do costs align with planned activities?

Consider the adequacy of baseline developed:

- Are all assumptions and stipulations reasonable, and do they include a justifying source?
- Were system performance characteristics recorded (e.g., lighting intensities, temperature set points)?
- Were energy calculations closely reviewed?
- Are savings estimates sound and reasonable?
- Were utility- or weather-based models validated?

Evaluate the quality of performance period activities:

- Are meaningful ongoing performance period data going to be used to calculate savings?

- What is being verified? Is this sufficient to support the savings guarantee?
- Will key variables affecting energy use be measured for this ECM? How often?
- Will single post-installation measurements apply to all years in the performance period? If so, how valuable are the data used?
- How likely are these data to change over the performance period? Is a detailed basis for future adjustments provided?
- Based on which party has accepted ongoing responsibility for each item, is this approach appropriate?

Review the strategy for conducting O&M for this ECM:

- Are O&M activities sufficiently detailed to demonstrate the level of effort?
- Are responsibilities allocated as suggested by the R&R Matrix?
- Are reporting requirements adequately defined?
- Is a strategy provided for verifying O&M savings and potential impacts of O&M deficiencies?

## 4.4 Reviewing Post-Installation and Annual Reports

Evaluating the post-installation and annual reports is more straightforward than reviewing M&V plans, as the level of qualitative assessment and engineering judgment required is considerably less. These reports document the results of the activities defined in the M&V plan.

The post-installation report documents the results of verification activities conducted by the ESCO after project installation but prior to project acceptance. This report documents any changes in the project scope and energy savings that may have occurred during construction and reports the expected first-year energy and cost savings. Keep in mind that in many applications of M&V Option A methods, measurements are only taken once following installation. Subsequent activities may be limited to inspections to verify an ECM's "potential to perform". The post-installation report is therefore a particularly critical document for ECMs using Option A M&V.

Similarly, each year during the performance period the contractor submits an annual report that documents the execution and results of the periodic M&V activities prescribed in the M&V plan (e.g., measurements, inspections, savings calculations, and O&M activities), as well as any items that may require additional follow-up. The annual report is the basis for determining if the annual savings guarantee has been met, and for determining if any "true-up" of payments is required in the case of any savings shortfall.

M&V for performance-based projects needs to show that the overall savings guarantee has been met but does not necessarily need to determine the actual savings for each ECM. The total level of cost savings for the project must meet or exceed the guaranteed cost savings for that performance year. If the project fails to meet the contractor's guaranteed annual savings as verified by the M&V documents, the customer adjusts the payment schedule, as necessary, to recover the customer's overpayments in the previous year and to reflect the lower performance level into the current year.

### 4.4.1 Prescriptive and Qualitative Evaluation Items

The first step in evaluating an M&V report is to complete the appropriate review checklists, as discussed in Section 4.2. A thorough evaluation of each measure, however, also requires some qualitative and engineering judgment as well, as discussed below. All findings resulting from review, including completed checklists, should be included in the evaluation report, as discussed in Section 4.2.3.

An evaluator should answer the following key questions:

- Were all activities required by the M&V plan followed?
- Was the content of the submitted report complete?
- Were the guaranteed savings for the project met? If the savings guarantee is not fulfilled for the performance year, how will the shortfall be remedied in future years?
- Does the agency's evaluation agree with the results of the ESCO's M&V report?

Understand any changes in the project's performance:

- Have savings levels for any ECMs increased or decreased significantly from year to year? Is an explanation of why changes occurred in these savings sufficient? If not, why not, and what corrective actions will or should be taken? By whom?
    - Note any changes in scope or performance, or results that differ from the post-installation or previous year's report
    - Note that ECMs using Option A methods may not show a change even if there are performance problems.
- Did the report provide useful feedback on the performance of each measure?
- Did the report verify the potential of the ECMs to save in future?
- Are there any performance problems, O&M issues, or deficiencies that need to be addressed? By whom?

Review the savings calculations:

- Were calculations submitted in electronic format?
- Are calculations transparent and understandable?

- Was the prescribed savings calculation methodology used? Did the reviewer verify the math in the savings calculations?
- Were rates shown in the final proposal used, and were rate adjustment factors applied correctly?
- Is the basis for any adjustment valid, and have the adjustments been consistently and uniformly applied?

# References

FEMP provides a wealth of guidance on all phases of ESPC projects on its web page at https://www.energy.gov/femp/resources-implementing-federal-energy-savings-performance-contracts.While developed for federal ESPCs, much of the guidance will be of use to any ESPC customer. Documents specifically related to M&V include the following:

***FEMP M&V Guidelines v. 4.0 (2015)***

https://www.energy.gov/sites/prod/files/2016/01/f28/mv_guide_4_0.pdf

***Example M&V Plan for an ESPC Project* (2007)**

Example Measurement & Verification Plan for a Super ESPC Project (energy.gov)

***How to Determine and Verify Operating and Maintenance (O&M) Savings in Federal Energy Savings Performance Contracts (2018)***

https://www.energy.gov/sites/prod/files/2018/03/f49/om_savings_guidance.pdf

***Reviewing Measurement & Verification Plans for Federal ESPC Projects (2007)***

www1.eere.energy.gov/femp/pdfs/6_3_reviewingmvplans.pdf

***Reviewing Post-Installation and Annual Reports for Federal ESPC Projects (2007)***

www1.eere.energy.gov/femp/docs/reviewing_annual_pi_reports.doc

***Including Retro-commissioning in Federal Energy Saving Performance Contracts (2003)***

https://www.energy.gov/femp/articles/including-retro-commissioning-federal-energy-savings-performance-contracts

***Commissioning Guidance for ESPCs (2015)***

http://energy.gov/sites/prod/files/2015/05/f22/ph3_cx_guide_23.pdf

***Guide to Government Witnessing and Review of Measurement and Verification Activities (2014)***

https://www.energy.gov/sites/default/files/2014/03/f14/agency_witness_accept.pdf

***Guidance on Utility Rate Estimations and Weather Normalization in Performance Contracts (2019)***

https://www.energy.gov/sites/prod/files/2019/02/f60/femp_escalation_guidance_2019.pdf

***Enhancing Performance Contracts with Monitoring-Based Commissioning (2021)***

https://www.energy.gov/sites/default/files/2021-07/enhancing-performance-contracts-mbcx.pdf

FEMP
Federal Energy Management Program

**For more information, visit:**
energy.gov/femp

DOE/GO-102023-6067 • September 2023