EVERGREEN ECONOMICS

Bonneville Power Administration

Strategic Energy Management Persistence Evaluation – Final Report

Submitted by Evergreen Economics

In partnership with SBW Consulting

November 30, 2022

1 EXECUTIVE SUMMARY

This document presents the results of an evaluation of the lifetime of measures associated with Bonneville Power Administration's (BPA's) Strategic Energy Management (SEM) offering for industrial facilities. A team led by Evergreen Economics (which includes SBW Consulting) conducted this research.

The objective of this evaluation was to estimate the effective useful life (EUL) and remaining useful life (RUL) for measures associated with BPA's SEM program. This concerns only the lifetime of reported measures/activities, not the savings that they yield.

1.1 METHODOLOGY

This study represents the population of utility industrial customers that were active in the program between 2015 and 2017 and receiving SEM consulting. The sample design targeted a 90/10 confidence level and precision and was developed based on an SEM database the evaluators developed based on SEM project files collected from BPA.

The sample frame contained 44 distinct sites. We used simple random sampling to select 15 sites for the SEM persistence study.

The data collection approach included a combination of sources such as interviews with SEM program implementation staff, interviews with onsite production management staff, and data collected from onsite systems. Ultimately, 108 SEM measures were included in the analysis. The SEM survival analysis was conducted using a multistep process including classifying SEM measures, estimating effective useful lifetimes, and estimating survival curves.

1.2 SUMMARY OF FINDINGS

This study was to apply EUL analysis techniques to inform estimates of measure[1] life for BPA's SEM program. There are two important caveats to consider when interpreting and applying these results to BPA's and other SEM programs.

1. Our study was designed to inform measure life (whether the measure is still active or not), which is distinct from measure savings persistence (i.e., whether the measure is still saving in future years what it was intended to in year one). Savings is an additional parameter to consider that may change over time and would need to be applied along with estimated measure life to get a more complete understanding of measure savings persistence.
2. The underlying information we used to develop a database from which to draw the study sample is based on SEM project completion reports after a year of

[1] BPA typically identifies SEM as a "measure", which includes all of its constituent activities. The purpose of this study was to assess the life of all of the activities that were likely to directly save energy, and not activities that are purely informational or organizational. The study team chose to define the smallest sampled unit in this study as a "measure" to be consistent with the other commonly sampled subunits in energy efficiency program impact evaluation.

participation and may underrepresent measure failures (i.e., measures tried by facilities but stopped immediately during the engagement year). Our measure life estimate may be biased upwards as a result of the database assembled from SEM program completion reports, which were the only uniformly available source of implemented measures during the studied period. The reports contained measures that the program was aware of, including some action items that were only attempted, although some attempted measures may not be represented.

Note that this concern only relates to the lifetime of reported measures/ activities, not the savings that they yield.

The study's key findings are:

Key Finding 1 - Overall EUL: The estimated EUL for BPA's SEM measures is 8.5 years, with 95 percent confidence bounds of 3.7 to 13.3 years. Of 108 measures assessed in the study (where program engagement was initiated between 2015 and 2017), only 13 were no longer in operation. On average, measures were removed within 1.7 years.

Key Finding 2 - Variation in EUL: The estimated EUL does not differ significantly based on type of SEM measure, equipment type, or industry type.

We recommend that BPA and its SEM program comprehensively track the specific SEM project and individual activity details listed in Appendix B to support robust SEM savings and EUL analysis in the future. There should be a follow up study to address the savings persistence for SEM.

2 INTRODUCTION

Bonneville Power Administration (BPA) contracted with Evergreen Economics (along with our subcontractor SBW Consulting) to conduct an assessment of the lifetime of measures associated with Bonneville Power Administration's (BPA's) Strategic Energy Management (SEM) offering for industrial facilities.

2.1 BACKGROUND

BPA began offering its SEM program to industrial facilities in 2010. Through the program, BPA provides long-term energy management consulting services to educate and train industrial energy users to (1) develop and execute a long-term energy-planning strategy and (2) permanently integrate energy management into their business planning. BPA's SEM program was one of the nation's first large-scale deployments of an SEM program in the industrial sector, having engaged 74 projects by the end of 2018.

An in-depth evaluation of the SEM program was completed in 2016. However, the issue of persistence was not addressed by that evaluation beyond the sampled participation period. In this research, persistence refers to the lifetime of a specific measure listed in the SEM plan, defined as how long the measure continues to be in operation.

2.2 DEFINITION OF AN SEM MEASURE

BPA typically identifies SEM as a "measure", which includes all of its constituent activities. The purpose of this study was to assess the life of all of the activities that were likely to directly save energy, and not activities that are purely informational or organizational. The study team chose to define the smallest sampled unit in this study as a "measure" to be consistent with the other commonly sampled subunits in energy efficiency program impact evaluation.

2.3 STUDY OBJECTIVES

The objective of this assessment was to estimate the effective useful life (EUL) and remaining useful life (RUL) for measures associated with BPA's SEM program. The outputs of this assessment are estimates of EUL, and whether or not the SEM measures persist after participants leave the SEM program.

3 METHODOLOGY

This section summarizes the methods used to conduct this assessment, and is organized as follows: Sample Design, Database Preparation, Data Collection, and Study Analysis.

3.1 SAMPLE DESIGN

BPA's evaluation and reporting policies include an established target for impact evaluation, striving for evaluations that attain a relative error of 10 percent at the 90 percent confidence level, with a minimum acceptable level of 80/20.

To evaluate the persistence of Strategic Energy Management (SEM)-initiated measures, we randomly sampled 15 Energy Management (EM) Program participants from those that were active in the program between 2015 and 2017.[2] We excluded measures that were conducted under the Track and Tune program, because this option shifted more SEM measures to the implementer, which is less representative of the current design of SEM.[3] The primary sampling unit of this study is a site (i.e., a single business entity that has completed one or more SEM measures between 2015 and 2017, as documented in the completion reports).

