LBNL -54835

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Independent Review of Estimated Load Reductions for PJM's Small Customer Load Response Pilot Project

Prepared for
PJM Interconnection, LLC

Principal Authors
G. Heffner, M. Moezzi and C. Goldman

Ernest Orlando Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 90R4000
Berkeley CA 94720-8136

Environmental Energy Technologies Division

June 2004

This work described in this paper was funded by the office of Electric Transmission and Distribution, Electric Markets Technical Assistance Program of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

LBNL-54835

Independent Review of Estimated Load Reductions for PJM's Small Customer Load Response Pilot Project

Prepared for
PJM Interconnection, LLC

Principal Authors

G. Heffner, M. Moezzi and C. Goldman

Ernest Orlando Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 90R4000
Berkeley CA 94720-8136

June 2004

This work described in this paper was funded by the office of Electric Transmission and Distribution, Electric Markets Technical Assistance Program of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Acknowledgements

We would like to thank Susan Covino (PJM) and Larry Mansueti (DOE) for their support of this project. This study was funded by the Office of Electric Transmission and Distribution, Electric Markets Technical Assistance Program of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Table of Contents

List of Figures and Tables

Acronyms and Abbreviations

ADDF	Adjusted Diversified Demand Factor
ALM	Active Load Management
AMR	Automatic Meter Reading
CSP	Curtailment Service Provider
DOE	U.S. Department of Energy
DR	Demand Response
ELRP	Economic and Emergency Load Response Programs
ETS	Electric Thermal Storage
EWH	Electric Water Heater
FERC	Federal Energy Regulatory Commission
HV	High-Voltage
ISO	Independent System Operator
LBNL	Lawrence Berkeley National Laboratory
M&V	Measurement and Verification
MV	Medium-Voltage
PURPA	Public Utility Regulatory Policy Act

Abstract

This study describes the results of a low-cost approach used to measure reported load reductions from a residential electric water heater (EWH) load control program operated as part of PJM Interconnection's Demand Response small customer pilot program. Lawrence Berkeley National Laboratory (LBNL) conducted this independent review of the engineering estimates for EWH load control reported by a Curtailment Service Provider (CSP) at PJM's request. LBNL employed low-cost measurement and verification (M&V) approaches that utilized existing interval metering equipment to monitor results for a series of load control tests. The CSP collected hourly load data for two substations and several hundred households over a six-week period in October and November 2003. During this time period, the CSP operated its electric water heater load control program during pre-specified test periods in the morning, afternoon and early evening. LBNL then analyzed substation and premise-level data from these tests in order to verify the diversified demand reductions claimed by the CSP for customers participating in the EWH load control program.

We found that the observed load reductions for the premise-level data aggregated over all households in the two participating electric cooperatives were, respectively, 40%-60% less and 3 % less-10% higher than the estimated diversified demand reduction values assumed by the CSP, depending on whether observed or normalized results are considered. We also analyzed sub-station level data and found that the observed load reductions during the test periods were significantly lower than expected, although confounding influences and operational problems significantly limit our ability to differentiate between control-related and non-control related differences in substation-level load shape data. The usefulness and accuracy of the results were hampered by operational problems encountered during the measurement period as well as insufficient number of load research grade interval meters at one cooperative. Given the larger sample size at one electric cooperative and more statistically-robust results, there is some basis to suggest that the Adjusted Diversified Demand Factor (ADDF) values used by the CSP somewhat over-state the actual load reductions. Given the results and limitations of the M&V approach as implemented, we suggest several options for PJM to consider: (1) require load aggregators participating in ISO DR programs to utilize formal PURPA-compliant load research samples in their M&V plans, and (2) continue developing lower cost M&V approaches for mass market load control programs that incorporate suggested improvements described in this study.

1. Introduction

ISOs that oversee and administer various wholesale electricity markets are attempting to ensure that these markets provide comparable opportunities for supply-side and demand-side resources to participate, consistent with FERC policy direction. PJM Interconnection began operating its Economic and Emergency Load Response Programs (ELRP) in June 2002. These programs allow for customer-based resources, including on-site generators and participant load reductions, to receive payments in exchange for measurable load reductions (during emergency events as declared by the system operator) or when market prices provide incentives for end-user participation. One issue that ISOs have had to address in implementing Demand Response programs are concerns that program rules and eligibility requirements arbitrarily exclude certain types of loads, such as small customers without integrated hourly metering, from participation. In order to better understand this issue and possible solutions to it, PJM has undertaken a two-year, 100 MW pilot program targeted at smaller customers.[1] In this pilot program, qualifying Curtailment Service Providers (CSP) operating in PJM's control area may receive payments for load curtailments as part of PJM's demand response program.

One component of this small customer pilot program seeks to identify and test new approaches for measuring and verifying (M&V) load reductions for customers that do not have interval meters. PJM approached the US Department of Energy (DOE) and Lawrence Berkeley National Laboratory (LBNL) regarding a collaborative effort in this area.[2] In this report, LBNL provides initial results of a novel and inexpensive approach to measuring and verifying reported demand reductions for participants in load control programs that do not have interval meters.

