Into the Great Wide Open: A Comparison of M&V 2.0 and Traditional Evaluation Methods for a Small Business Direct Install Program

Joseph Dolengo, National Grid
David Barclay, NMR Group, Inc.
Scott Walker, NMR Group, Inc.
Julian Ricardo, NMR Group, Inc.

ABSTRACT

Measurement and verification (M&V) 2.0 promises a bright future of instantaneous evaluation without the need for sampling customers, using census-level analysis instead. What role does traditional M&V play in this future? Should it serve to validate emerging techniques? In this paper, the authors explore the results of concurrent M&V and M&V 2.0 metering studies performed on lighting system retrofits at a sample of small business direct install customers from one Northeastern state. Data for this paper includes a variety of metering techniques, from analysis of building- and circuit-level advanced metering infrastructure (AMI) data to more traditional approaches involving photocell loggers and billing analysis.

The evolving realm of M&V merits further scrutiny to better assess its costs, accuracy, and feasibility compared to traditional approaches. The investigation of simpler systems like the lighting retrofits considered here lays out methods to quantify the potential for new forms of coordination and aims towards broader comparative analysis of new and existing M&V methods in the future. Results indicate that both traditional and M&V 2.0 methods benefit from greater coordination among administrators, implementers, and evaluators. Nonetheless, further work is necessary to address how the M&V 2.0 techniques we considered apply to more complicated measures.

The paper reports on the findings of this study, still underway, with a focus on the potential for greater coordination between implementers and evaluators when installing and retrieving data collection equipment, coordinating sampling, and managing customer contacts. We expect the study to provide additional insights into improving resource allocation in future program evaluations, especially with real-time data collection and analysis.

Introduction

This paper assesses the alignment of different metering technologies and estimates based on reference values in quantifying energy savings from lighting retrofits. Among the various models and technologies discussed, the common thread is a comparison of results derived using sub-hourly interval data from networked devices against existing approaches for calculating program-related lighting energy savings. We investigated three approaches relying upon advanced metering technologies: (1) monthly billing data; (2) daily and hourly interval meter data; and (3) circuit and subcircuit meter data. In turn, we calculated the ratio and root-mean-square-error (RMSE) of savings estimates from these methods against contractor estimates and two different savings baselines: (1) concurrently installed photocell-based data loggers, and (2) reference hours-of-use figures (HOU) from the New York State Technical Resource Manual (NYTRM v5, 2017) in conjunction with site-level wattage figures. Beyond probing the alignment and variability of resulting savings estimates, we also analyzed the influence, if any, of building end-uses, duration of data collection, and number of fixtures on savings estimates.

We found the closest alignment with both baseline estimates when using subcircuit meter data and contractors' HOU estimates. Setting aside questions of cost for now, these results suggest that circuit-based data models can at least supplement more-established M&V methods with regard to making accurate usage measurements. Assuming that circuit-level metering becomes more widespread, methods for calculating savings from real-time or near real-time interval data might allow for performing M&V with less time spent installing and retrieving temporary equipment to monitor usage. This study, though, provides inconclusive evidence as to whether circuit-level metering would improve the accuracy of lighting savings estimation as compared to using data loggers or deemed values based on contractor estimates. By comparison, monthly billing, daily interval, and hourly interval models produced results comparable to subcircuit, logger, and deemed values at one-third (34%) of tested sites.

Methodology

Measurement and verification (M&V) provides an empirical basis for determining whether the savings projections of an energy-efficiency program reflect realized savings, and to what extent. M&V can provide value to end users in the form of increased investment to state and utility energy-efficiency programs, and in the form of compliance with legislation. M&V can also provide value to stakeholders in electric resource markets that commodify efficiency as a resource.

The M&V techniques we compare in this paper test the validity of savings estimates from metered lighting usage. These methods include:

- account-level billing analysis
- whole building metering
- subcircuit metering
- photocell on/off logging from a representative sample of circuits

In order to provide accurate M&V of lighting projects, all unnecessary variables that could influence usage measurements must be eliminated, or at least mitigated in their contribution to final savings calculations. As such, the guidelines for the study follow from the principles described by Pacific Northwest National Laboratory for standardized M&V of lighting retrofit projects in buildings. These include the following:

- A new and/or sufficiently well-maintained baseline lighting system.
- Consistency in surrounding structures, occupants, and weather (unless weather variables factor into the savings model).
- Minimum accuracy requirements for metering equipment.
- Installation of metering equipment such that only the light usage from tested technologies is captured.