The sample frame contained 44 distinct sites, from which we drew our sample of 15 sites. There was an average of 11 SEM measures listed for each site, with a range of 2 to 38 measures per site. The database included measures such as:

- Non-incentivized capital measures that were rolled into an SEM engagement: small motors, variable speed drives, lighting improvements.
- Changes in process: adding a measurement step to reduce process time, changing boiler setpoints, increasing batch size.
- Changes in behavior: Training staff to operate equipment more efficiently, scheduling staff to do routine leak maintenance.

For the 15 sites that were randomly selected, all of the SEM measures listed were assessed to see if the same measures from previous years of SEM engagement were duplicated, and when found were flagged for exclusion. The measures were also reviewed to make sure they were implemented and expected to save energy between 2015 and 2017, and then the dates of first and last implementation (which may have begun prior to 2015 and ended after 2017) were appended.

---

[2] The relevant projects have an SEM start date sometime between January 2015 and December 2017, as documented in the completion reports.

[3] The following programs were included in the sample frame: SEM, High Performance Energy Management (HPEM), and Refrigerator Operator Coaching (ROC).

Note that there were no explicit energy savings listed for individual SEM measures in the BPA database (IS2.0) or the SEM completion reports.[4]

3.2 DATABASE PREPARATION

BPA (or its implementation contractor) enters data into the program database (referred to as IS2.0) each year during each SEM site's engagement period. However, only total savings are recorded. For the persistence study, the Evergreen team needed to determine whether specific SEM measures were still operational. BPA and/or its implementation contractor documented individual SEM measures in each site's annual completion report PDF file for any engagement prior to 2017. The files are of similar but not identical structure, making it impractical to extract the measure descriptions using a computer program.

The Evergreen team developed a database by manually entering the data from the completion reports, including all SEM measures, key personnel, previous engagements, and other related data from 2011-2019. This allowed the evaluation team to organize the measures by name, date, and type to support research and analysis. Some key fields for measure sampling and analysis fields are shown in Appendix B. These fields reflect measures that were expected to lead to savings in SEM regression analysis and measures that were to be deducted from SEM savings, respectively.

3.3 DATA COLLECTION

Our data collection tasks were tailored to meet the needs of the SEM persistence evaluation. The data collection tasks for the evaluation were as follows:

- **Change inventory.** For each sample participant, we created a list of SEM-initiated measures from our SEM database. We performed a detailed review of sampled participants' reports and tracking data to ensure we had all reported details of onsite measures. Data were provided by BPA to support this database task. Any measures that required additional description or implementation information were flagged for follow-up with program staff.
- **Verification plan.** For measures with sufficient detail, evaluation staff determined the least intrusive, but still reliable, method to determine:
    - The status of each measure/if the measure is no longer in place and/or it has ended;
    - When the measure was started and when the measure was stopped, estimated to the nearest quarter year;
    - The reason the measure was no longer in place and/or had ended such as a change in management, change in production, shifting to another similar measure, or a direct result of the COVID-19 pandemic; and

---

[4] The BPA database (IS2.0) provides energy savings related to SEM, but these are in aggregate for a site, not listed for each measure. The descriptions of measures extracted from the completion reports do not have sufficient detail to categorize or rank SEM measures for the full sample frame. This categorization and ranking (of relative impact on energy usage) was conducted prior to the data collection and analysis tasks.

    - If a measure was modified, whether the savings qualitatively increased or decreased.

  For measures that were no longer in place and/or had ended, we had to rely on participant staff or their vendors to tell us when the measure was no longer in place and/or had ended. For measures that were still operational, we estimated how long the measure would remain operational and whether site staff actively monitor its status (if that is appropriate).

- **Energy Smart Industrial program staff review.** Program staff reviewed the sampled measure inventory and verification plans:
    - Energy Smart Industrial (ESI) SEM program staff added detailed descriptions and implementation dates where needed for clarity.
    - ESI SEM program staff reviewed and commented on verification plans to improve data collection.
    - The evaluation team made revisions and improvements to the verification plans.
- **Collect measure data.** With the assistance of ESI program staff (and utility staff as relevant), we contacted the sampled customers and collected the needed data regarding the status of each measure and the dates reflecting that any measure was no longer in place and/or had ended. Utility staff were only asked to notify end users and assist with contacts as needed. In all cases, evaluation staff relied on file review findings, customer staff providing specifications, control system trend data and screen prints, or taking photos or videos and sending them to the evaluation team. In one instance, the persistence of a majority of measures onsite could not be reliably verified by onsite staff, and the site was replaced with a backup sample site.

3.4 STUDY ANALYSIS

To estimate the survival model necessary for developing an estimate of the effective useful life (EUL) of SEM measures, the Evergreen team gathered the following characteristics for each SEM-initiated measure from the sample of former program participants:

- Installation (or begin) date of SEM-initiated measure
- Whether the SEM-initiated measure ended or is still in place
- If the measure ended, the approximate date the measure ended

With this information, we created “time-to-event” variables and conducted survival analysis to estimate an overall EUL for SEM measures and EULs for each SEM measure type. In addition to estimating EULs, we estimated the remaining useful life (RUL) of SEM measures based on the age (i.e., time since the first documented date of implementation during program intervention) of the measure. See Appendix A for more detail on the study analysis methodology.

4 FINDINGS

This section presents descriptive information of SEM measures and the results of the assessment of the persistence of SEM measures[5] implemented through Bonneville Power Administration's (BPA's) Strategic Energy Management (SEM) offering for industrial facilities.