The generally accepted approach to measuring and verifying load impacts of non-interval metered customer load control programs is installation of a statistically-representative sample of load research recorders at the premises of program participants. This method is currently used to measure and verify the load impacts of PJM's Active Load Management (ALM) Program. This measurement and verification approach can be expensive and time-consuming, especially for load aggregators participating in pilot programs that may be short-lived.[3] LBNL worked closely with PJM and a CSP in the pilot program to test an alternative M&V approach that would not require investment in large numbers of load research recorders and could yield results without the need for extensive analysis over an entire summer or winter season.

The approach applied here provides a relatively low-cost, quick-turn-around "snapshot" of the aggregate impacts of a load control program. It does so by measuring the aggregate demand impact of load control at the Medium-Voltage (MV) substation transformer bank or the High-Voltage (HV) wholesale delivery point. In applying this approach it is critical to choose an MV network serving a customer population with a high saturation of program participants, so that the

[1] PJM Emergency Load Response Program, FERC Electric Tariff First Revised Sheet No. 256A.

[2] LBNL is already providing technical assistance to regional grid operators and state/federal policymakers on demand response policy, technology, and program design issues, notably in program evaluation and identification of emerging technologies and demand response strategies for both interval and non-interval metered customers.

[3] A PURPA-compliant load research survey for a small customer load control program comprising 50,000 participants can cost $50,000-$75,000, which may outweigh the potential benefits to the CSP.

aggregate effect of many small load impacts is observable.[4] In addition to the MV-level aggregate impact measurements, the premise-level interval metering capabilities of several electric cooperatives that are working with the CSP in this program were mobilized and a small (though not statistically-representative) number of interval meters were placed on the premises of program participants served by the selected MV networks. This provides a second independent source of data for measurement and verification.

LBNL considers this measurement and verification approach to be novel because it relies on multiple approaches, each using an independent method (engineering estimates, substation-level load data, and premise- or end-use level load research, and – possibly- time-series or other modeling techniques), to verify load reductions. Such an approach, although not based on a statistically representative sample, may yield results that provide a suitable basis for payments while costing much less than a PURPA-compliant load research study. The loss in statistical precision may be acceptable if the M&V approach is sufficient to build confidence among stakeholders and provides a suitable basis for settlement and program valuation.

2. Small Customer Pilot Program Characteristics

The CSP participating in the small customer pilot program operates an integrated load management system that serves the needs of rural electric cooperatives located in 5 PJM zones. This integrated system comprises approximately 45,000 load control switches, the vast majority of which are located on residential consumer end use devices, delivering an estimated 35 MW of load reduction in summer (50 MW in winter) through control of electric water heaters, water pumps and electric thermal storage space heaters. The CSP is also participating in PJM's pilot Demand Response program.

LBNL worked closely with the CSP and two of its client electric cooperatives to identify target substations suitable for measuring the aggregate impacts of load control and to develop a regimen of short-duration load control tests that did not interfere with ongoing system operations, were outside the busy summer and winter seasons, and minimized the intervening effect of seasonal commercial and agricultural loads on these predominantly rural residential networks.

The two substations selected are both important delivery points (and therefore metering points) for wholesale service from PJM-member load-serving entities. Table 1 provides summary information on the two substations for which data was collected, including (1) the customer mix and the number of customers with electric water heating (EWH) load control devices; (2) the total coincident peak demand at each substation, with the residential component broken out separately; (3) the number of premise-level load research monitoring devices that were utilized at each sub-station; and (4) the estimated load reduction per electric water heater control point for

[4] See "Distribution Substation Load Impacts of Residential Air Conditioner Load Control", Transactions on Power Apparatus and Systems (IEEE Power Engineering Society), Spring 1985 for an example of estimating impacts of load control programs using substation data.

each of the intended load control test times, as provided by the CSP and expressed as an Adjusted Diversified Demand Factor (ADDF).[5]

Table 1. Characteristics of Participating Electric Cooperatives

Rural Network	(1) Customers	(2) Peak Demand (kW)	(3) Premise-level Load Control Data	(4) Estimated ADDF (kW/participant)
Electric Coop # 1 • Total: • Residential: • EWH LC group	 4,400 4,350 631	 6,500 6,250	215 households with hourly interval AMR	Oct. 2-3 pm: -0.65 Oct. 3-4 pm: -0.675 Nov. 7-8 am: -0.85 Nov. 8-9 am: -0.925 Nov. 6-7 pm: -1.175 Nov. 7-8 pm: -1.175
Electric Coop # 2 • Total: • Residential: • EWH LC group	 895 895 243	1,990 1,990	9 households with GE TMR-92 15-minute interval meters	Oct. 2-3 pm: -0.8 Oct. 3-4 pm: -0.85 Nov. 7-8 am: -1.0 Nov. 8-9 am: -0.925 Nov. 6-7 pm: -1.05 Nov. 7-8 pm: -1.05

3. Measurement Approach

Synchronized load control testing for the two rural residential networks took place over a six-week period from early October to mid-November 2003. Three separate load control tests, defined by start time and duration, were dispatched in order to provide three "point estimates" of load reductions that could be compared to the input assumptions for electric water heater diversified demand factors (ADDF) used by the CSP in the load control strategy tables that determine the estimated load reductions. The load control tests were:

- Load Control Test A: Two hour control 2-4 pm, Tuesdays and Thursdays in October.
- Load Control Test B: Two hour control 7-9 am, Tuesdays and Thursdays in November
- Load Control Test C: Two hour control 6-8 pm, Tuesdays and Thursdays in November.