Building and Retrofit Typology

We monitored 32 sites as part of this research, using both networked metering devices and standalone data loggers to collect information about lighting usage. Figure 1 shows the distribution of building types after developing a representative sample of customers and building types.

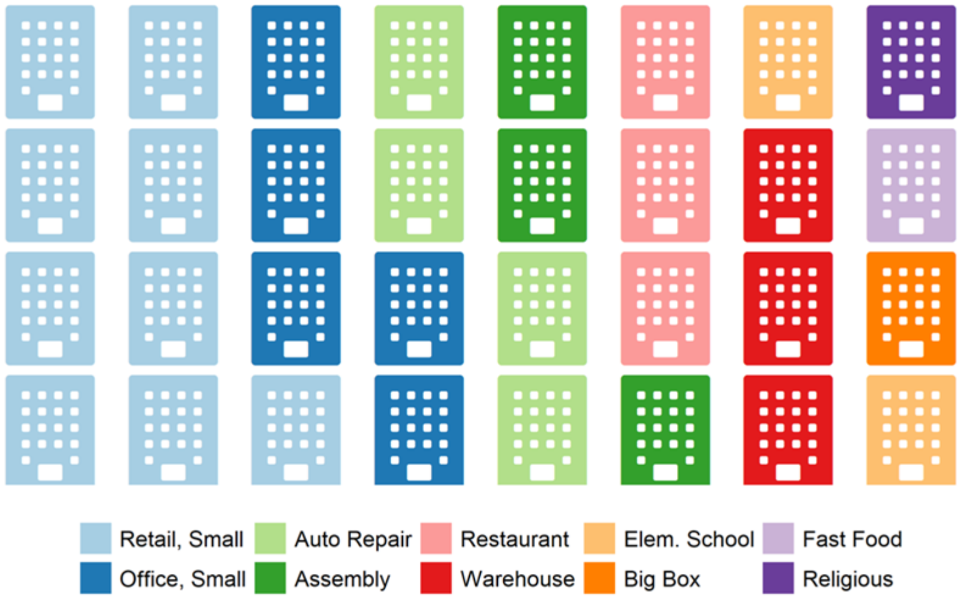

Figure 1. Typography of sites studied as part of this research, adapting the categories from those listed in the NYS Technical Resource Manual (NYTRM v5 2017).

The figure adapts the categories listed in the New York State Technical Resource Manual (NYTRM) for classifying buildings and assigning them reference levels of lighting usage. We selected candidates on a rolling basis based on their agreement to have equipment installed on-site, as well as the age and state of their electrical equipment. We also performed preliminary checks on the quality of data captured at each site in conjunction with the entities responsible for modeling savings from interval data.

Altogether, there were 309 data loggers monitoring the lighting usage of 312 retrofitted fixtures across 32 sites, with a median logging period of 84 days. Among these 312, 164 were T8 fluorescents swapped for T8 LEDs and 129 were T12 fluorescents swapped for T8 LEDs, leaving 19 cases where lights were replaced with other LED lamps. Overall, 97% of existing lamps were either fluorescent T8 or T12, and all replacement lamps were LEDs.

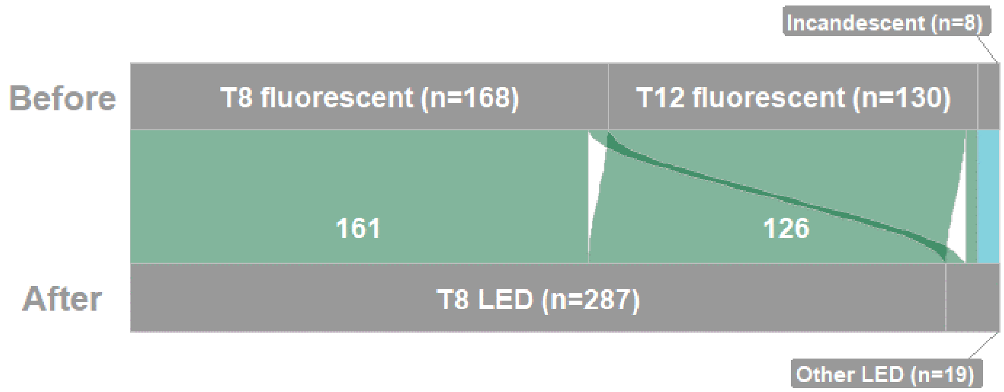

Figure 2. Typography of lighting retrofits studied as part of this research, showing the overwhelming presence of fluorescent-to-LED retrofits.