4.1 CHARACTERIZATION OF SEM MEASURES

Table 1 describes the study sample frame (sites active in the program between 2015 and 2017) by industry. The most common industry is manufacturing, followed by wastewater.

Table 1: Sample frame by industry

Industry	*Total Site Count*	*Count of Sites in Sample*
Manufacturing – Total	22	7
Manufacturing – Paper	8	2
Manufacturing – Building Materials	3	1
Manufacturing – Other	11	4
Wastewater	7	5
Public Administration	5	
Refrigeration Storage	3	2
Transportation	2	
Food Storage/Distribution	2	
Distilling/Food Processing	1	1
Industrial Products	1	
Mining	1	
Total	**44**	**15**

[5] BPA typically identifies SEM as a "measure", which includes all of its constituent activities. The purpose of this study was to assess the life of all of the activities that were likely to directly save energy, and not activities that are purely informational or organizational. The study team chose to define the smallest sampled unit in this study as a "measure" to be consistent with the other commonly sampled subunits in energy efficiency program impact evaluation.

Table 2 presents the count of measures associated with sampled sites by industry. A total of 224 measures were initially assessed, with 108 ultimately included in the study sample.

Table 2: Sampled customers

Site ID	Industry	Measures Included in the Sample	Measures Removed From Sample	Total Number of Measures Assessed
1003	Manufacturing	3	0	3
1008	Manufacturing	6	9	15
1004	Manufacturing	5	3	8
1013	Manufacturing	2	24	26
1014	Manufacturing	4	5	9
1010	Manufacturing	2	15	17
1022	Manufacturing	1	11	12
1005	Wastewater	21	17	38
1006	Wastewater	19	3	22
1017	Wastewater	6	0	6
1020	Wastewater	7	1	8
1007	Wastewater	3	9	12
1018	Refrigeration Storage	22	9	31
1019	Refrigeration Storage	2	5	7
1011	Distilling/Food Processing	5	2	7
Total		**108**	**116**	224

Table 3 summarizes the reasons that measures were removed from the sample—with most being removed because they were a duplicate of another measure already included in the sample, or because the measure was not associated with energy savings (e.g., attending a meeting, a study that did not result in an implemented measure). About 10 percent lacked sufficient data, including three sites that recently closed.[6]

Table 3: Reasons for measures removed from sample

Removal Reason	*Measures Removed*
Duplicate measure	50
Non-energy saving measure	30
Incentivized capital	24
Insufficient data	12
Total	**116**

Table 4 shows the SEM measure classifications for the 108 measures included in the analysis. Most of the measures were classified as Operations, which have to do with how a facility runs its primary processes. For example:

- A wastewater treatment plant that shuts down half the aeration basins during low effluent periods can reduce its energy use.
- A food packing plant insulates its canning vats.
- A brewery switches to a cooled brine instead of supply water in their cooling coil.

These activities were classified into three categories:

1. Operations: change how the process/facility/equipment is run because of SEM program intervention.
2. Physical repairs: one-time repairs of equipment because of SEM program intervention with no scheduled follow up maintenance. Could also include replacement of equipment
3. Routine maintenance: Scheduled ongoing upkeep activities because of SEM program intervention.

[6] Note that adding the measures from these three sites into the analysis as failures did not change the average effective useful life (EUL) estimate.

Table 4: SEM measure classification

SEM Measure Type	*Measures*
Operations	84
Physical Repairs	13
Routine Maintenance	11
Total	**108**

4.2 MEASURE PERSISTENCE ANALYSIS

The Evergreen team conducted time-to-event analysis (commonly referred to as "survival" analysis) to estimate the EUL of SEM measures implemented at industrial facilities in the Northwest. The "event" of interest for our analysis was the termination of an SEM measure due to failure, removal, or any other reason. Our expectation was to define EUL as the *median* length of time (in years) that an SEM measure is in place and/or is still being practiced.[7] However, as we describe below, The vast majority of the SEM measures we evaluated are still in place and operational, requiring the evaluation team to instead define EUL as the average (mean) length of time that an SEM measure is in place and/or being practiced.[8]

In addition, we estimated the remaining useful life (RUL) of SEM measures. We define RUL as the difference between the current time-since-implementation of an SEM measure and the expected time-to-termination of the SEM measure, also measured in years. The RUL represents the length of time we would expect an SEM measure to continue to be in place and operating. Whereas the EUL is an estimate of the expected service life of an SEM measure at the time of installation, the RUL is an estimate of the remaining service life of an already-implemented SEM measure. As such, the RUL accounts for the fact that the SEM measure has survived up to a specific point. To develop estimates of the overall RUL for SEM measures and RULs for each SEM measure, we used a parametric survival analysis approach. Unlike the nonparametric Kaplan-Meier approach, which is completely data driven, the parametric survival model requires an assumption of the underlying distribution for SEM measures, which allows us to develop a mathematical function to estimate RULs.

---

[7] As such, the median based EUL represents the age at which half of the equipment would still be in operation and half would have already failed. Alternatively, EUL could be defined as the average length of time (in years) that equipment is in operation. In general, the two approaches do not differ greatly; however, a mean-based EUL is typically a little higher than the median-based EUL.

[8] To conduct the survival analysis, we used the Kaplan-Meier estimator, which is a non-parametric method. As such, the Kaplan-Meier estimator requires that the event of interest—termination of the SEM activity—has occurred for at least one-half of sampled SEM measures.

We considered five alternative distributions for the parametric survival model (Weibull, exponential, normal, log-normal, and gamma). Using the data collected from onsite evaluations, we tested the performance of the different distributions and chose the exponential distribution, which provided the best fit to the data. A defining characteristic of the exponential distribution is the assumption of a constant failure rate, which suggests that, regardless of the amount of time in which an SEM measure has been in place, the expected RUL of the measure is the same as when it was first initiated. In other words, regardless of the "age" of an SEM measure, its RUL is equal to its EUL.