During this six week time period, each Wednesday was designated as a "baseline" day during which no load control was to be activated. This "test-baseline" approach allowed us to create several definitions of baseline (or comparison) loads against which to compare the loads for the load control test days.[6] The CSP and two cooperatives collected data for the entire test period, which was then sent to LBNL for reformatting and analysis.

Operating difficulties were encountered during the course of the tests. In particular, following completion of the tests, the CSP reported that the Electric Thermal Storage (ETS) load control program had also been inadvertently dispatched during many of the October and November load

[5] Adjusted Diversified Demand Factor (ADDF) is the estimated diversified demand reduction for load control of a typical electric water heater; it varies by hour, season and day type (weekend vs. weekday).
[6] A mid-week baseline is preferred for mid-week test days in order to avoid the "Monday-Friday" effect, when daily load curves are often shaped differently and often have lower magnitudes.

control test periods. This had a serious confounding effect on our ability to analyze the sub-station level data, as the estimated demand impact of the ETS program is **larger** than its sister Electric Water Heater (EWH) program for Electric Cooperative #1 (see Table 2). However, the estimated demand impact of the ETS program for Electric Cooperative #2's network is about 35% of the expected magnitude of the estimated EWH program impact. In analyzing sub-station level data, LBNL did not attempt to adjust the measured results of the EWH load control program to account for the effect of ETS operation, as we had no independent basis for measuring estimated ETS impacts.

Table 2. Impact of Operations Difficulties on Measurement and Verification Study

		October Tests		November Tests			
		2-3 pm	3-4 pm	7-8 am	8-9 am	6-7 pm	7-8 pm
Elec Coop	EWH Est. Load Impact (kWh)	410	426	536	584	741	741
# 1	ETS Est. Load Impact (kWh)	755	755	755	755	755	755
	ETS/EWH Rel. Size	1.84	1.77	1.41	1.29	1.02	1.02
Elec Coop	EWH Est. Load Impact (kWh)	194	207	243	225	255	255
# 2	ETS Est. Load Impact (kWh)	73	77	85	79	89	89
	ETS/EWH Rel .Size	0.37	0.37	0.35	0.35	0.35	0.35

A less serious problem was the operation of electric water heater load control during a few of the Wednesday periods set aside as "baseline" or benchmark days. This occurred only once, during the week of November 3, and for analysis purposes the adjacent Monday and Friday was used as the "baseline" day for load reduction measurements on Tuesday and Thursday, respectively, of that week.

4. Results

LBNL estimated the load reduction due to EWH load control by measuring the differences in hourly usage patterns between the "Load Control Test" days (Tuesdays and Thursdays) and the "Baseline" day (Wednesday). LBNL also applied a normalization technique to the raw observations in order to take into account potential differences between the Test and Baseline day (e.g., temperature differences and other non-random as well as "random" variation). The observed results together with the normalized results provide a good estimate of the effect of load control while attempting to allow for some intervening influences and effects (see Appendix A for description of the normalization method). Tables 3 and 4 summarize the results for Electric Cooperatives # 1 and 2, respectively. Each row represents an independent comparison of a Load Control Test Day [identified in columns (1) and (2)] and a Baseline Day.

Columns (3) through (8) provide results based on observations at the premise level. Column (3) shows the observed mean value across households of the difference (expressed in kW) between each pair of Load Control Test and Baseline Days for the specified load control test period. This observed mean difference is calculated for the entire population (N) of premise-level interval meters available at each electric cooperative. A negative value means the load values for the Load Control Test day are lower than the corresponding values for the Baseline day, thus signifying a load reduction.[7] Column (5) provides the corresponding "normalized" mean difference across the interval metered data. LBNL calculated a *p*-value (using a paired t-test) based on both the observed and the normalized differences in load between each test and baseline period for all of the individual premise-level data pairs.[8] All of the premise-level data comparisons considered to be statistically significant (*p*-value less than 0.1) are shown in italics.[9]

Column (7) provides the mean total daily difference in electricity usage (in kWh) between each pair of Load Control Test and Baseline Days. This value is the integral of the hourly difference between the two daily load shapes and provides a rough indicator of the presence of intervening effects (e.g., significant differences in temperature, usage patterns or other factors that may have occurred between the test and control day) that could skew or obscure the data pair comparisons.

Column (8) provides a convenient arithmetic aggregate impact of each EWH load control test based on the per-premise mean values [i.e., Column (3)] times the number of EWH program participants on each network: 631 customers for Cooperative # 1 and 243 customers for Cooperative # 2.

Column (9) shows the results for substation/MV network level data between test and baseline periods. This is provided as a simple mean difference between pairs of data points over fifteen-minute intervals contained within each of the load control test periods. A negative value means the load values for the Load Control Test day are lower than the corresponding values for the Baseline day. The values in Column (8) and Column (9) can be directly compared to see the differences in aggregate load reduction obtained from the two independent sources of data used in this study.