We selected a subset of representative fixtures to monitor with data loggers based on the size and number of distinct circuits at each site. Where possible, we paired all circuits at a site with a logger. Occasionally, fixtures did not permit logger installation (e.g., some screw-in fixtures) or presented too great an intrusion for customers. In these instances, we collaborated with implementers to ensure that loggers captured lighting usage from a majority of site fixtures.

Metering Hardware

The photocell-based logging equipment deployed to monitor light changes consisted of Onset HOBO UX90-002x Data Loggers. Data loggers were installed in lighting troffers on a representative sample of newly retrofitted fixtures, with an average of ten loggers at each site.

In addition to photocell data loggers, two separate energy measurement strategies were installed at each site. To meter building-level electricity usage, a kW meter was installed on the main electrical panel of each building, using current transformers (CTs) to measure the amperage on each main feed and voltage taps to measure the voltage for each phase. The installation of the kW meter occurred at the beginning of each site's M&V installation. The CTs and voltage taps supplied data to a unit that processed it before transmitting it to servers via cellular connection for aggregation. This aggregate data was then updated at regular five-minute intervals.

In addition to building-level metering, current transducers were installed to sub-meter the amp-hours of all lighting circuits in the facility. This required a site mapping that was used to develop switch-level scopes for each facility. By measuring the usage on each circuit and testing switches in the facility, the wattage of each switch could be determined and associated with the appropriate circuit. Light logger installation occurred during switch-to circuit installation. At the completion of this process, all identified lighting circuits were sub-metered at the electrical panel.

After switch-to-circuit association, the installed CTs transmit amperage data every six seconds from each measured lighting circuit. This data is received via a locally installed gateway for data processing and calculation.

Models

Table 1. Comparison of Data Requirements for Analytic Methods

	Data Type				
Technique	Project	Metered	On/Off	Weather	References
Billing	X	X		X	CalTRACK
Whole building metering	X	X		X	CalTRACK
Subcircuit metering	X	X			PNNL
Lighting loggers	X		X		PNNL

Account-Level Monthly Billing Analysis

Billing analysis follows the CalTRACK Site-Level Monthly Weather Normalized model for calculating energy savings.[*] This approach consists of a parametric model using ordinary least-squares to estimate daily energy use based on monthly (bill-period interval) data, and heating and cooling degree days (HDD and CDD, respectively), accounting for differences in pre- and post-intervention normal temperatures.

Daily Interval

Whole building meter analysis follows the CalTRACK Site-Level Daily Weather Normalized model for calculating energy savings. Pre-intervention average daily energy use is interpolated from pre-intervention monthly billing data, whereas post-intervention data is acquired at daily intervals. Otherwise, the approach involves a parametric model similar to monthly billing analysis.

Hourly Interval

Hourly data analysis involves a weather normalized parametric model for calculating energy savings similar to the model applied to monthly and daily energy usage data. Pre-intervention average hourly usage is interpolated using monthly data, but post-intervention data is acquired at hourly intervals.

Circuit and Subcircuit Metering

An initial visual inspection of pre-retrofit data helps to determine the typical weekly operating schedule for each site, in addition to discussions with the customer or facilities manager. From there, subcircuit meter readings in amps[†] are taken at six-second intervals, aggregated at hourly intervals, and flagged depending on whether they occur during operating or non-operating periods. Daily savings estimates are then taken as the difference between actual post- retrofit usage and the appropriate operating or non-operating pre-retrofit average. Deemed

[*] Documentation of the open-source CalTRACK project describes it as “a set of methods for calculating site-based, weather-normalized, metered energy savings from an existing conditions baseline and applied to single family residential retrofits using data from utility meters.”

[†] Measured amperage data from CTs multiplied by a nominal voltage value, multiplied by a power factor of 1.

savings for any un-metered 24/7 subcircuits are included afterward, consistent with how this lighting usage was accounted for with the photocell loggers.