4.2.1 ESTIMATED EFFECTIVE USEFUL LIFE (EUL)

For the analysis, we used the Kaplan-Meier estimator, which is the most common non-parametric approach for estimating survival functions. This estimator does not require assumptions regarding the shape of the underlying survival distribution. However, estimated EULs may be biased toward longer life expectancies when a large proportion of observations are censored.[9] This may be an issue in this study, as we found that only 13 of the 108 measures that we included in the analysis have failed.

Using the Kaplan-Meier estimator, we developed an overall estimate of the EUL for SEM measures, as well as EULs for each measure type, equipment type, and industry group. We used the log-rank test to determine if there were statistically significant differences in survival times between measure types, equipment types, and industry groups.

Table 5 shows the estimated EUL for all SEM measures (8.5 years) and the estimated EULs for the three different measure types. We also estimated RULs using parametric survival analysis. We found that the exponential model best fit the underlying data. A characteristic of the exponential model is a constant failure rate, which implies that, regardless of the length of time in which an SEM measure has been in place, we would expect it to remain in place (its RUL is equal to the EUL). While such an assumption may not be reasonable for capital equipment that wears out over time at an increasing or decreasing rate, for SEM measures, which focus on optimizing processes and behaviors, it may be reasonable to assume that once in place, the SEM measure stays in place until the process is no longer needed or a superior process is implemented. Regardless, the RUL of the SEM process is a function of exogenous forces.

---

[9] For the study of SEM persistence, censoring (or more precisely, "right censoring") simply means that the event of interest (removal/failure of the SEM activity) has not yet occurred. In other words, the SEM activity is still in place and operating.

Table 5: Effective and remaining useful life, measured in years

Segment	EUL	RUL*	95% Confidence Interval Lower & Upper Bounds		n
All Measures	**8.5**	**8.5**	**3.7**	**13.3**	**108**
Operations	8.6	8.6	4.1	13.0	84
Physical Repairs	7.6	7.6	1.3	14.0	13
Routine Maintenance	8.7	8.7	2.2	15.3	11
Log-Rank Test of no difference between survival rates: ** Chi-Square = 1.74, Significance = 0.42					

* RULs were estimated using a parametric survival model, which requires a distribution to be assumed for the underlying population. Commonly used distributions include the Weibull, exponential, normal, log-normal, and gamma distributions. Using the data collected from onsite evaluations, we tested the performance of the different distributions and found that the exponential distribution best fit the data. A characteristic of the exponential distribution is a constant failure rate, which suggests that, regardless of the time in which an SEM measure is in place, one would expect its RUL to be equal to its EUL.
** The log-rank test indicates that the EULs of individual SEM measures do not differ from the EUL of all SEM measures.

The estimated EULs for the individual measure types are not statistically significantly different based on the log-rank test, which is a test of the probability of failure between the three measure types at any time point.

Table 6 shows the estimated EUL by type of equipment impacted by the SEM measure. The EUL for Refrigeration Support Services (6.3 years) appears to be substantially less than the EULs for production processes and “All Other” Support Services (9.1 years); however, the estimated EULs for each of the three equipment types fall within the confidence intervals of the equipment types. In addition, the results of the log-rank test indicate that the EULs for the three equipment types do not differ.

Table 6: Effective and remaining useful life by affected equipment

Segment	EUL	RUL*	95% Confidence Interval Lower & Upper Bounds		n
All Equipment	**8.5**	**8.5**	**3.7**	**13.3**	**108**
Production Processes	9.1	9.1	4.8	13.3	51
Support Services – Refrigeration	6.3	6.3	3.3	9.3	24
Support Services – All Others	9.1	9.1	2.2	16.0	33
Log-Rank Test of no difference between survival rates: ** Chi-Square = 3.6, Significance = 0.16					

* RULs were estimated assuming the underlying population is distributed exponentially with a constant failure rate, suggesting that, regardless of the time in which an SEM measure is in place, one would expect its RUL to be equal to its EUL.
** The log-rank test indicates that the EULs of individual equipment types do not differ from the EUL of all equipment.

Table 7 shows the estimated EUL by industry impacted by the SEM measure. The results of the log-rank test indicate that the EULs for the three industries do not differ.

Table 7: Effective and remaining useful life by industry

Segment	EUL	RUL*	95% Confidence Interval Lower & Upper Bounds		n
All Industries	**8.5**	**8.5**	**3.7**	**13.3**	**108**
Manufacturing (including food processing, distilling)	10.7	10.7	2.1	19.3	28
Wastewater	7.5	7.5	3.9	11.0	24
Refrigeration Storage	7.8	7.8	4.9	10.6	56
Log-Rank Test of no difference between survival rates: ** Chi-Square = 3.1, Significance = 0.21					

* RULs were estimated assuming the underlying population is distributed exponentially with a constant failure rate, suggesting that, regardless of the time in which an SEM measure is in place, one would expect its RUL to be equal to its EUL.
** The log-rank test indicates that the EULs of individual industries do not differ from the EUL of all industries.

4.2.2 SURVIVAL CURVES

Figure 1 shows the smoothed survival curves for all SEM measures and for the three SEM measure types. The survival curves were constructed using a parametric survival model, which was necessary due to the small number of measures for which we have complete information. That is, of the 108 SEM measures evaluated by the Evergreen team, only 13 were found to be no longer in place and operating. On average, measures were removed within 1.7 years, as shown in Table 8 below.