5. Discussion

5.1 Electric Cooperative # 1

5.1.1 Substation (MV) Level

We believe that results from the substation/MV network-level data for Electric Cooperative # 1 (e.g. Column 9 of Table 3) should be discarded, for the following reasons:

---

[7] The values in Column (3) can be compared directly to the per-unit electric water heater diversified demand factors for each of the load control periods of interest (see Table 1).

[8] The pair-wise t-test assesses differences between paired observations, via a one-sample t-test on differences computed for each pair. The estimated difference (i.e., estimated load reduction) is the result of simple averaging of the observed pair-wise differences. The p-value is the statistical significance of a t-test on the distribution of differences; a *p*-value of less than 0.1 means there is a 90% confidence that the magnitude and sign of the mean is not the result of a random distribution but is due to our experimental design.

[9] Though we think the t-tests are a good description of these differences, note that the formal requirements for the t-test are not met by this configuration of data, and the p-values and estimated differences should be taken as heuristic rather than as formal statistical estimates. For example, the pairs are not independent observations nor are the data necessarily normally distributed, so a non-parametric paired t-test may be more appropriate. Nor is the hour-to-hour dependence of load values accounted for by this method. In particular, the t-tests are not able to distinguish between control-related and non-control related differences in load shapes.

- The inadvertent operation of the ETS load control program significantly impacts our ability to assess the impact of the EWH load control, because the ETS load control is expected to yield twice as much load control as the EWH program. We do not believe that netting-out the estimated load of the ETS program is a sufficient remedy, as this simply raises the question of verifying estimated load impacts for another load control program.
- The relatively low saturation of 15% of EWH control switches coupled with the fact that there are ~50 non-residential customers on the MV system (see Table 1) of Electric Cooperative # 1 makes it more difficult to spot the impact of the residential load control test in the substation data.

5.1.2 Premise Level.

A primary reason that Electric Cooperative # 1 was included in this M&V study was that a large number of residences that participated in the EWH program also were fitted with AMR (Automatic Meter Reading) hourly interval meters. This large sample of 215 households with premise-level data ultimately produced the most statistically robust results, which were also not subject to the confounding effects of inadvertent ETS program operation.[10] Fifteen of 18 observed mean difference comparisons and 18 of 18 normalized difference comparisons were signed correctly (i.e., signifying a measured load reduction), and eleven or twelve mean difference comparisons were statistically significant at a 90% confidence interval (see Table 3), based on the observed and normalized mean difference comparisons, respectively. Figure 1 compares the average load reductions for the 215 premises with the estimated ADDF values for each comparison period: average load reductions in each comparison period that are statistically significant are shown with lighter shaded bars. Fig. 1 also provides a visual overview of the underlying variability in the observed usage data across comparison periods, as well as the effect that normalization has on this variability. Generally speaking, normalization had the effect of reducing the load impact values for each control strategy, especially in those cases where the observed impacts were significantly higher than the estimated (ADDF) values. However, in several cases where the observed values were incorrectly signed or very low, the normalization adjusted the impact value sufficiently closer to the estimated (ADDF) value to at least have the correct sign. The overall effects of normalization were to: (1) reduce the variability of the impact values; and (2) produce average load impacts that were even lower than the estimated (ADDF) values.

[10] LBNL has verified through the CSP that none of the 215 EWH participants whose premise level data are included in this analysis participated in the ETS load control program.

The average observed load reduction over the 18 control-versus-baseline comparison periods (see Table 5) is **-0.55 kW/participant**, with a standard error of 0.041 and a 90% confidence interval of (-0.618, -0.482), while the average normalized load reduction is **-0.41 kW.participant**, with a standard error of 0.05 and a 90% confidence interval of (-0.523, -0.372).[11] These values are significantly lower than the average ADDF values estimated by the Curtailment Service Provider.

11 The standard error statistically describes the variability of a computed value, in this case, the computed mean difference between interval meter readings during the load control test period and the corresponding baseline period. It is computed as $(s^2/n)^{1/2}$ where s is the standard deviation (calculated as the sum of the square of the differences between each observation and the mean of all the observations), n is the product of the sample size and the number of observations (18), and s^2 is the statistical variance across the individual differences. Thus the standard error increases as variance increases and as sample size decreases. The standard error can be used to compute a confidence interval for the mean difference, which describes the range within which the "true" value of the difference is expected to fall. A 90% confidence interval for the mean, for example, is the range within which we would expect the computed mean to fall 90% of the time, if the same experiment were repeated many times. The magnitude of the confidence interval will generally be larger depending on how high you set the confidence level, how large the standard error is, and the number of observations (sample size). For more discussion *see Theoretical Statistics*, D. R. Cox, D.R. and D.V. Hinkley, 1974, published by Chapman & Hall.