Photocell Loggers

We extract the retrieved loggers' on/off data and merge with site-level information, including pre- and post-retrofit light fixture wattages, quantities, and descriptions. We then supplement the logger data with variables representing building end-uses, operating hours, lighting hours of use (HOU) as estimated by implementers, and circuit-level identifiers. A separate conference paper, "Time to Move On" (2018), describes the procedure for retrieving and analyzing logger data in greater detail.

Results

We calculated two sets of savings estimate ratios at each site by separately normalizing savings estimates from each method with multiple baselines. The first baseline reflected hours of use (HOU) captured with data loggers, while the second was calculated using HOU values from NYTRM, which reflect findings from two 2008 lighting usage studies utilizing data loggers. We derive mean savings ratios for each method and baseline. As expected, the logger-TRM and TRM-logger ratios are inverses of one another on a site-by-site basis.[‡]

Across both baselines, the results from each method fell into one of two disparate categories, as determined by RMSE, where we define error as the difference between any measured savings estimate ratio and the ideal ratio of 1. Figure 2 compares each method's performance based on this metric, quantifying the stark differences in alignment with baseline savings estimates. RMSE values from the subcircuit model and contractor estimates fell between 0.7 and 0.9, showing a gap of an order of magnitude in alignment when compared to monthly billing, daily interval, and hourly interval models' RMSE. Values for the latter group fell between 5.2 and 20.

Overall, subcircuit model results aligned closest with the data logger savings estimates, which are assumed to be the most accurate measure of program-induced lighting savings. This method also produced results that were the most similar to the NYTRM baseline savings estimates, and on par with contractor estimates. By comparison, monthly billing, daily interval, and hourly interval models produced results comparable to subcircuit, logger, and deemed savings at 11 of 32 sites (34%), spanning 5 of 11 building types: offices, auto repair shops, warehouses, assembly spaces, and big box stores.

[‡] This relationship does not necessarily hold when we aggregate the results, as shown in Tables 2 and 3. For this reason, we also compare the root-mean-squared-error (RMSE) of the savings estimates.

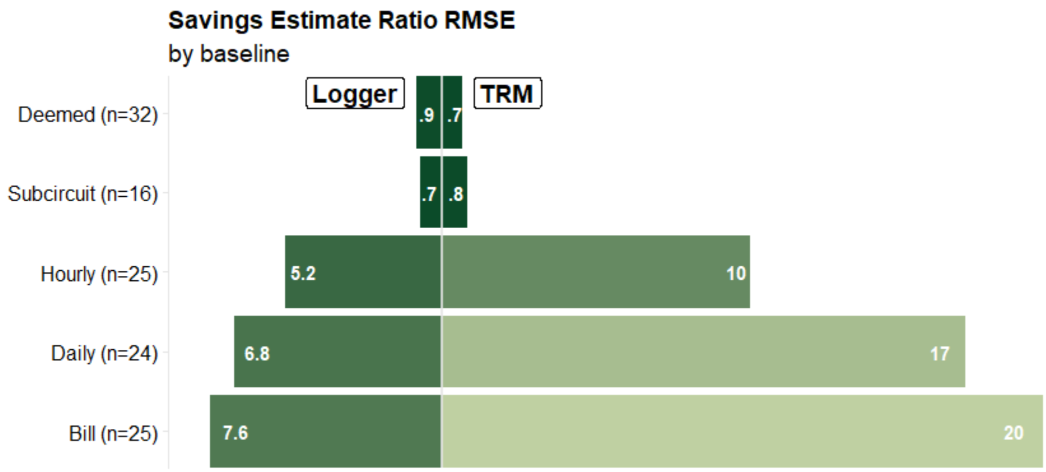

Figure 2. Comparison of savings estimate ratio root-mean-squared-error by building type and method, using data logger savings (left) and NYTRM (right) as the baselines.

Logger baseline

Altogether, the subcircuit model performed best in replicating the savings estimates derived from data loggers, with a mean savings estimate ratio of 0.97 (n=16 sites). Savings estimates calculated using contractor and NYTRM HOU values followed in terms of alignment, with mean ratios of 1.56 (n=32 sites) and 1.70 (n=32 sites). On average, the hourly, daily, and monthly billing models produced savings estimate ratios at least a factor of two greater: 3.78, 4.84, and 5.12, respectively. Looking at the RMSE associated with each method provides a similar picture of alignment among different methods, though ultimately both estimators are sensitive to outlier savings ratios >1.

