



SEM Energy Modeling Method Selection Guide

Northwest Strategic Energy Management Collaborative

Submitted by the NW SEM M&V Workgroup

6/14/2019

Contents

- Contents2
- Preface.....3
- Acknowledgements4
- Executive Summary5
- Method Selection6
- Method 1: Forecasting.....8
- Method 2: Backcasting10
- Method 3: Mean Model11
- Method 4: Pre-Post Model12
- Method 5: Standard Conditions14
- Method 6: Chaining16
- Method 7: Bottom-up Analysis on Individual SEM Measures or Calculated Savings Approach.....18
- Summary.....19
- Appendix A – Glossary20
- Appendix B - Metrics of Model Fitness.....21
- Appendix C - Analytical Options23
- Appendix D - Key References.....25

Preface

The Northwest Strategic Energy Management (NW SEM) Collaborative was formed in 2011 by the Bonneville Power Administration (BPA), Energy Trust of Oregon (Energy Trust), and the Northwest Energy Efficiency Alliance (NEEA) for the purpose of identifying and addressing market barriers to the adoption of Strategic Energy Management (SEM) in the industrial sector. Over the first three years of the collaborative, workgroups were formed to address topics such as solutions for small-to-medium industrials, market analysis and planning, and energy tracking and savings protocols.

The NW SEM Collaborative has since expanded to also include the commercial sector.

At the fall 2015 NW SEM workshop, a new workgroup was formed to build on the prior work of the Energy Tracking and Savings Protocols (ETSP) workgroup. Referred to as the Measurement and Verification (M&V) workgroup, this team sought to inventory and describe alternative options to regression-based forecasting models. This document is the first deliverable from that effort, and draws from existing protocols and the implementation experience of the team. The solutions and methods summarized within do not represent a comprehensive list, and the reader is encouraged to utilize the referenced sources for more detail.

Acknowledgements

This document represents the collective input of the following individuals who provided written descriptions of the various M&V methods documented within or who provided comments and suggestions on this document. Special thanks to Chad Gilless and Steve Martin who developed the concept for this paper and led the initial efforts, and to Sara York and Tina Schnell who led the drafting, editing, and finalization efforts.

- Andrew Bernath (Illume Advising)
- Andrew Wood (DNVGL)
- Chad Gilless (Stillwater Energy)
- Jacob Schroeder (Cascade Energy)
- Jennifer Hockett (Cadmus)
- Jim Stewart (Cadmus)
- Kati Harper (Energy Trust of Oregon)
- Lonny Peet (Nexant)
- Marissa Lee (Ecova)
- Phil Degens (Energy Trust of Oregon)
- Robert Greenwald (Prism Engineering)
- Sara York (Cascade Energy)
- Steve Martin (Cascade Energy)
- Tina Schnell (AESC)
- Todd Amundson (Bonneville Power Administration)
- Warren Fish (Northwest Energy Efficiency Alliance)

Executive Summary

Utilities, demand side energy efficiency programs, and corporations are adopting Strategic Energy Management (SEM) to drive low-cost operational, maintenance, and behavior-based energy efficiency improvements. In order to ensure the long-term success of initiatives and programs, measurement and verification (M&V) of savings is paramount. These programs demonstrate savings to justify costs and motivate continuous savings.

Traditional M&V methods use pre-SEM engagement (baseline) data to create an energy model which then forecasts energy use during an engagement or reporting period based on historical usage and external energy drivers. The observed or actual measured energy use throughout SEM is then compared to this forecast and the savings are calculated based on the difference. Although this method is generally preferred during SEM engagements, complexities often render the forecast-based savings calculations unsuitable and post-engagement. Evaluators may prefer to use different methods when all data is available at the time of analysis.

The purpose of this paper is to outline the selection and application of a variety of SEM M&V methods. First a method selection decision tree provides guidance to determine the appropriate M&V method under common circumstances. Detail on the application, merits, demerits, and special considerations for each of seven methods is then provided. The appendices include references, standard metrics for model fitness (statistical verification), and discussion of some common analytical choices.

The following is a brief summary of the seven methods included in this paper:

- Method 1 - The Forecasting Method: uses baseline data to create a predictive energy model. This is the preferred approach when baseline independent variable (energy driver) data is available. Assuming data is available, the challenge with this method is that if conditions significantly change during the reporting period, the model may not be able to predict accurately, in which case a revision or alternate method is needed.
- Method 2 - The Backcasting Method: follows the same methodology as forecasting but uses the reporting period to create the model. Energy use is then predicted for the baseline period. This method is useful when finer interval data is available during the reporting period (i.e. electrical meters are upgraded from monthly manual readouts to automated daily reads).
- Method 3 - A Mean Model: is applicable when there is insufficient variation in the independent variables or energy usage over time to create a predictive energy model. Use per period is then used as a model. The baseline is simply the average energy consumption for the facility during a defined baseline interval. This method should only be used when other modeling approaches are not effective and independent variables do not fluctuate much.
- Method 4 - The Pre-Post Method: uses both the baseline and reporting periods to generate one model. One of the model variables is an indicator variable for the reporting period, the coefficient of which is used to calculate energy savings. This method is useful when the baseline data set is too small to adequately capture the impact of key independent variables.
- Method 5 - The Standard Conditions Method: models are specified for the baseline and each reporting period. A set of “standard” conditions are specified and used as model inputs. Thus the baseline model energy use is adjusted for the standard conditions as is the energy use for the reporting period. An example of when this method is useful is when measures are highly weather dependent.

- Method 6 - The Chaining Method: is useful for situations where key independent variables, such as production, expand beyond the range observed in the baseline and reporting periods. In this case a single forecasting or backcasting model cannot adequately quantify savings. Instead a model is specified using an intermediate period. The model is then applied twice, once to backcast to the baseline period and once to forecast the reporting period performance.
- Method 7 - A Bottom-up Analysis: is essentially the calculation of energy savings through engineering principles and calculations for each of the identified and implemented measures. This method is useful when whole facility energy modeling provides inadequate analysis and when sufficient data is available for the implemented measures.

SEM stakeholders including program administrators, evaluators and implementers can benefit from a greater understanding of the methods available to calculate energy savings beyond the common forecasting method. In addition to the summaries presented in this paper, key references are included in Appendix D of this document to provide readers further background.