Table 3: Load Control Test Results for Electric Cooperative #1

Premise-level Data (N = 215):								Substation-level Data:
1	2	3	4	5	6	7	8	9
Date	**Test Period**	**Observed Average Load Reduction (kW)**	**Observed *p*-value**	**"Normalized" Average Load Reduction (kW)**	**"Normalized" *p*-value**	**Observed Difference in Total Daily Consumption (kWh)**	**Estimated Aggregate EWH Load Reduction (kW) [n= 631 houses]**	**Observed Difference in Substation Load (kW)**
10/7/2003	2pm-4pm	-0.16	0.23	-0.21	0.11	**1.59**	-103	366
10/9/2003	2pm-4pm	*-0.56*	*<0.001*	*-0.46*	*<0.001*	-3.22	-352	442
10/14/2003	2pm-4pm	-0.17	0.14	-0.16	0.22	-2.39	-110	-312
10/16/2003	2pm-4pm	0.16	0.23	-0.06	0.69	**2.60**	99	45
10/21/2003	2pm-4pm	*-1.08*	*<0.001*	*-0.69*	*<0.0001*	**8.22**	-681	-2143
10/23/2003	2pm-4pm	*-2.61*	*<0.001*	*-1.37*	*<0.0001*	-9.48	-1648	965
11/4/2003	7am-9am	*-1.09*	*-0.04*	*-0.75*	*<0.0001*	-3.01	-685	-1071
11/4/2003	6pm-8pm	*-0.48*	*0.01*	*-0.43*	*0.01*	-3.01	-300	828
11/6/2003	7am-9am	*-0.93*	*<0.001*	*-0.58*	*0.00*	-0.58	-585	-2109
11/6/2003	6pm-8pm	*-0.31*	*0.06*	-0.03	0.88	-0.58	-196	-112
11/11/2003	7am-9am	-0.17	0.35	*-0.49*	*0.01*	**7.25**	-107	1135
11/11/2003	6pm-8pm	*-0.32*	*0.10*	*-0.40*	*0.03*	**7.25**	-203	-907
11/13/2003	7am-9am	*-1.02*	*<0.001*	*-0.87*	*<0.0001*	-1.10	-644	-168
11/13/2003	6pm-8pm	*-0.53*	*0.01*	*-0.59*	*0.00*	-1.10	-337	-199
11/18/2003	7am-9am	-0.16	0.36	-0.27	0.12	**4.62**	-101	603
11/18/2003	6pm-8pm	0.00	1.00	*-0.39*	*0.06*	**4.62**	1	1800
11/20/2003	7am-9am	*-0.63*	*0.00*	*-0.51*	*0.00*	**2.54**	-394	234
11/20/2003	6pm-8pm	0.19	0.35	-0.19	0.36	**2.54**	123	-162

Observed Results: 15/18 Signed Correctly; Normalized Results: 18/18 signed correctly
Italics = significant at p=0.10 level (90% confidence)
Observed Results: 11/18 significant @ 90% confidence level; Normalized Results: 12/18 significant @ 90% confidence level
Bold=positive difference in daily sum kWh (i.e., test day is higher than baseline day)

Table 4: Load Control Test Results for Electric Cooperative #2

Premise-level Data (N = 9)								Substation level data
1	2	3	4	5	6	7	8	9
Date	**Test Period**	**Observed Average Load Reduction (kW)**	**Observed *p*-value**	**"Normalized" Average Load Reduction (kW)**	**"Normalized" *p*-value**	**Observed Difference in Total Daily Consumption (kWh)**	**Estimated Aggregate EWH Load Reduction (kW) (n =243 houses)**	**Observed Difference in Substation Load (kW)**
10/7/2003	2pm-4pm	0.14	0.78	-0.45	0.38	0.61	34	-41
10/9/2003	2pm-4pm	-0.42	0.47	-0.37	0.46	0.09	-102	-84
10/14/2003	2pm-4pm	0.19	0.55	0.42	0.17	0.25	47	-70
10/16/2003	2pm-4pm	0.14	0.81	-0.03	0.96	**1.37**	34	-92
10/21/2003	2pm-4pm	*-0.97*	*0.08*	-0.48	0.17	-1.41	-236	-212
10/23/2003	2pm-4pm	-0.33	0.63	-0.31	0.49	1.89	-80	85
11/4/2003	7am-9am	*-2.76*	*0.06*	-2.00	0.28	-0.04	-671	-188
11/4/2003	6pm-8pm	*-1.97*	*0.05*	*0.02*	0.98	-0.04	-479	-261
11/6/2003	7am-9am	*-3.73*	*0.05*	-2.22	0.14	**0.65**	-906	-228
11/6/2003	6pm-8pm	*-1.82*	*0.07*	-0.79	0.43	**0.65**	-442	-93
11/11/2003	7am-9am	-0.71	0.40	-1.54	0.14	**1.81**	-172	67
11/11/2003	6pm-8pm	-1.26	0.16	-1.20	0.22	**1.81**	-306	-27
11/13/2003	7am-9am	-1.88	0.13	-1.17	0.28	-3.36	-457	-144
11/13/2003	6pm-8pm	-0.92	0.25	-1.09	0.34	-3.36	-223	116
11/18/2003	7am-9am	0.13	0.90	*0.05*	0.96	**1.27**	30	57
11/18/2003	6pm-8pm	-0.75	0.40	-0.15	0.83	**1.27**	-182	-42
11/20/2003	7am-9am	-1.42	0.07	-0.94	0.18	-1.43	-345	-100
11/20/2003	6pm-8pm	*-1.22*	*0.06*	*-1.13*	*0.09*	-1.43	-296	-167

Observed Results: 14/18 signed correctly; Normalized Results: 15/18 signed correctly
Italics = significant at p=0.10 level (90% confidence)
Observed Results: 6/18 significant @ 90% confidence level; Normalized Results: 1/18 significant @90% confidence level
Bold=positive difference in daily sum kWh (i.e., test day is higher than baseline day)