*Table 8: Average age of SEM measures that were removed (n=13)**

Category	Sub-Category	Average Age
SEM Measure Type	Operations	1.7
	Physical Repairs	1.8
	Routine Maintenance	1.0
Equipment	Production Processes	1.3
	Support Services – Refrigeration	1.0
	Support Services – All Others	2.1
Industry Group	Manufacturing	1.7
	Wastewater	1.0
	Refrigeration Storage	2.0
Overall		1.7

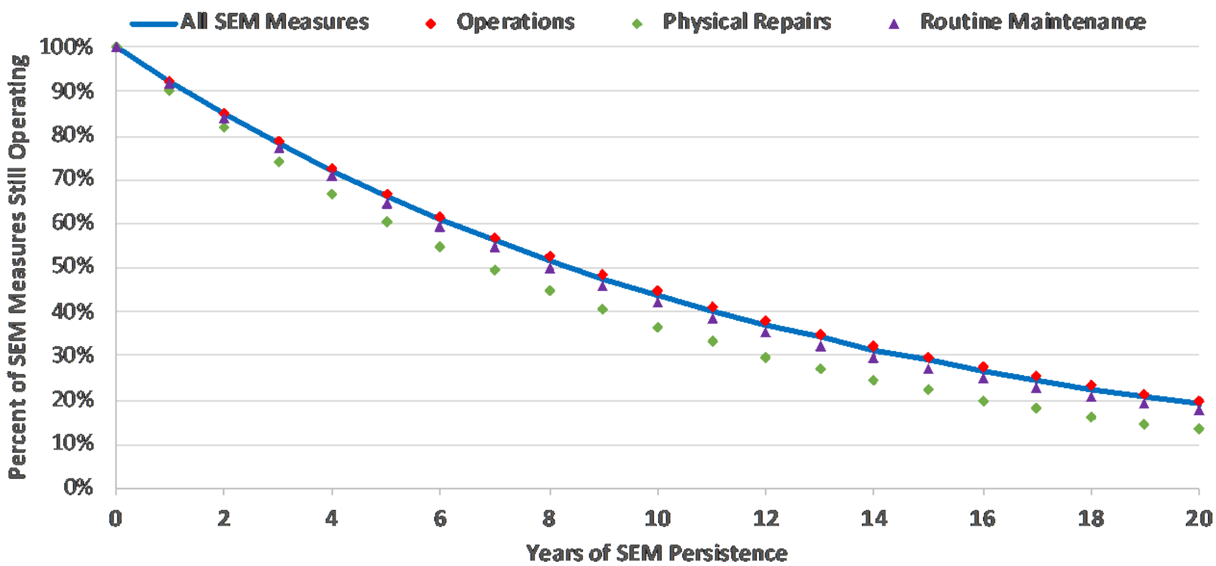
* The median age of an SEM measure that was removed was one year.

For the 13 SEM measures that were no longer in place, we know the year of installation and/or initiation, that they are no longer in place, and the year in which they were removed and/or discontinued. For these 13 measures, we have complete time-to-event information.

For the other 95 SEM measures we only know the year in which the measure was installed and/or initiated and that it is still in place and operational. Since no event has occurred—i.e., the SEM measure has not been removed and/or discontinued—we do not have complete time-to-event information for these measures. The statistical term for a situation in which the outcome of an observational unit is only partially known is *censoring*.[10] As a non-parametric approach, the Kaplan-Meier method does not predict the time-to-event of censored observations. Instead, to estimate the time-to-event for censored observations, the Evergreen team utilized a parametric survival model, which allowed us to develop the survival curves shown in Figure 1.[11]

The survival curves for the Operations and Routine Maintenance measure types are effectively indistinguishable from the survival curve for All SEM Measures, while the survival curve for Physical Repairs is slightly steeper. We expect that half of SEM measures would still be in place after 8.5 years (i.e., the EUL) and that about 20 percent of SEM measures would be in place after 20 years.

Figure 1: Smoothed survival curves for SEM measures



[10] For time-to-event data, the activity end date is censored for all measures that are still in place at the end of the "period of observation." For our purposes, the "period of observation" would be the years between the date of an activity begun through the respective utility program and the date in which we would verify whether the activity has ended or is still in place.
[11] Please see Appendix A for more information of the parametric survival model the Evergreen team used to develop the survival curves shown in Figure 1.

5 KEY FINDINGS AND RECOMMENDATIONS

This section first presents important caveats to consider when interpreting the study results, and then summarizes the key findings from the study and offers a recommendation for facilitating more robust Strategic Energy Management (SEM) measure savings and effective useful life (EUL) analysis.

5.1 STUDY CAVEATS

This study was intended as an initial effort to use the available data and apply EUL analysis techniques to inform estimates of measure life (i.e., date of first documented measure implementation until measure failure/deactivation) for BPA's SEM program. There are two important caveats to consider when interpreting and applying these results to BPA's and other SEM programs.

1. Our study was designed to inform measure life (whether the measure is still active or not), which is distinct from measure savings persistence (i.e., whether the measure is still saving in future years what it was intended to in year one). Savings is an additional parameter to consider that may change over time and would need to be applied along with estimated measure life to get a more complete understanding of measure savings persistence.
2. The underlying information we used to develop a database from which to draw the study sample is based on SEM project completion reports after a year of participation and may underrepresent measure failures (i.e., measures tried by facilities but stopped immediately during the engagement year). Our measure life estimate may be biased upwards as a result of the database assembled from SEM program completion reports, which were the only uniformly available source of implemented measures during the studied period. The reports contained measures that the program staff assumed were worth reporting on as potentially impacting savings and may not represent all SEM measures attempted.

   Note that this concern only relates to the lifetime of reported measures/ activities, not the savings that they yield.

Given the stated issues with the program measure data that we used to draw the study sample, BPA may consider choosing a conservative estimate of measure life from this study, such as on the lower end of the estimated confidence bounds.

5.2 KEY FINDINGS

The key findings from this study are:

Key Finding 1 - Overall EUL: The estimated EUL for BPA's SEM measures is 8.5 years, with 95 percent confidence bounds of 3.7 to 13.3 years.