Table 2 compares the average savings estimate ratio[§] and RMSE across all tested methods. The results support Figure 2, showing that the subcircuit model, contractor estimates, and NYTRM values were in greater alignment with logged lighting usage than the other methods. Comparing savings estimate ratios, the subcircuit model showed even greater alignment with the lighting logger results than either the contractor or NYTRM estimates.

Comparing rates after categorizing buildings by end-use complicates the outlook for which methods are most accurate in estimating savings. Figure 3 illustrates the extent to which these aggregated site-level savings estimate ratios are equal to or fall within 5% of the logger-based determination of savings. As with Figure 2, Figure 3 shows that the circuit-based model tends to align more closely with logger-based savings.

[§] Calculating the savings estimate ratio involves dividing the savings determined using one method (e.g. hourly metering data) by savings determined using a different method (e.g. logger data), referred to here as the baseline method.

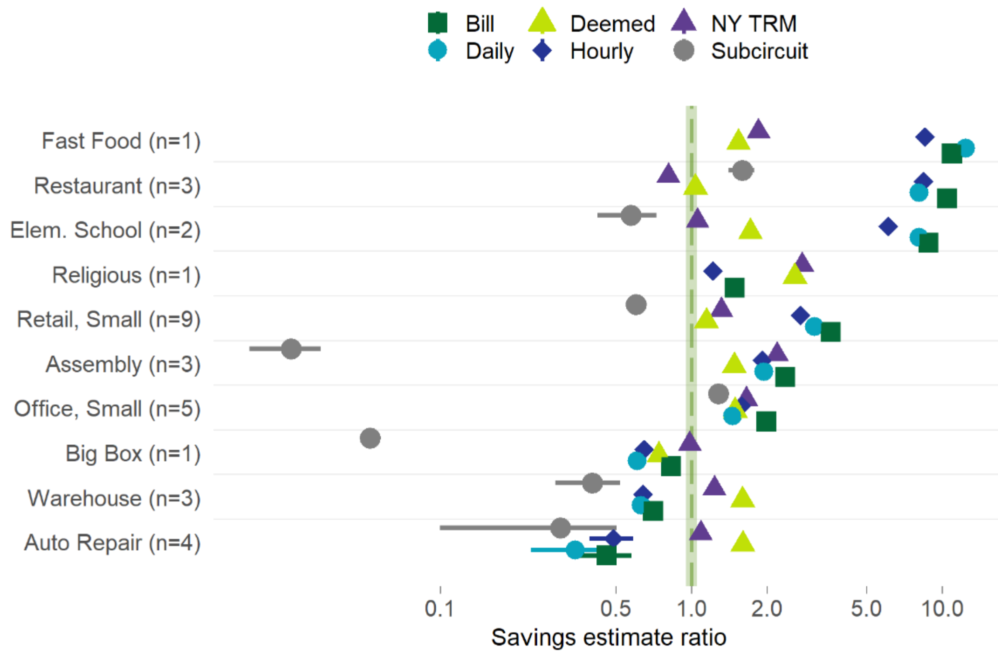

Figure 3. Comparison of savings estimate ratio by building type and method, using data logger savings as the baseline. The shaded region covers savings estimate ratios within 5% of the ideal value of 1.

Table 2. Summary of Savings Estimate Ratio by Method, Logger Baseline

	Deemed	NYTRM	Subcircuit	Hourly	Daily	Bill
n (sites)	32	32	17	25	24	25
Mean Savings Est. Ratio	1.56	1.70	0.97	3.78	4.84	5.12
RMSE	0.87	1.45	0.73	5.16	6.84	7.62

TRM baseline

Comparing savings estimate ratios derived using the TRM methods, we again found that subcircuit estimates overall aligned more closely with the baseline savings estimates than the others, with the same exceptions for offices, auto repair shops, warehouses, assembly spaces, and big box stores.

Table 3 and Figure 4 summarize the average savings estimate ratio and RMSE for each method analyzed here, using savings based on the NYTRM figures to normalize savings estimates. The data loggers, subcircuit model, and contractor estimates performed best in replicating the savings estimates derived using the NYTRM operating hours, with respective mean ratios of 0.89 (n=32 sites), 0.88 (n=17), and 1.17 (n=32).

As with Table 2, Table 3 indicates that the subcircuit model and contractor estimates tended to align more closely with baseline estimates than the hourly and daily interval models, or the monthly billing model.