Method Selection

All seven methods outlined in this paper rely on similar statistical validation and share the goal of quantifying energy savings as the key output. However, the decision of when to use different methods can be a source of discussion between various SEM stakeholders. Generally, the selection is the modeler's preference; however there are three key criteria which delineate between methods:

1. *When is the model estimate being used?* Forecasting and mean model methods can be developed before a reporting period begins. Backcasting, pre-post, standard conditions, and chaining can only be developed after a reporting period to quantify historical savings.
2. *Are independent variables available that correlate with energy consumption?* If no statistically significant variables can be identified, the methods are limited to mean model or bottom-up analysis.
3. *Is total energy savings for the reporting period sufficient?* If so, pre-post and standard conditions methods are good options. However, if understanding when savings occur (e.g. seasonal opportunities) or finer interval savings are needed, then forecast, mean model, and backcast would be preferable.

The Method Decision Tree on the following page explores these criteria in more detail. Additional details regarding each modeling method are provided in the corresponding Method Section. The Appendix contains a glossary, metrics of model fitness, analytical options, and key references for further reading.

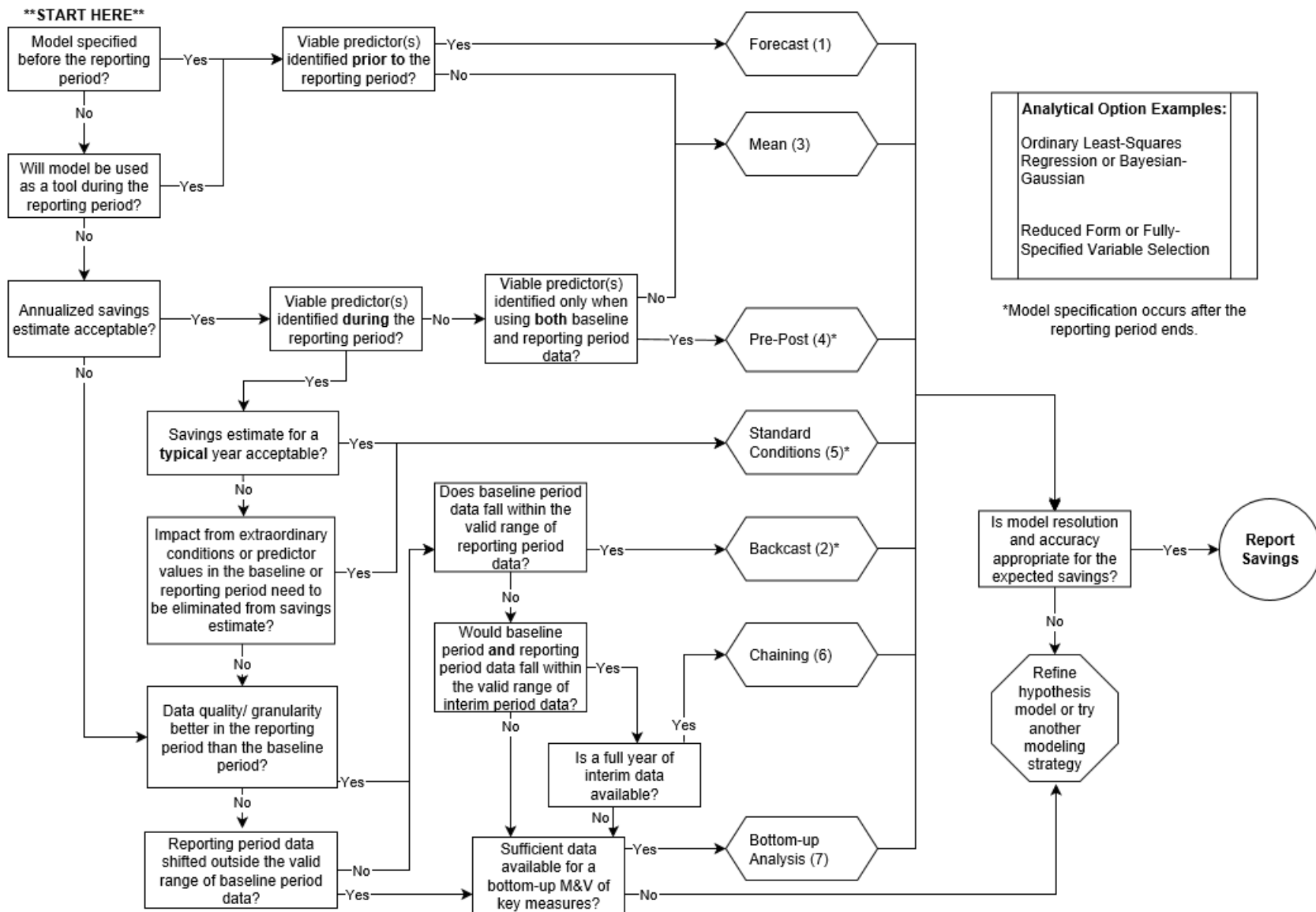


Figure 1: Decision Tree for Modeling Methodology Selection

Method 1: Forecasting

The forecasting approach involves creating a predictive energy model with baseline period historical data. The predictive energy model is then applied in the ensuing SEM reporting period to estimate what energy consumption would have been if the facility had not implemented SEM-related improvements. Forecasting adheres to IPMVP Whole Building Section C, and ASHRAE 14 Annex D requirements for modeling energy savings and provides timely reporting period progress visibility, which has made it the primary method used by implementers to estimate SEM savings.

Building the Model

Development of a model using the forecast method involves the following steps:

1. Define the ideal baseline period (generally full-year increments). Usually the baseline period should end as the reporting period begins, typically at the beginning of an SEM program.
2. Collect energy data via billing records or metering.
3. Identify energy consumption characteristics of facility and data sources for independent and dependent variables.
4. Collect independent variable data (production output, weather, occupancy, etc.).
5. Clean and prepare data.
6. Evaluate relevance of hypothesis independent variables using graphical or other methods (e.g. correlation matrix). If correlation is not identified, return to Step 4.
7. Develop a general model with the following structure:

$$E_{Model} = \beta_0 + \beta_1 \times (\text{production variable}) + \beta_2 \times (\text{weather variable}) + \beta_i \times (\text{other predictors})$$

8. Develop and review competing models using a stepwise (or other) selection process against the standard metrics of statistical fitness.
9. Assess whether new variables add to explanatory power of model and stop once best results are obtained.

Additional Model Diagnostics (i.e. statistical power to detect specific savings)

- Coefficient of Variation: Input to Fractional Savings Uncertainty value < 0.20 [ESI]
- Autocorrelation < 0.50 [ASHRAE 14]
- Fractional Savings Uncertainty < 50% at 68% Confidence Level [ASHRAE 14]
- Net Determination Bias < 0.005% [ASHRAE 14]
- Independent variable or energy driver data values < three standard deviations from the mean of the data that were used in the model fitting process [SEP 2017].