Of 108 measures assessed in the study (where program engagement was initiated between 2015 and 2017), only 13 were no longer in operation. On average, measures were removed within 1.7 years.

Key Finding 2 - Variation in EUL: The estimated EUL does not differ significantly based on type of SEM measure, equipment type, or industry type.

5.3 RECOMMENDATION

We offer one recommendation to support more robust evaluation of SEM measure life.

> **Recommendation** - We recommend that BPA and its SEM program comprehensively track SEM engagements in the SEM program database and include the specific fields listed in Appendix B.

BPA (or its implementation contractor) currently enters data into the program database during each SEM site's engagement period. However, only total savings are recorded. For the persistence study, the Evergreen team needed to determine whether specific SEM measures were still operational. BPA and/or its implementation contractor document individual SEM measures in each site's annual completion report PDF file for any engagement prior to 2017.

Key data from the completion reports could be entered routinely into the program database, including all SEM measures, key personnel, and previous engagements, which would facilitate evaluation and measure life research and analysis. Some key fields for measure sampling and analysis are shown in Appendix B. These fields reflect measures that were expected to lead to savings in SEM regression analysis and measures that were to be deducted from SEM savings, respectively. Documenting all measures, active or otherwise, will also reduce likelihood of bias when determining EUL and savings persistence. Identifying a complete history of successful and unsuccessful measures may benefit engineering staff by identifying measures with the highest rates of success. It is not the opinion of the evaluation team that the evaluation taxonomy should necessarily be used; any reliable and documented database should be sufficient to aid in supporting SEM programs and evaluation.

APPENDICES

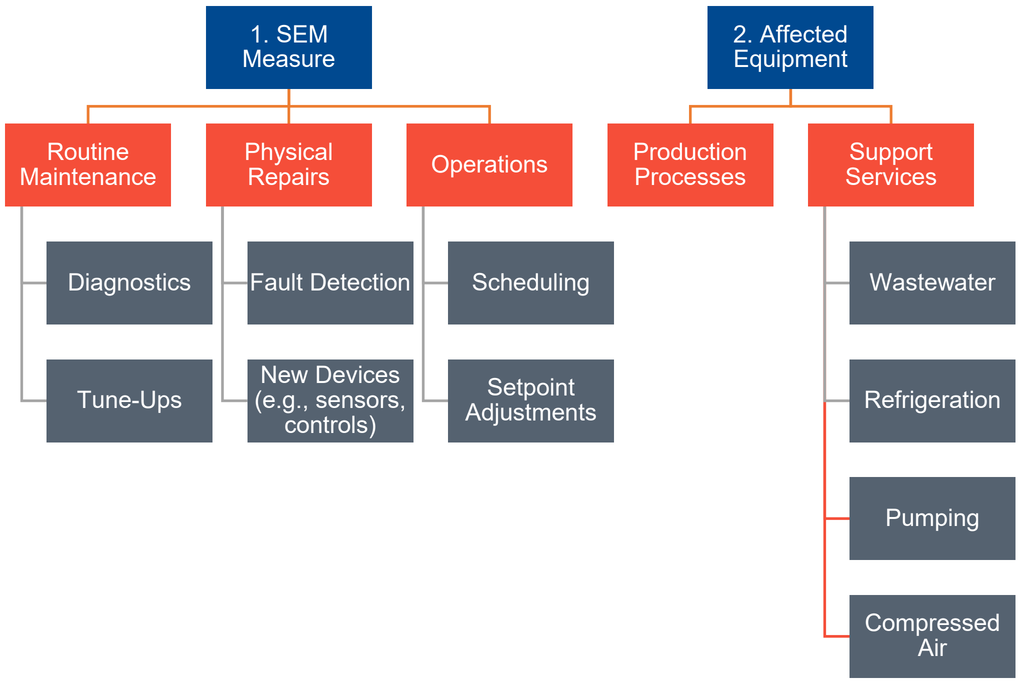
APPENDIX A: DETAILED ANALYSIS METHODOLOGY

This section provides more detail on the study analysis approach.

CLASSIFYING SEM MEASURES INTO MEASURE GROUPS

The Evergreen team examined the list of SEM-initiated measures and developed a scheme for classifying these measures into like groups based on factors such as the type of SEM measure and affected equipment, as shown in Figure 2.

Figure 2: SEM Change classification framework



GATHERING INFORMATION TO ESTIMATE LIFETIMES OF SEM MEASURE GROUPS

To estimate the survival model necessary for developing an estimate of the effective useful life (EUL) of SEM measures, the Evergreen team gathered the following characteristics for each SEM-initiated measure from the sample of former program participants:[12]

- Installation (or begin) date of SEM-initiated measure
- Whether the SEM-initiated measure ended or is still in place

[12] The Evergreen team also developed estimates of the EUL for each of three different activity types (Operations [including delamping], Physical Repairs, and Routine Maintenance) to test whether EUL differs between activity type.

- If the measure ended, the approximate date the measure ended

With this information, we created the "time-to-event" variables necessary to conduct survival analysis:

- **Status:** A binary variable that equals 1 if the event has occurred (i.e., the SEM-initiated measure is no longer in place and/or it has ended), else 0.
- **Time:** Length of time in years between initiation of the measure and the event
  - If event = 1 → the time is the difference between the date the measure was installed (or begun) and the date the measure ended.[13]
  - If event = 0 → the time is the difference between the date the measure was installed (or begun) and the date the status of the measure was checked.