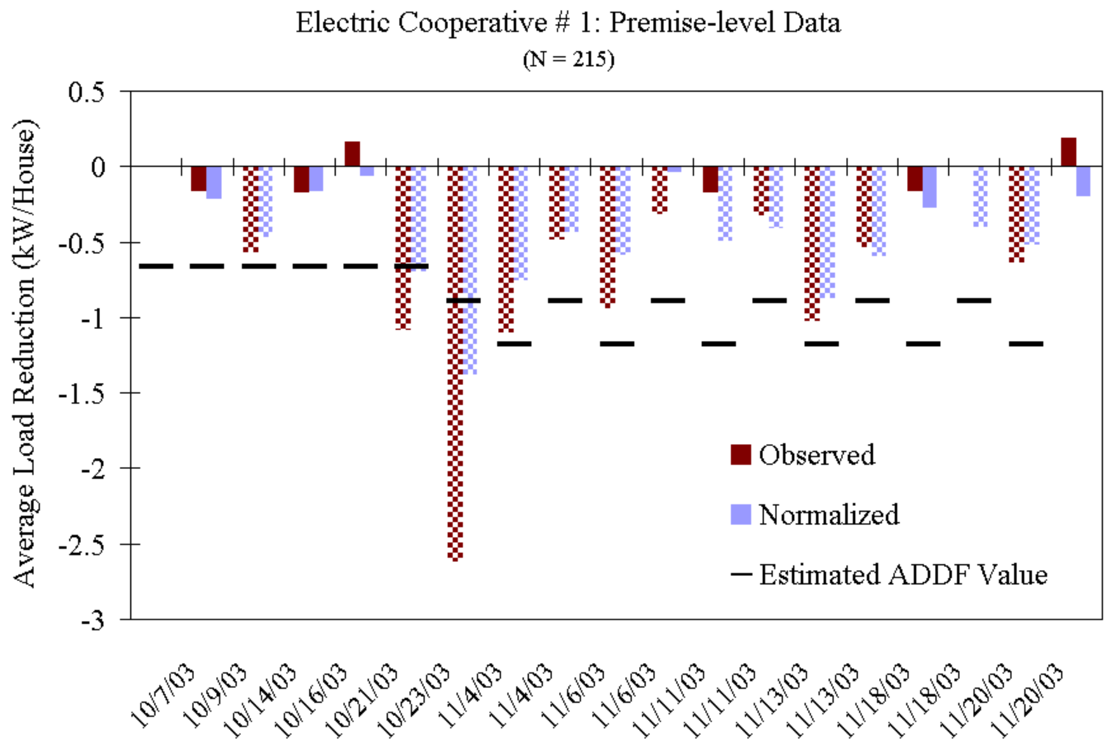

Figure 1: Observed Load Reduction Vs Estimated ADDF Value - Electric Cooperative # 1

Care should be taken in interpreting these results, because of the possibility of non-control related differences in load shapes between test and control days. For example, we observe relatively high mean daily kWh differences between test and control days for the period October 21-23 and to a lesser extent for the November 11 Veteran's Day holiday (see Column 5).[12]

Figure 2 shows a typical average load curve derived from the premise level data for Electric Cooperative # 1.

[12] A positive mean daily usage difference means that usage overall was higher for all customers on the test day vs. the comparison day. A negative mean daily usage difference means that electricity usage overall in all hours was lower on the test day vs. the comparison day. One possible method for accounting for these general daily differences is to normalize the loads between load control periods and comparison periods, e.g., according to the respective load values for the hour before the load control begins.