Estimating Savings

Apply the identified energy driver data to the reporting period (e.g. production and ambient condition data) to calculate the predicted energy use had the SEM engagement not occurred.

Energy savings are estimated by summing the differences between the actual energy and model-predicted energy during the reporting period. This calculation is sometimes referred to as the cumulative sum of the residuals (or CUSUM). If appropriate for SEM program requirements, apply adjustments for capital projects or non-routine adjustments (e.g. new load, plant expansion). Ideally these adjustments will be based on sub-metered data from new loads.

$$E_{Savings} = E_{Model} - E_{Reporting} \pm Adjustments$$

Usability

This method is the preferred approach to model SEM engagements throughout the reporting or engagement period when reliable independent variable data is available and demonstrates a statistically significant relationship with energy consumption during the baseline period. This methodology can be further enhanced by using shorter time intervals (daily and weekly in place of monthly) for a more detailed understanding of when energy use is affected.

Merits:

- The energy model can be developed before the implementation of SEM measures (i.e. ex-ante), and can be used to monitor, target and report on performance during the reporting period. For these reasons, forecasting tends to be the default method for SEM implementers.
- Different operating regimes (e.g. weekday versus weekend or high versus low production periods) can be specified to allow for a more robust model.
- This method can be further enhanced by applying offsets for seasonality and other changes in consumption in the baseline period.

Demerits:

- The valid range for independent variables is limited by baseline period conditions. If conditions change significantly during the reporting period, then alternative methods should be applied (i.e., backcasting, chaining, or bottom-up analysis).
- The number of data points may be less than desired, particularly in a monthly application. This limits the number of independent variables that can be included in the model and may limit the statistical power to detect SEM savings.

Forecasting Subset: Key Performance Indicators

Key performance indicator (KPI) based models are special cases of forecasting models. Their application is limited as KPI models are less robust than standard forecasting. While KPI models may provide an intuitive point of reference for on-site personnel, their application is more appropriate for monitoring rather than M&V. However, in some limited circumstances the following two methods may be appropriate.

- KPI Average or Ratio models specify a ratio with intercept equal to zero. Although simple in concept, a ratio model is only applicable when (1) baseload is negligible, (2) energy use has a strong linear relationship with a single independent variable, (3) energy use goes to zero as the variable goes to zero, and (4) the independent variable range and distribution is expected to remain consistent. In this case, a forecasting model would include one independent variable and an intercept of zero.

- KPI Bin models specify ratios for individual binning categories and a separate base energy load. Some examples of binning categories are paper grades, lumber specifications, and product codes. Base energy load is the energy consumed when production or the binning category is equal to zero. Binning categories may be appealing if there is no continuous numerical data that statistically represents how energy is consumed. Notable issues include model bias at high and low production rates, increased model maintenance requirements, lack of predictive capability outside the observed range, and difficulty separating production-dependent load from base energy load.

Method 2: Backcasting

The backcasting method uses data from the reporting period to specify a model and predict the baseline period energy consumption. To avoid seasonal bias, the reporting period should be one full year during which SEM participation was in progress. The backcasting approach is particularly useful when data in the reporting period is more granular than data in the baseline period (e.g., daily versus monthly). It may also be useful if the range of the independent variables in the reporting period is broader than the range observed in the baseline period. However, note that significant changes in energy savings during the reporting period may complicate backcasting model development. Care should be taken to ensure that any changes in savings are not artificially attributed to seasonal changes in the variable inputs.

Building the Model

Model development for backcasting generally follows the same guiding principles as forecasting. The main difference is the backcasting model cannot be specified until after the SEM reporting period ends. Please refer to the forecast method section for general model development and form.

Estimating Savings

Apply the identified energy driver data to the baseline period (e.g. production and ambient condition data) to calculate the predicted energy use had the SEM engagement started during the baseline period.

Energy savings are estimated by summing the differences between the actual energy and model-predicted energy during the baseline period.

$$E_{Savings} = E_{Baseline} - E_{Model} \pm Adjustments$$

Usability

Backcasting should be used when the reporting period energy data is more granular than energy data in the baseline period (e.g., daily vs. monthly). It may also be useful if the range of the independent variables in the reporting period is broader than the range observed in the baseline period.

Merits:

- If only monthly energy data is available in the baseline period, but daily energy data is available for the reporting period, a backcasting model takes advantage of the more granular data. The resulting model is more robust and may be able to identify independent variables that would have been missed with a monthly forecasting model.

- When the range of an independent variable is broader in the reporting period than it was in the baseline period, a backcasting model can accommodate the wider range seen in the reporting period (e.g. in the second month of the SEM engagement, an industrial facility with one manufacturing line began to ramp production up by 15% over the previous year's numbers. This results in higher production ranges than seen in the baseline; a backcasting model could expand the allowed energy driver range and ensure the high production days are within the model limits).
- Model development process is basically identical to that of a forecasting model.
- Model may be applied to future years as a forecasting model or, in the case of multi-year engagements, using the Chaining Method.

Demerits:

- Real-time calculation of energy savings is challenging.
- Limited use as a participant engagement tool. Since the model is based on the complete reporting period, it has limited use outside of post hoc evaluation.
- Reporting period savings are based on baseline period conditions rather than those of the reporting period. So, in a case where savings would scale with the independent variables, a large change in production may not translate to a corresponding change in savings.
- If energy savings build gradually through the reporting period, it may be difficult to develop a strong model. Reductions in energy use may be artificially attributed to coincident variations in the independent variables and not be identified as savings.

Method 3: Mean Model

The Mean Model approach may be necessary when there is insufficient variation in the independent variables or energy consumption over time to detect their effect on energy consumption (e.g., production or energy is constant), or, when there is insufficient correlation between hypothesized variables and energy consumption (e.g., production or proxies for production are not correlated with energy consumption). In some instances, this is indicative of a plant or process that is dominated by a high base energy load, i.e., a significant component of the plant or process does not vary considerably and cannot be explained by other energy driver data.

Building the Model

Primary steps:

1. Attempt the forecast modeling method as described in the forecasting method section.
2. Once potential predictors are shown to not meet standard statistical criteria, and/or energy use is shown to be fairly constant, the "mean-model" baseline is simply the average energy consumption for the facility during the defined baseline period interval.

This approach requires thorough documentation of baseline operating conditions so that changes in energy intensity observed during the reporting period can be properly assigned to the SEM engagement versus other changes in plant operation.