However, two principal differences in the results for the NYTRM baseline are worth highlighting. First, subcircuit estimates showed greater alignment than contractor estimates in terms of the mean savings ratio, but the latter produced a lower RMSE. This suggests that while both methods produced estimates centered around the baseline value, the distribution of contractor estimates around it was narrower. Meanwhile, logged levels of lighting usage produced the closest average savings ratio to the ideal value of 1 across all methods, and a RMSE comparable to the contractor estimates'.

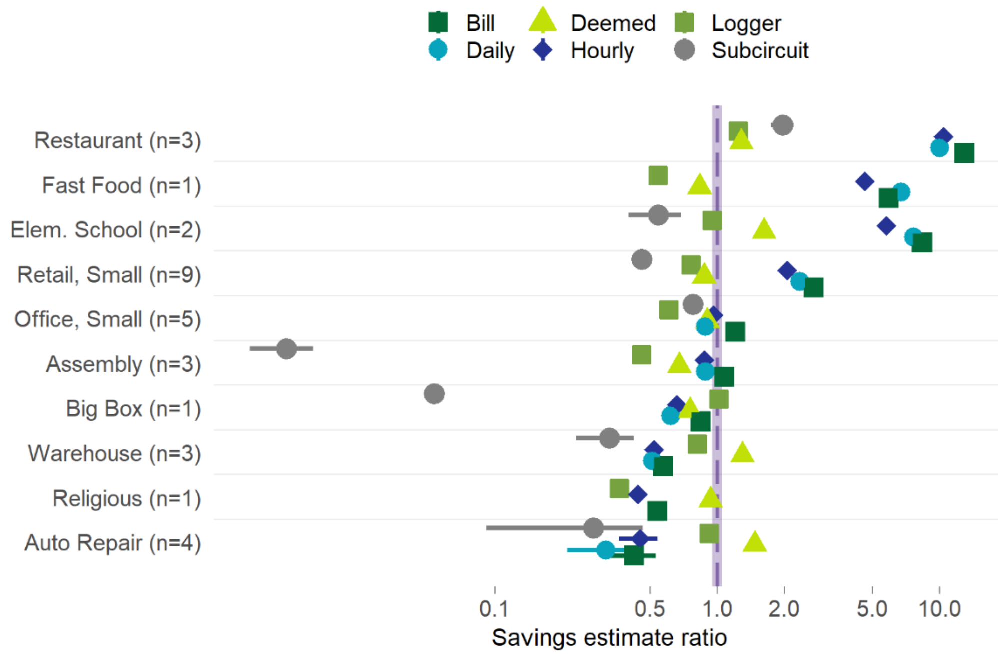

Figure 4. Comparison of savings estimate ratio by building type and method, using NYTRM savings as the baseline. The shaded region covers savings estimate ratios within 5% of the ideal value of 1.

Table 3. Summary of Savings Estimate Ratio by Method, TRM Baseline

	Logger	Deemed	Subcircuit	Hourly	Daily	Bill
n (sites)	32	32	16	25	24	25
Mean Savings Est. Ratio	0.89	1.17	0.88	4.54	6.59	7.17
RMSE	0.68	0.67	0.80	10.10	17.14	19.68