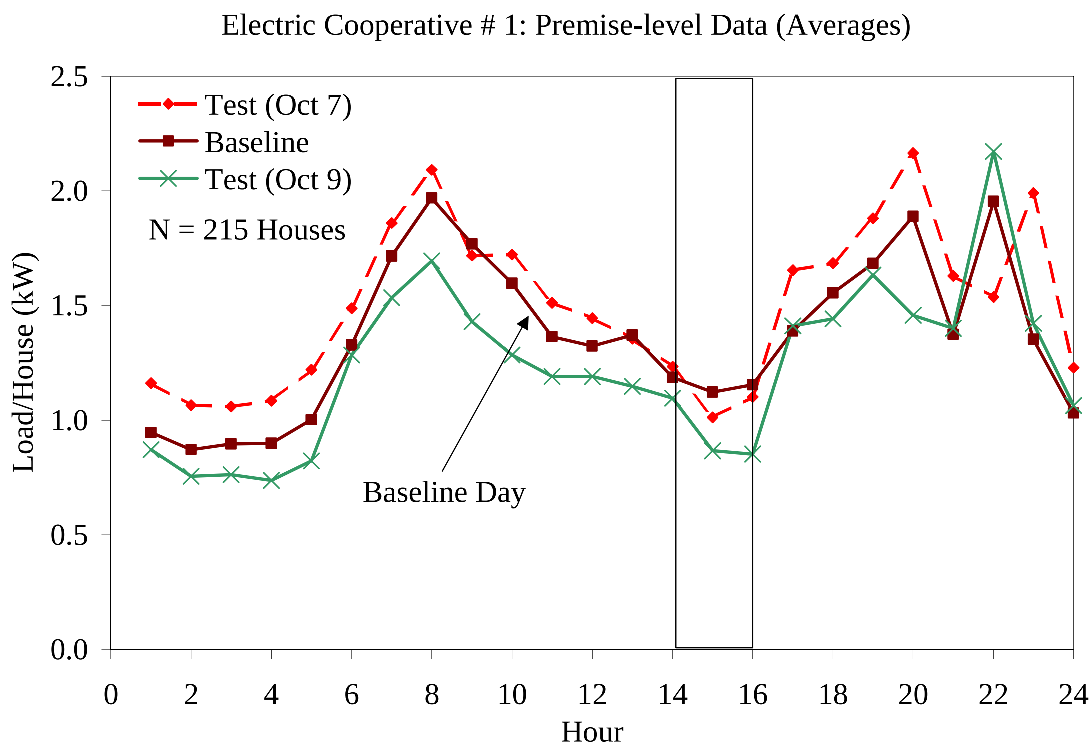

Figure 2: Typical Premise Level Load Shape - Electric Cooperative # 1

By comparing the two Test Days (October 7 and 9) with the Baseline Day (October 8) during the Load Control Test Period of 2-4 pm it is easy to see the characteristics of electric water heater load control – a sharp reduction in load at 2 pm followed by a "rebound effect" after control is released at 4 pm.

5.2 Electric Cooperative # 2

5.2.1 Substation (MV) Level

The substation/MV network served by Electric Cooperative # 2 is more conducive to comparing the results from premise-level and substation-level metering (see Table 4). This is because the network is small, entirely residential, and has a higher saturation of EWH load control devices (about 27% of all customers served were program participants). Consequently, it could be somewhat easier to "see" and measure load control impacts at this substation, because there are fewer confounding effects.

We found that the substation/MV level data from this cooperative is better behaved but the aggregate load impacts are difficult to track here as well. Of the eighteen simple mean difference comparisons that can be observed, only four are incorrectly signed (i.e., positive; see Table 4, Column 9). However, the magnitude of the aggregate load reduction impact is significantly lower than that estimated by the CSP. The average load reduction for the eighteen observations (regardless of sign) is -52.2 kW, which is significantly lower than the load reductions estimated by the CSP, which range between 200-250 kW for the aggregate substation impact **even before adjusting for inadvertent ETS program operation.** Adjusting for ETS operation on this substation would effectively negate any measured EWH load impact at the substation level.

In Figure 3 we can see both the promise and the difficulties encountered in considering the substation- or MV-level load data. During the week of November 10 we can clearly see the load shape impacts of electric water heater load control – rapidly reduced load at the time control is instituted followed by a rapid rebound effect after control is relinquished – only on the two test days of November 11 and 13. This is true for both the morning (7-9 am) and evening (6-8 pm) load control strategy tests. In contrast, the baseline load shape shows no load shape distortions during the control test periods (i.e., Nov. 12). However, using these data to estimate the absolute magnitude of the load impact due to load control is complicated by unknown and intervening variables that cause the test day load shapes to often be higher than the baseline day load shapes. This problem of incorrectly signed load impacts is greater for substation/MV-level data than for premise-level data, even without introducing the extra conundrum of other load control operations.