Additional Model Diagnostics

- Post-baseline, independent variables should remain within +/- three standard deviations from the mean [ESI], or more than 10% outside the range [ETO] of values recorded during the baseline, whichever is more reasonable.
- A best practice is to perform a statistical test on the baseline and reporting period data sets to confirm equivalent distributions (Chi-square test) and statistical significance of any reported change in the mean consumption (student T-test). If the T-test does not show a significant difference at the 80% confidence level [ESI], the program administrator or evaluator should consider whether savings should be reported for the SEM engagement.

Estimating Savings

Collect energy consumption data during the reporting period at the defined interval. A best practice is to continue collecting data for the independent variables that were not statistically significant, but by the physics of the process, may ultimately prove to be significant (e.g. during the SEM engagement as measures are implemented and the facility eliminates “waste” that is masking these effects). Savings is calculated in a mean model the same way it would be with a forecast methodology, provided there is a significant difference between the energy use in the baseline and reporting periods. A regression analysis may be performed again on the post-implementation data set to confirm whether statistical relationships have changed.

$$E_{Savings} = E_{Baseline_Mean} - E_{Reporting} \pm Adjustments$$

Usability

A mean model should only be used when none of the hypothesized variables are statistically significant in a regression analysis. Before selecting this option, efforts should be made to identify other variables, ensure that the data set is of high quality, and understand the physics of why a specific plant's or process' energy consumption would not correlated with the hypothesis independent variables as expected.

Merits:

- Simple methodology for the facility or an auditor to understand.
- No statistical software or detailed calculations are needed; therefore it is relatively easy to create.
- Provides performance transparency during the SEM engagement.

Demerits:

- Last-case scenario in terms of estimating energy use due to oversimplification of facility operations.
- Does not normalize energy consumption when changes occur, and so risk of over/underestimation during a reporting period is higher than with other methods.

Method 4: Pre-Post Model

At their most basic, pre-post models may be used to estimate the average energy savings of a single reporting period. More advanced versions may also be able to estimate the average relationship between key variables

and energy savings. Pre-post models may be useful when the baseline data set is too small to adequately capture the impact of key independent variables, since the baseline and reporting period are combined into a single regression with a variable for the reporting period in this methodology. Since a pre-post model cannot be used to inform SEM participants of estimated savings during the reporting period, it is less desirable than a predictive model during SEM implementation.

Building the Model

1. Identify the baseline period and reporting period.
2. Develop a model with the following general structure:

$$E_{Model} = \beta_0 + \beta_1 (\text{production variable}) + \beta_2 (\text{weather variable}) + \theta_0 (\text{reporting indicator})$$

where θ is the coefficient of the reporting indicator.

If the magnitude of savings is expected to be dependent upon variables inputs, it may also be appropriate to test interactive effects between selected variables and the reporting indicator, as shown below:

$$E_{Model} = \beta_0 + \beta_1 (\text{production variable}) + \beta_2 (\text{weather variable}) + \theta_0 (\text{reporting indicator}) + \theta_1 (\text{production} \times \text{reporting indicator}) + \theta_2 (\text{weather} \times \text{reporting indicator})$$

Further detail and helpful examples of the pre-post method are provided in referenced papers [Luneski] [Stewart].

Estimating Savings

1. Identify key independent variables and solve for model coefficients.
2. Conduct robustness and sensitivity checks on the regression model.
3. Calculate the reporting period savings. If there are no interaction terms, savings are equal to the coefficient of the reporting indicator variable, θ , times the number of valid time intervals in the reporting period.

$$E_{Savings} = \theta_0 \times (\# \text{ of Intervals in Reporting Period}) \pm \text{Adjustments}$$

Usability

A pre-post model may be appropriate when the baseline data set is too small to adequately capture the impact of key independent variables. In some cases, it may also be able to accommodate variables that fall outside the valid range of the baseline period or that did not exist in the baseline period to begin with. In these latter cases, care should be taken to ensure that the impact of the changes in independent variables is not conflated with the estimate of the reporting period savings.

Merits:

- Including the reporting period in the regression may allow the model to identify significant independent variables that appeared negligible in the baseline period.
- When model variables during the reporting period fall outside the valid range observed in the baseline period, a pre-post approach accommodates the full range observed in both the baseline and reporting period.

- For models with limited data, a pre-post approach can increase the size of the data set. This increases the degrees of freedom of the model and improves the ability of the model to include all key independent variables and estimate savings.
- This method eliminates the increasing uncertainty that results from forecasting serially correlated data beyond the baseline [Luneski].
- If the model includes terms to account for interaction between the reporting period and independent variables, a pre-post model can estimate the relationship between those variables and the energy savings.

Demerits:

- Limited use as a participant engagement tool. Since this model uses a dataset spanning both the baseline and reporting period, it can only be built at the end of the reporting period. This limits its use outside of post hoc evaluation.
- A pre-post model assumes a level savings impact across the reporting period. If the true savings varies throughout the reporting period (e.g., with action items being implemented gradually, or seasonal HVAC savings), model coefficients will only reflect an average impact across the whole reporting period.
- If any major changes to the model variables are coextensive with the reporting indicator, it is difficult to distinguish energy savings from the effect of the change in the variable inputs. Such cases cannot typically be resolved with regression analysis alone.
- Vulnerability to inaccurate, incomplete, or inconsistent data. Since θ must be calculated simultaneously with other coefficients, missing or erroneous data in either the baseline or testing period can impede the use of this method.

Method 5: Standard Conditions

The standard conditions method can be used to estimate energy savings under a standard set of “normal” conditions for each independent variable. Using observed data from the baseline and reporting period, the modeler specifies separate energy models for the baseline and reporting period [SEP 2017]. Savings are then estimated by applying a set of standard conditions data to each model, and then taking the difference of the energy consumption estimates from each model.

This method requires the designation of representative standard conditions data, which should be documented along with the model specifications. For example, an ambient temperature variable will require the selection of a suitable weather station with available typical meteorological year (TMY3) data, and a production variable will require a determination of a typical production year.

Standard conditions can help mitigate the risk of over or under-claiming energy saving in situations where energy driver range and frequency can vary significantly from year to year. For example, a destination hotel will seasonal tourists located in an area with extreme weather conditions could be a good candidate. Both weather and occupancy will likely be key energy drivers during model specification. If most savings opportunities occur in cooler weather when occupancy is low, energy savings would likely vary year-to-year.