SUMMARIZING LIFETIMES OF SEM MEASURE GROUPS

Using the time-to-event data described above, the Evergreen team conducted survival analysis to estimate an overall EUL for SEM measures and EULs for each SEM measure type. Survival analysis methods are used to analyze data when the outcome of interest is the time until an event occurs and are a better choice than standard methods of statistical modeling such as linear or logistic regression, which do not account for both the status of an event and the timing of when the event occurred, nor do standard regression methods adequately account for the censoring characteristic of time-to-event data.[14] When censoring is present in the data, estimates of the true time-to-event will be underestimated unless statistical methods specifically designed for time-to-event data are used. Survival analysis methods utilize the information on all observations—those that have experienced the event and those with censored data—to provide unbiased estimates of the future time-to-event for each censored observation.

The general survival function below defines the probability of survival (i.e., that "the event" has not occurred) at time *t*:

$$S(t) = Pr\{T \geq t\} = 1 - F(t) = \int_{t}^{\infty} f(x)dx,$$

[13] While we anticipate measuring time in years, we will not constrain the time to be measured in whole years.

[14] In statistics, the term *censoring* refers to the circumstance in which the value or outcome of an observation is only partially known. For time-to-event data, the activity end date is censored for all measures that are still in place at the end of the "period of observation." For our purposes, the "period of observation" would be the years between date of an activity begun through the respective utility program and the date in which we would verify whether the activity has ended or is still in place.

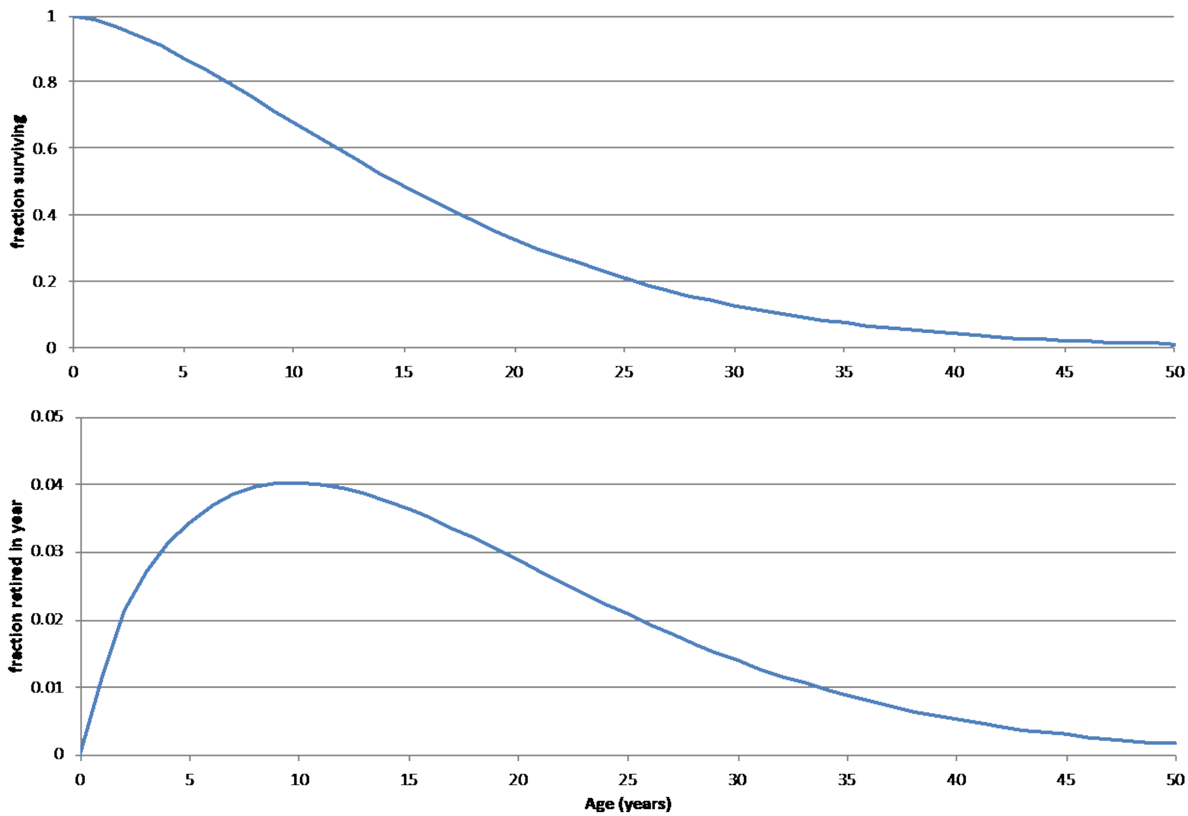
Alternatively, the hazard function below characterizes the instantaneous rate of a measure ending (i.e., the probability “the event” will occur) at each point along the survival function:[15]

$$\lambda(t) = \lim_{dt \to 0} \frac{Pr\{t \leq T < t + dt | T \geq t\}}{dt}$$

The survival function and the hazard are related, and if one is known, the other can be computed. Figure 3 shows an example of a smoothed survival function and corresponding hazard function. The survival function (upper figure) shows the proportion of a population expected to survive over a 50-year time period. The lower figure shows the proportion of the population expected to end the measure (i.e., experience the event) each year. In this example, the hazard rate grows through age 10 and then begins to decline. The survival function and hazard function are inversely related.

Figure 3: Example of a survival function (upper) and hazard function (lower)

[15] The numerator of the hazard function is the conditional probability that the event has occurred given that it has not occurred before; the denominator is the width of time interval (e.g., day, month, year).



For the analysis, we used the Kaplan-Meier estimator, which is the most common non-parametric approach for estimating survival functions. This estimator does not require assumptions regarding the shape of the underlying survival distribution; however, estimated EULs may be biased toward longer life expectancies when a large proportion of observations is censored.[16] Using the Kaplan-Meier estimator, we developed an overall estimate of the EUL for SEM measures and EULs for each measure type. We used the log-rank test to determine if there were statistically significant differences in survival times between measure types.