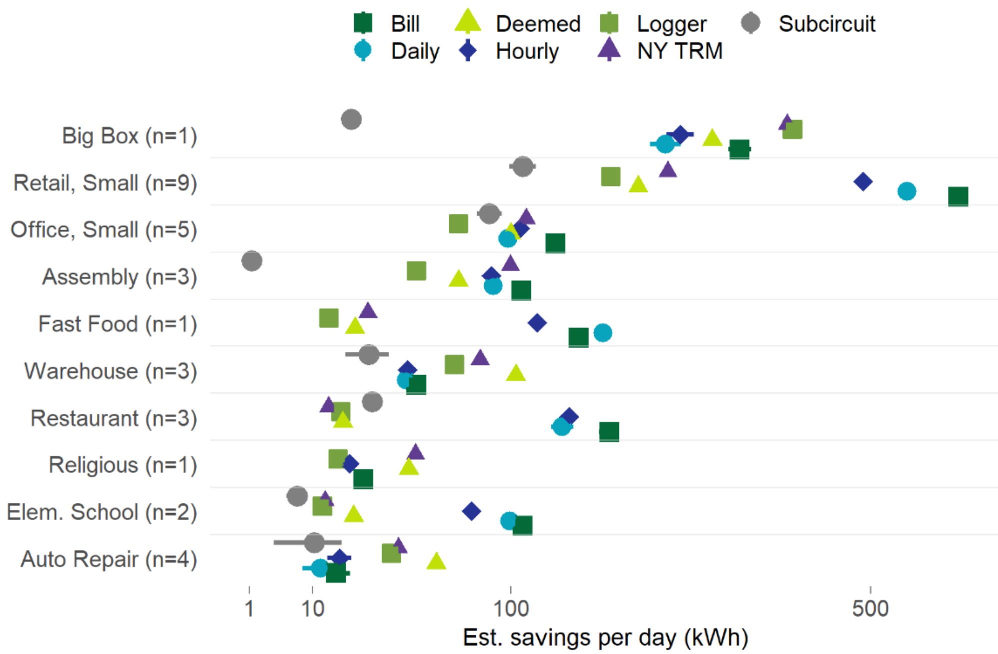

Figure 5. Comparison of savings by building type and method

Conclusions

This study provides inconclusive evidence as to whether circuit-level metering would improve the accuracy of lighting savings estimation as compared to using data loggers or deemed values. At most of the facilities participating in the study (n=32), circuit-based data models produced savings estimates comparable by the average and RMSE of savings ratio to those based on deemed values (e.g. TRM calculation methods) and traditional loggers. By comparison, monthly billing, daily interval, and hourly interval models only did so for 11 sites (34%). These spanned 5 of 11 building types—offices, auto repair shops, warehouses, assembly spaces, and big box stores—but probing for correlations between model performance and business type (or load shapes, occupancy patterns, etc.) would require larger sample sizes and/or additional metadata. An ongoing analysis of existing, and additional sites in 2018, will in part aim to determine whether the findings reported here carry over in a larger sample size.

Still, we can say that the feasibility of calculating savings from real-time or near real-time circuit-level interval data, and thereby performing M&V with less time spent installing and retrieving temporary equipment to monitor usage, depends upon a number of factors beyond their alignment with existing methods. For one, our investigation does not quantify the cost-effectiveness of each compared method. The equipment currently used in networked metering is often more expensive than data loggers and other existing methods—especially when factoring in soft costs (i.e. setup, training), maintenance needs, and potential to reuse devices at multiple sites. Furthermore, relying on advanced metering may present issues stemming from the difficulty of disaggregating lighting loads from non-lighting loads at certain sites.

Collaboration between implementers and evaluators facilitated data validation and identification of missing data, which may help improve AMI implementation and models going forward—but a firmer accounting of the costs associated with deploying the networked equipment, beyond the scope of this paper, is needed to compare against data loggers and other existing approaches.

References

CalTRACK Technical Documentation. 2017. *CalTRACK Site-level Daily Weather Normalized, Metered Energy Savings Estimation*. http://docs.caltrack.org.

———. 2017. *CalTRACK Site-level Monthly Weather Normalized, Metered Energy Savings Estimation*. http://docs.caltrack.org.

Kleinhenz, et. al. 2013. "Comparison of Metering and Verification Methodologies of Compressed Air Systems for Utility-Based Energy-Efficiency Programs: A Case-Study," *ACEEE Summer Study on Energy Efficiency in Industry* 4:1-13. Washington, DC: ACEEE. http://aceee.org/files/proceedings/2013/data/papers/4_214.pdf.

National Renewable Energy Laboratory. 2013. *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*: *Commercial and Industrial Lighting Evaluation Protocol*. https://www.energy.gov/eere/about-us/ump-protocols.

New York State Joint Utilities. 2017. *New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs – Residential, Multi-Family, and Commercial/Industrial Measures*. Version 5. http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/72c23decff52920a85257f1100671bdd/$FILE/ATTV3NKB.pdf/TRM%20Version%205%20-%20January%202018.pdf.

Ricardo, et. al. 2018. "Time to Move On: An Examination of Metering Periods for Small Business Direct Install Participants," *ACEEE Summer Study on Energy Efficiency in Buildings.* Washington, DC: ACEEE. *In press*.

Richman, EE. 2012. *Standard Measurement and Verification Plan for Lighting Retrofit Projects for Buildings and Building Sites*. Washington: Pacific Northwest National Laboratory. PNNL-21983. https://www1.eere.energy.gov/buildings/publications/pdfs/alliances/lighting_measurement_evaluation_protocol.pdf.