A second example could be a water distribution system. Water demand will peak in hot, dry periods. If the system runs near full capacity at these times, the potential for energy savings will likely be lower than during cooler, wetter months. An extended drought could reduce the potential for energy savings and a very rainy year

could increase energy savings potential. If standard conditions are set, savings calculations would rely on one set of weather conditions, reducing annual variation.

Building the Models and Defining Standard Conditions

The process of specifying the baseline period model and the reporting period model is fundamentally similar to the methodology described in the forecast section. Here two models must be specified, and the independent variables included in each model may vary. However, major differences between the two models should align with known changes in energy consumption characteristics. In most cases models will be developed from the same number of observations (e.g. 365 days), as exceptions may lead to overrepresentation of certain operating modes and seasonal bias.

A standard conditions data set must be defined for each independent variable. Chronologically, this should happen early in the process when the baseline model is specified. However, the primary function of the standard conditions data set is for producing the savings estimate, which occurs at the conclusion of the reporting period. Common methods of defining standard conditions for typical categories of independent variables are outlined below.

- Weather: TMY3 data from a nearby weather station is recommended. Program administrators may specify or allow other data sources to define standard weather conditions.
- Production: Multiple years of production data may be used to represent a typical production year. If representative historical data is not available (e.g. new construction), other sources may be proposed such as approved production plans validated against more recent observations.
- Categorical variables: Designating operating modes such as weekend/weekday or non-production (e.g.) may be self-evident or need to be established with input from plant staff.

Estimating Savings

To estimate energy savings the standard conditions data set is applied to both the baseline and reporting period models. The difference between the two predicted consumption values is the energy savings estimate. Key steps include:

1. Estimate baseline energy consumption at standard conditions by inputting standard conditions data into the baseline model.
2. Estimate reporting period energy consumption at standard conditions by inputting standard conditions data into the reporting period model.
3. If applicable, calculate non-routine adjustments.
4. Estimate energy savings at standard conditions using the following equation:

$$E_{Savings} = E_{Baseline_Model} - E_{Reporting_Model} \pm Adjustments$$

5. CV(RMSE) < 25%, [SCE] [ASHRAE 14].

Usability

The standard conditions method can be used to estimate energy savings for an average year versus the actual year. Standard conditions could relate to baseline data (the backcasting method), or the reporting period data (the forecasting method), a different period entirely, or some average/typical profile, in which case the data may come from a wide range of time periods (as with TMY3) [SCE].

Merits:

- Applying standard conditions to the baseline and reporting period models reduces variation in saving estimates caused by production or weather-dependent mechanism of savings. For example, a facility that installed a new rooftop condenser designed to save energy during cool ambient conditions may show less savings during an abnormally hot summer and more savings during a cool summer. Applying standard conditions temperature data would mitigate annual variation.
- The standard conditions method provides the SEM participant with a model that is reflective of current operating characteristics (versus a forecasting model, which may be somewhat obsolete). The reporting period model can then be used to provide a predictive forecast in the subsequent year.

Demerits:

- Requires creating multiple models, one for the baseline and one for the reporting period. In concept, the reporting period model would need to be updated as often as the desired monitoring frequency.
- Point estimates of savings (i.e. energy impact of making a change at a given time) include an additional source of variation. Specifically, the difference between the standard condition value and actual value of each independent variable for a given observation. For this reason, standard conditions models are generally limited to total annual estimates of energy savings.

Method 6: Chaining

The chaining method uses an intermediate period between the baseline and reporting period to develop and specify an energy model. The model is then applied twice: first to backcast the baseline and then to forecast the reporting period. This method can be useful when neither forecasting nor backcasting is appropriate as a standalone method. This method is most applicable to multi-year SEM engagements.

For a multi-year SEM program, chaining can extend the useful life of a backcasting model, which improves program cost-effectiveness. For example, a two year program that would normally use the same forecasting model for both reporting years has a participant upgrade the electrical metering directly before the Year One reporting period. The new electrical data allows the program to upgrade a monthly forecasting model to daily backcasting model, which is used to estimate Year One savings. To estimate Year Two savings, the program could either generate a new backcasting model based on Year Two reporting period data or the program could leverage the specified backcasting Year One model via the Chaining Method. In this scenario, cumulative energy savings would be the sum of backcast savings (equal to Year One energy savings) and forecast savings during Year Two's reporting period.

This method is also relevant when independent variables change over time, and there is an intermediate period between the baseline and reporting period under consideration which would more accurately represent the energy driver range.

Chaining may be appropriate under certain conditions for a single year engagement. However, more straight-forward methods such as model adjustments and pre/post analysis are likely more appropriate solutions.

Building the Model

Model development follows the same guiding principles as the forecast and backcast methods. Please refer to those sections for general model development. Additional considerations for the chaining method include the following:

1. Identify and define the three periods (baseline, intermediate, and reporting) in full-year allotments if possible to minimize seasonal bias.
 - The intermediate period must bridge or chain the energy driver ranges for both the baseline and reporting period datasets. The intermediate period may overlap with other SEM implementation (e.g., Year One reporting period in the summary example).
 - If the system can be accurately modeled using less than one year of data, the rationale must be clearly documented and sufficient observations must be included to justify the number of independent variables included.
2. Preference should be given to model forms previously used to claim savings (i.e. if a forecasting model was originally generated, the same independent variables should be considered for the chaining model).
3. If a backcasting model has been generated for previous reporting, it may be applicable to future year savings using the chaining method.

Estimating Savings

The model developed using the intermediate period data is applied to both the baseline period (backcast) and to the reporting period (forecast). Energy savings from both periods are summed to determine the total energy savings to allocate to the reporting period [SEP 2017].

Cumulative energy savings are the sum of the savings from the backcast and forecast method applications. Please refer to the backcasting and forecasting sections of this paper for more information.

$$E_{Savings} = E_{Savings_Backcast} + E_{Savings_Forecast} \pm Adjustments$$

Usability

The main application for this method is to accommodate for changes in operational conditions that are substantially different in the baseline and reporting period, or conditions that exhibit unexpected (generally non-linear) trends in energy driver rates that are observed in one period but not the other.

The chaining approach takes advantage of the two most frequently used modeling techniques. The explanation is more complex than other methods, which may be challenging for implementers and SEM participants to understand. SEM implementers should be prepared to spend additional time clarifying how savings are calculated if using this method.

This method may also be applicable if a backcasting model is used to quantify energy savings as the same model may be applied to forecast future savings without re-baselining or specifying a new model, as may occur in some multi-year SEM programs.