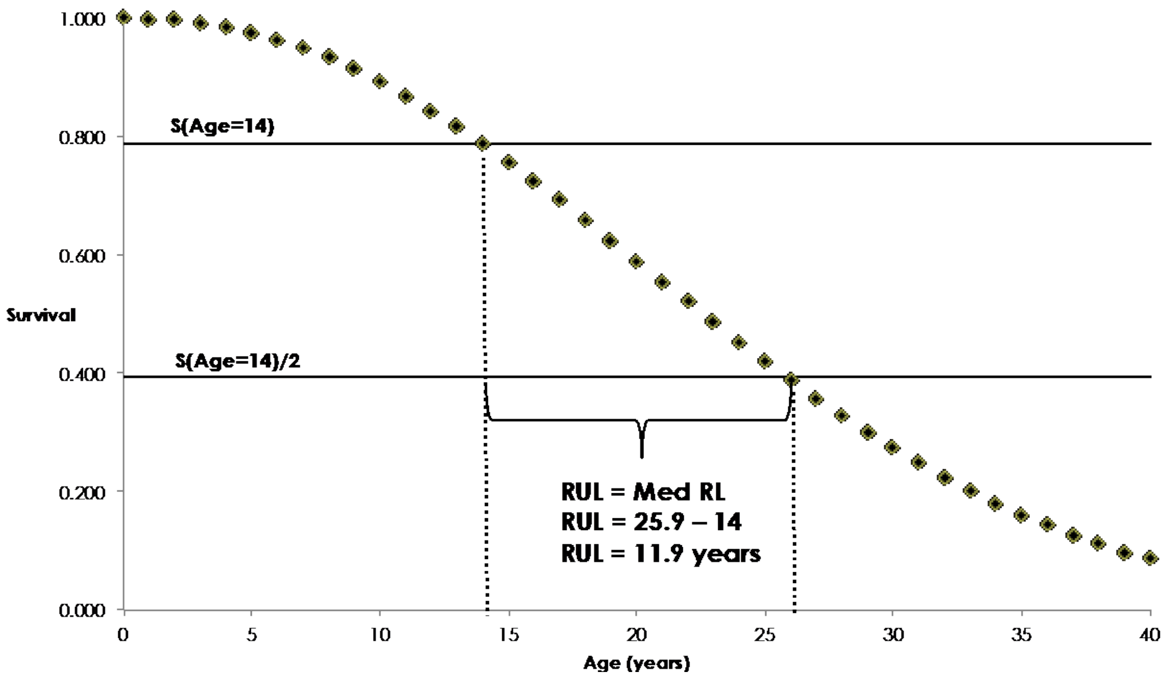
In addition to estimating EULs, we estimated the remaining useful life (RUL) of SEM measures based on the age (i.e., time since implementation) of the measure). Figure 4 shows how a survival function is used to estimate RUL based on computing median residual life (MRL) of a measure. In this example, the measure has been operational for 14 years. Computing the MRL requires four steps:

1. Determine the survival probability at a given age. Example: Survival (age=14) ≈ 0.8
2. Divide the survival probability from Step 1 in half. Example: 0.8/2 ≈ 0.4

[16] For the study of SEM persistence, censoring (or more precisely, "right censoring") simply means that the event of interest (removal/failure of the SEM activity) has not yet occurred. In other words, the SEM activity is still in place and operating.

3. Determine the age that corresponds with the survival probability calculated in Step 2. Example: approximately 26 years
4. Subtract the current age of the measure from the age determined in Step 3. Example: 26–14 ≈ 12 years

Figure 4: Computing median residual life based on a survival curve



Source: Early Replacement Measures Study, Phase II Research Report, A report to the Regional Evaluation, Measurement and Verification Forum facilitated by Northeast Energy Efficiency Partnerships, prepared by Evergreen Economics, Michaels Energy, and Phil Willems.

APPENDIX B: KEY FIELDS FOR SEM MEASURE DATABASE

Key fields that would be useful for the BPA SEM program to comprehensively collect in its program database are listed in the following two tables. Table 9 represents fields used by this study, which could be the basis of follow up SEM studies. Table 10 includes details necessary for cross verification of capital measure deduction from SEM savings regressions.

Table 99: SEM measure fields

Fields	*Short Description*
Program Sample ID	Unique ID assigned for each SEM activity
Site	End User/Site
Program Year	Which year of SEM engagement
Utility	Utility or District serving end user
Program	Type of SEM engagement
Prior Engagement	If end user was in a different program, what was it?
Energy Team (Comma Separated)	Key end user personnel
SEM Start	Start date for SEM savings analysis period
SEM End	End date for SEM savings analysis period
Measure Name	Description of SEM activity
Measure Occurrence	Number of times measure attempted if repeated
Operations	Operations SEM measure designed to change energy use in day-to-day processes
Maintenance	Maintenance SEM measure designed to keep equipment running efficiently and in good repair
Capital	Capital SEM measure involving equipment investment
Awareness	Information SEM measure designed to influence behaviors
Completion Status	Was the measure completed
Completion -Year	Year measure was completed
Completion -Month	Month measure was completed

Table 1010: Capital measure fields

Fields	*Short Description*
Capital Project Index	Unique ID assigned for each SEM related activity
Company	End User/Site
Program Year	Which year of SEM engagement
Date Year	Year measure was completed
Utility	Utility or District serving end user
Implementer	Who implemented the capital measure
Program Type	Type of SEM engagement
Prior Engagement	If end user was in a different program, what was it
Energy Team (Comma Separated)	Key end user personnel
SEM Start	Start date for SEM savings analysis period
SEM End	End date for SEM savings analysis period
Capital Project Name	Descriptive name of capital project
Capital Project ID	Unique incentivized project ID in utility or BEETS/IS2 database