Merits:

- Establishes a method of extending independent variable ranges when reporting period ranges fall outside the permitted range established during the baseline period.
- Calculations are a combination of two established and documented methods (forecast and backcast).
- May also be used to account for non-linear energy driver relationships or change points that were not observed in the original baseline data due to insufficient data ranges.
- Inclusive of energy savings generated between the baseline period and the intermediate period.

Demerits:

- Qualitative explanation of how energy savings are calculated is more complex than either the forecast or backcast method, which may increase SEM implementer time commitment to explanations and responding to questions from SEM participants, program administrators, and data review auditors.
- It is best if the intermediate period used to develop the energy model is one full year, which could delay program results.
- Real-time calculation of energy savings is challenging.
- Increases the SEM modeller time input as two data sets must be reviewed.

Method 7: Bottom-up Analysis on Individual SEM Measures or Calculated Savings Approach

All of the previous methods described thus far specify top-down, whole-facility or whole-system energy models. Bottom-up analysis of measure-by-measure energy savings may be the only alternative to estimating savings during SEM engagements when top-down, whole-facility energy modeling efforts prove unsuccessful. An example of this is when a baseline model is no longer valid due to changes in production or process, or reliability of data sources diminishes. Engineering principles and calculations can be used in conjunction with other methods to provide the bottom up confirmation of energy savings in this scenario.

Model Alternative Process

1. Obtain list of completed SEM opportunities.
2. Determine which opportunities are associated with direct energy saving outcomes (e.g. an energy team meeting does not save energy but reducing the compressed air system pressure does).
3. Confirm with SEM participant that opportunities have been completed, that they are still in effect, and the date of implementation.
4. If possible, perform facility visits to verify opportunities (all or a subset) have been completed.
5. Calculate energy savings based on engineering principles and calculations.
6. The savings of each opportunity in operation (or for the period it was in place) are summed together to obtain a bottom-up estimate.
7. The savings can be adjusted if the facility visit finds that more or less of the opportunities are still in place relative to what was confirmed by the interview.

Estimating Savings

Engineering equations are used to estimate savings from the energy saving actions that the participant reports having implemented due to SEM.

Examples of simple equations are:

$$E_{Savings} = (Power_{Baseline} - Power_{Reporting}) \times (Hours_of_Operation)$$

$$E_{Savings} = (Power) \times (Hours_of_Operation_{Baseline} - Hours_of_Operation_{Reporting})$$

Usability

If independent variable tracking is not consistently maintained, or if changes within the process render the original model invalid, the system may have too many interacting variables for successful application of the previous methods. The bottom-up approach may be the only other alternative to estimating savings. Depending on magnitude of energy savings and the durability of the opportunity, data collection (including short-term data logging) may be appropriate to support the energy engineering calculations.

Merits:

- It provides a greater indication to the source of savings and provides feedback to implementer on what type of actions are being taken and maintained.
- Energy savings calculations are not impacted by non-programmatic events that influence energy consumption but cannot be isolated via a top-down approach.

Demerits:

- Cost-effective application of this method requires careful selection of which measures to calculate which may result in conservative energy savings estimates.
- Can miss the holistic impact of behavioral change, interactive impacts, and small or complex measures.

Data monitoring requirements need to be scaled with the magnitude of savings and programmatic risk when claiming these savings

Summary

It is the authors' hope that this document provides insights into the variety of modeling methods available to SEM stakeholders, including program administrators, evaluators and implementers, along with guidance for when each method is best applied. Additional information on the standard regression model best practices and statistical criteria are readily available in several published documents listed as key references in Appendix D.

Appendix A – Glossary

American Society of Heating, Refrigeration and Air-Conditioning Engineers (ASHRAE): A society that sets standards for energy efficiency, building systems, indoor air quality, refrigeration, and sustainability within the built environment.

Base energy load: Constant energy consumption incurred by equipment that runs regardless of production or space use. Examples include lights, some fans, circulation pumps, and other uncontrolled equipment.

Baseline period: The period before the engagement with the participant, used to compare to the energy consumption during the reporting period.

CUSUM: Cumulative sum of residuals. It is calculated as the sum of the differences, over time, between actual energy use versus predicted energy use during the reporting period.

Energy driver: An independent variable, which has a direct relationship to energy consumption. Also referred to as a model parameter or energy predictor.

Energy model: A mathematical model where the dependent variable (energy) is regressed on the independent variables, which are said to determine its value. The energy model is the primary method for claiming energy savings for an SEM engagement.

Engagement period: The period that a participant is enrolled in the SEM program, where SEM projects are being implemented. Also referred to as the implementation or performance period.

Indicator variables: An artificial variable created to account for two or more distinct categories/levels. Generally, indicator variables are assigned a value of 0 or 1 to account for different modes of operations. Indicator variables may be used to represent seasonal changes, energy projects during the baseline period, or other step-changes.

International Performance Measurement and Verification Protocol (IPMVP): Provides best practices available for verifying the results of energy efficiency, water efficiency, and renewable energy projects in commercial and industrial facilities.

Intermediate period: The period between the baseline and reporting period used to specify a model when implementing the chaining method.

Measurement and verification (M&V): Method of reliably determining energy savings. Savings are determined by comparing before and after data in a standardized, methodical way.

Operations and maintenance (O&M): In the context of SEM, many energy saving opportunities are tied to operational or maintenance changes which tend to be implementable at low- to no-cost.

Participant: The recipient of the SEM program. Also referred to as the customer, facility or site.

Reporting period: Energy consumption during this period is compared to energy use during the baseline period to calculate/quantify energy savings.

Strategic Energy Management (SEM): A holistic approach to managing energy use in order to continuously improve energy performance and achieve persistent energy and cost savings over the long term [CEE].

Appendix B - Metrics of Model Fitness

Standard Metrics of Model Fitness

This section explains the most common metrics used for assessing model fitness, their ranges, and reference sources. Additional criteria may be listed under a particular model method.

- **Coefficient of Determination (R-square):** Range of possible values from 0 to 1. Mathematically defined as the ratio of explained variation of dependent variable to the total variation of dependent variable. Hence, R-square is often thought of as a goodness-of-fit test. Note that R-square > 0.75 is generally considered a rule-of-thumb for an acceptable model using monthly billing data, and whether or not higher frequency data is used, a high R-square value is not enough to say the selected model fits the data well, nor that a low R-square indicates a poor model.
 - R-square > 0.50 [SEP 2012] (with additional criteria requirements)
 - R-square > 0.75 [IPMVP]
- **Adjusted R-square:** Range of possible values from negative (usually not) to 1. Adjusted R-square adjusts the R-square value to account for the number of explanatory terms in a model relative to the number of data points and includes a penalty for each additional variable (larger penalty relative to small number of data points, lesser penalty with large number of data points). Since the addition of an independent variable will always result in an increase in a model's R-square, an adjusted R-square value provides insight to whether additional independent variable(s) improve the model fit from one model specification to another by more than random chance.
- **P-value of coefficients (P-value):** P-value can be considered the probability of a non-relationship between a dependent variable and the independent variable. Therefore a small P-value indicates the independent variable is a significant (important) predictor of the model's dependent variable.
 - At least one independent variable must have a P-value < 0.10 [SEP 2012]
 - Each independent variable must have P-value < 0.20 [SEP 2012] or < 0.05 [ETO], whichever is more reasonable
- **T-Statistic of coefficients (T-stat):** T-stat is calculated from the ratio of an independent variable's coefficient multiplier in model equation to its standard error, and hence a measure of the significance for each coefficient (therefore, each independent variable) in the model. Per IPMVP Volume 1 (2012), a rule of thumb is that the absolute value of a t-stat 2.0 or greater implies that the estimated coefficient is significant relative to its standard error, and therefore a relationship does exist between the dependent variable and independent variable (related to the coefficient).
 - Include variables when T-Stat > 2.0 [IPMVP], [ETO]
- **P-value of model based on F-test:** For a simple regression (no change points) with a single independent variable, the F-test value is the same as the p-value for the independent variable, in that it is the probability that the model does not explain most of the variation in the dependent variable. Low p-values of the model based on the F-test are desirable. [In the Excel Regression tool output, *Significance F* is the whole-model equivalent of *p*-value for an individual variable.]
 - The model p-value, based on the F-test, < 0.10 [SEP 2012] (with additional criterial requirements)

Additional Metrics of Model Fitness

- **Net Determination Bias (NDB):** range of possible values from negative to positive, but typically discussed in terms of absolute value. NDB is defined as the ratio of summation of differences between model-predicted and actual dependent variable values to summation of actual dependent variable values. Per ASHRAE Guideline 14, a model is unacceptable (i.e., should be rejected) when its NDB is greater than 0.005%.
 - $NDB < 0.005\%$ [ASHRAE 14]
- **Coefficient of Variation (CV):** Coefficient of variation of the root mean squared error (RMSE) is the RMSE normalized by the average dependent variable value. It is used as a measure of the uncertainty of the model, where smaller CV values indicate lower uncertainty and can be more informative than R-square for situations where the dependence is relatively low.
 - $CV < 0.20$ [ESI]
- **Autocorrelation:** The residuals from a regression model are said to be independently distributed if the residual at time t is uncorrelated with the residual at time $t-1$, $t-2$, or any other period. If residuals are found to be correlated with one another, they are said to suffer from autocorrelation or serial correlation. Autocorrelation can be common in energy models, especially with data taken at short intervals.
 - Autocorrelation < 0.50 can be ignored [ASHRAE 14]
- **Multicollinearity:** this term describes a strong relationship between two independent variables. A key point is that allowing multicollinearity in a model can create a number of problems and lead to incorrect inferences from the model, such as affecting the precision of individual coefficients and understating the statistical significance of individual predictor variables.
 - Multicollinearity, as in R-square between two independent variables > 0.70 indicates the need to address [ESI]
- **Fractional Savings Uncertainty (FSU):** The relative precision of an energy savings estimate as the magnitude of the uncertainty relative to the estimate of energy savings. The key drivers of relative precision are: (1) size of the signal, a function of the model's ability to explain variation by observed usage; (2) amount of noise in the data, whereas savings from SEM engagements in facilities with noisy, or erratic, load patterns will be more uncertain than SEM engagements in facilities with more predictable load patterns; and (3) frequency of the data, in it should be easier to precisely measure effects with daily or hourly energy usage data than compared to monthly data. Yet, practitioners should consider the effects of autocorrelation on traditional statistical calculations when using hourly or sub-hourly data and take additional modeling steps to produce unbiased estimates of uncertainty than is necessary with lower frequency data.
 - $FSU < 50\%$ at 68% confidence level [ASHRAE 14]
- **Data collection range:** a model's independent variable data set should cover the full range of system or facility operating range conditions. One consideration for "outlier data" is applying a maximum, ± 3 sigma, yet due to the potential for increased model error at values other than the mean, professional judgment should be applied to the decision of valid range.
- **Independent variable data values:** the data values of independent variable should be less than three standard deviations from the mean of the data that were used in the model fitting process per SEP M&V Protocol.

Appendix C - Analytical Options

In addition to model method selection, there are analytical options which may be applied to most or all methods. Each analytical option produces valid results when properly applied. There are many analytical options, but the two most discussed by the authors of this paper are estimation methods for model creation and variable form.

Model Creation: Linear Regression Estimation Methods

The most common estimation method is ordinary least squares (OLS). OLS is readily accessible in Excel and other common evaluation software. OLS produces accurate results provided that model residuals have a mean of zero, variance is constant and there is limited autocorrelation. If errors follow a normal distribution, hypothesis testing can be used to determine the statistical significance of independent variables and confidence intervals for the models.

Other potential methods are gaining traction as machine learning and fine interval data are more readily available. For example, the aim of Bayesian Linear Regression is not to find the single “best” value of the model parameters or predictions, but to determine the posterior distribution for them. The two primary benefits of Bayesian Linear Regression correspond to the prior distribution and posterior distribution:

- Priors: With domain knowledge, or a guess for what the model parameters should be, that knowledge can be included in the model, unlike in the frequentist OLS approach which assumes everything there is to know about the parameters comes from the data. If there are no estimates ahead of time, non-informative priors can be used for the parameters such as a normal distribution.
- Posterior: The result of performing Bayesian Linear Regression is a distribution of possible model parameters and predictions based on the data and the prior.

As the amount of data points increases, the likelihood washes out the prior, and in the case of infinite data, the outputs for the parameters converge to the values obtained from OLS (Toward Data Science). Convergence diagnostics for the Markov chains, e.g., trace plots, Gelman and Rubin diagnostic—potential scale reduction factors, autocorrelation, effective sample size, and residuals (Coursera) Implementation requires statistical software (e.g. WinBugs, SAS, R). MCMC can, but does not necessarily, require significant computing resources and knowledge of specialized software.

Bayesian Merits

- Takes as input the standard baseline/training data that we would collect from an OLS model report, workbook, or similar.
- Bayesian statistics uses “credible intervals” rather than confidence intervals. In some cases, the credible intervals for predictions are narrower than the confidence intervals resulting from OLS regression (Educational Research). In theoretical work, credible intervals are not often calculated for the prediction of future events, but for inference of parameters – i.e., credible intervals of a parameter, not for the outcomes of the variable itself. However, particularly where applications are concerned with possible extreme values of yet to be observed cases, credible intervals for such values can be of practical importance.

Bayesian Demerits:

- Requires an initial “guess” for parameter estimates but can use uninformative prior if necessary.

- Outputs the distribution of model parameters, i.e., mean and uncertainty of each regression coefficient.
- May not be the best tool for implementers due to being tricky to implement (not currently available in Excel) and difficult for stakeholders to follow the process.
- In cases with high volumes of data or model parameters, could require considerable computing power.
- Proof of concept has not been performed or validated in SEM M&V applications for this method to date. It is an area for further research that should be explored in tandem with implementing other methods.

Independent Variable Selection: Reduced Form or Fully-Specified

The decision of how to select independent variables depends on many factors including available data, consistency across project portfolio, statistical requirements, engineering judgement, cost-effectiveness, and accuracy. Two methods are discussed here: reduced form and fully specified. Reduced form variable selection focuses on minimizing the number of variables evaluated and choosing only those with straight-forward relationships [Degens]. A fully-specified form will generally include a thorough review of potential independent variables and their statistical relevance with a goal of selecting the most statically sound that is also reasonable from an engineering cause/effect standpoint [Bernath].

Both forms emphasize reasonable statistics and good theoretical foundations in variable selection. The following table outlines merits of each approach.

Component	Reduced Form	Fully Specified
Model Buy-In	Readily interpreted by most parties	Addresses “what if” and “what about” questions
Number of Variables	One or two, facilitating tracking	One or more as needed to describe system
Statistical Fitness	Reasonable on a portfolio basis with well-reasoned variable selection. Some individual models may have lower statistics. Unlikely to over-specify model	High individual model statistical fitness or well-investigated and explained model if statistics are low
Data Interval	Finer interval data improves statistics. At long intervals, such as monthly, optimizes the time input to generate models with generally lower statistical fitness	Pairs selection of optimal interval and variables when fine-interval data is available
Accurate Savings	Accurate for facilities with few significant independent variables	Accurate for standard and unique facilities
Cost-effectiveness	Streamlines model development	Longer initial model development process with possible off-set of high confidence in predictive/backcasting accuracy

Appendix D - Key References

- [ASHRAE 14] The American Society of Heating, Refrigerating and Air-Conditioning Engineers Guideline 14-2014 – Measurement of Energy, Demand, and Water Savings. Atlanta, Ga.: American Society of Heating, Refrigerating and Air-Conditioning Engineers. Purchase at: https://www.techstreet.com/standards/guideline-14-2014-measurement-of-energy-demand-and-water-savings?product_id=1888937.
- [Bernath] Bernath, Andrew and Buffum, Maggie. Estimating Energy Savings Resulting from Strategic Energy Management Programs: Methodology Comparison. 2017 International Energy Program Evaluation Conference, Baltimore, MD.
- [BPA] Bonneville Power Administration M&V Protocols: Regression Reference Guide, V2.0. Available at: <https://www.bpa.gov/EE/Policy/IManual/Pages/IM-Document-Library.aspx>.
- [CEE] Consortium for Energy Efficiency. CEESM Strategic Energy Management Minimum Elements. 2014. Available at https://library.cee1.org/system/files/library/11283/SEM_Minimum_Elements.pdf.
- [Degens] Degens, Phil and Kelly, Anna. Strategic Energy Modeling: What's Good Enough? 2017 International Energy Program Evaluation Conference. Baltimore, MD.
- [Educational] Linear Regression vs Bayesian Regression. Educational Research Techniques. Available at: <https://educationalresearchtechniques.com/2017/10/18/linear-regression-vs-bayesian-regression/>.
- [ESI] Energy Smart Industrial Energy Performance Tracking (EPT) Team. 2018. *Monitoring, Targeting and Reporting (MT&R) Reference Guide – Revision 7.0*. <http://www.bpa.gov/EE/Policy/IManual/Documents/MTR-Reference-Guide-Rev7.pdf>. Portland, Ore: Bonneville Power Administration.
- [ETO] Production Efficiency Team. 2019. *Energy Intensity Modeling Guideline – V2.2*. <https://conduitnw.org/Handlers/conduit/FileHandler.ashx?RID=4708>. Portland, Ore: Energy Trust of Oregon.
- [Gelman] Gelman, Andrew, Goodrich, B., Gabry, J., Ali. I. 2017. R-squared for Bayesian regression models. Unpublished, available at: http://www.stat.columbia.edu/~gelman/research/unpublished/bayes_R2.pdf.
- [Heiner] Heiner, Matthew. Bayesian Statistics: Techniques and Models. Coursera. Available at: <https://www.coursera.org/lecture/mcmc-bayesian-statistics>.
- [IPMVP] International Performance Measurement and Verification Protocol Volume I: Concepts and Options for Determining Energy and Water Savings. EVO 10000 – 1:2012. Washington, D.C.: Efficiency Valuation Organization. Available at: <https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp>.
- [Koehrsen] Koehrsen, William. 2018. *Introduction to Bayesian Linear Regression*. Toward Data Science. Available at: <https://towardsdatascience.com/introduction-to-bayesian-linear-regression-e66e60791ea7>.
- [Luneski] Luneski, R.D. 2011. A Generalized Method for Estimation of Industrial Energy Savings from Capital and Behavior Programs. Industrial Energy Analysis 2011.

[SCE] Emerging Products, Customer Service, Southern California Edison. 2017. *Normalized Metered Energy Consumption Savings Procedures Manual*. Available at: <https://www.etcc-ca.com/reports/normalized-metered-energy-consumption-savings-procedures-manual?dl=1545443978>.

[SEP 2012] Superior Energy Performance, 2012, Superior Energy Performance Measurement and Verification Protocol for Industry. <https://betterbuildingssolutioncenter.energy.gov/iso-50001/sep-50001/sep-2012>. Oakland, Calif.: The Regents of the University of California.

[SEP 2017] Superior Energy Performance, 2017, Superior Energy Performance Measurement and Verification Protocol. <https://betterbuildingssolutioncenter.energy.gov/iso-50001/sep-50001/sep-2017#program>. Oakland, Calif.: The Regents of the University of California.

[Stewart] Stewart, James. 2017. Chapter 24: Strategic Energy Management (SEM) Evaluation Protocol. Available at: <https://www.nrel.gov/docs/fy17osti/68316.pdf>.