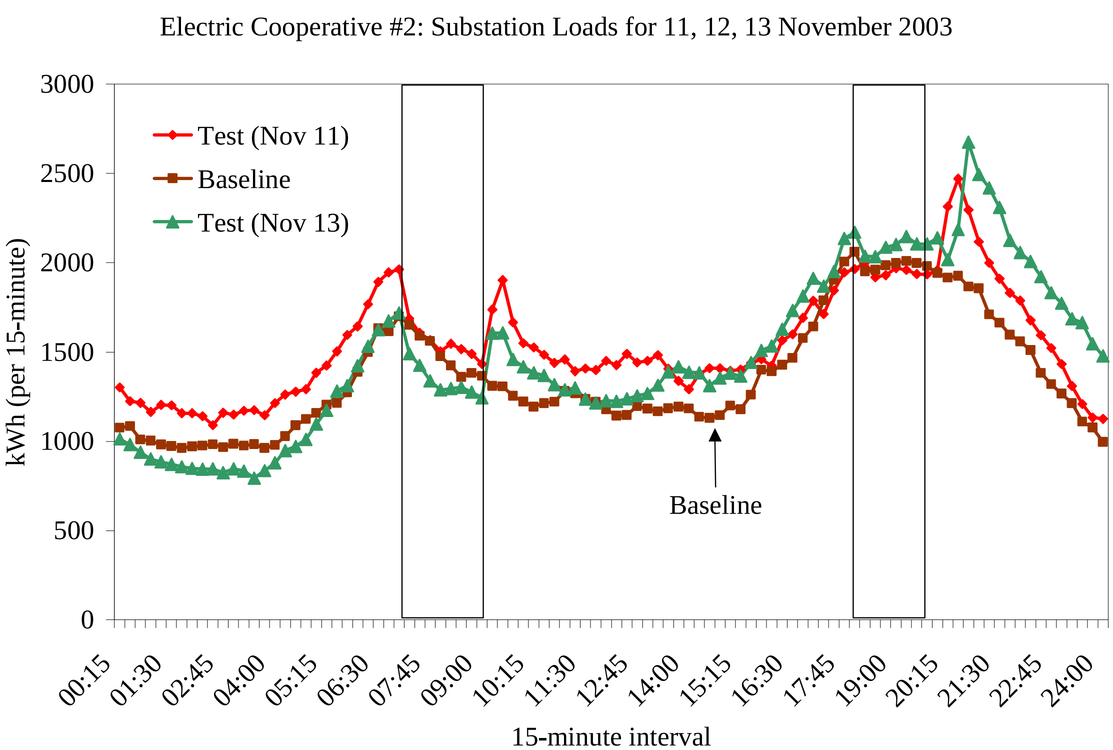

Figure 3: MV-Level Load Shape - November 11-13 Load Control Test on Electric Cooperative # 2

5.2.2 Premise Level

5.2.3 Only nine interval load recorders were available for placement on participating customers served by this substation. The observed and normalized premise-level results of the load control testing are shown in tabular format in Table 4 (Columns 3-6). Figure 4 compares the observed and normalized average load reductions for these nine premises with the estimated ADDF values for each comparison period; average load reductions in each comparison period that are statistically significant are shown with lighter shaded bars. Only six observed mean difference results and just one normalized mean different result was statistically significant, which is primarily a result of the small sample size. The observed mean difference results are signed correctly in 14 out of 18 load control tests, improving slightly to 15 out of 18 tests with the normalized results.[13] The normalization step does appear to modulate both the very low and very high individual load impact values. In fact, for Substation # 2 the normalized impact values are very close to the ADDF values, both for the individual load control strategies and for the overall average. The comparisons of load control test days with baseline days did not yield any evidence of significant unknown intervening variables, as indicated by the relatively low values of mean daily kWh difference (Column 7; Table 4). The average observed load reduction over the 18 control-versus-baseline comparison periods is **–1.08 kW/participant**, with a standard error of 0.195 and a 90% confidence interval of (-1.400775, -0.759225), while the average normalized load reduction is **-0.88 kW/participant,** with a standard error of 0.22 and a 90% confidence interval of (-1.26,-0.46). On average, the observed and normalized results as well as results for each of the three load control tests tend to support the ADDF values estimated by the CSP.

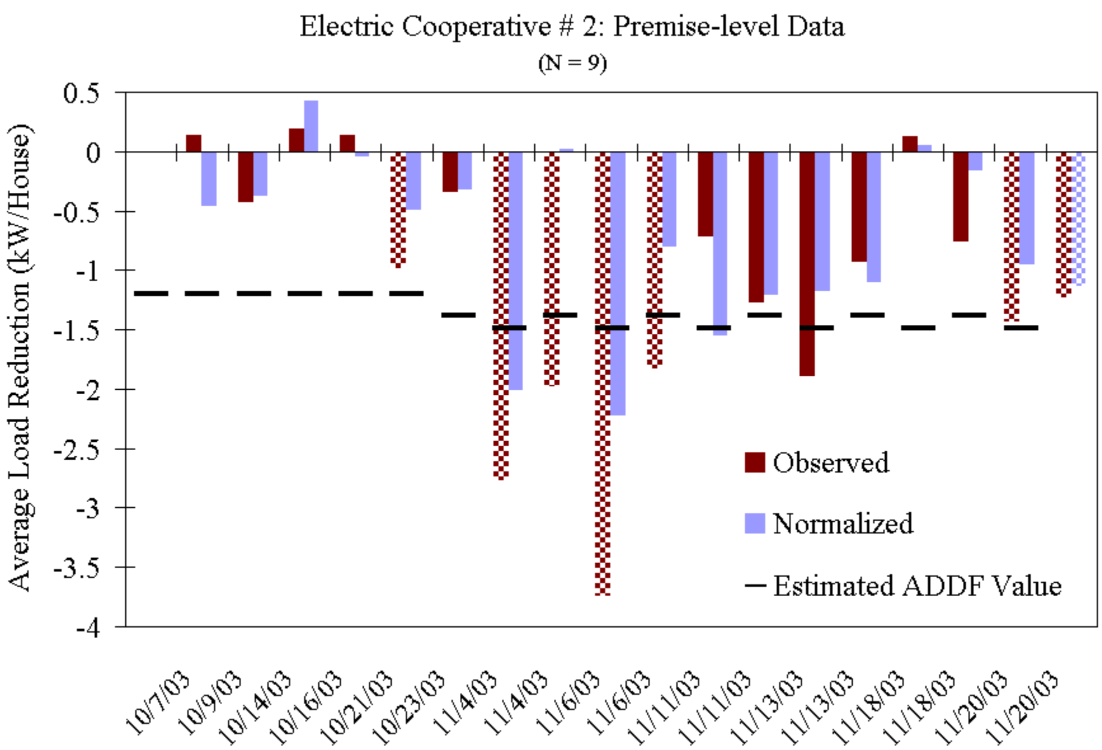

Figure 4: Observed Load Reduction Vs Estimated ADDF Value - Electric Cooperative # 2

[13] By "signed correctly" we mean the numerical result is consistent with the load-reducing effect expected.

Figure 5 provides the load shape for the week of November 10.

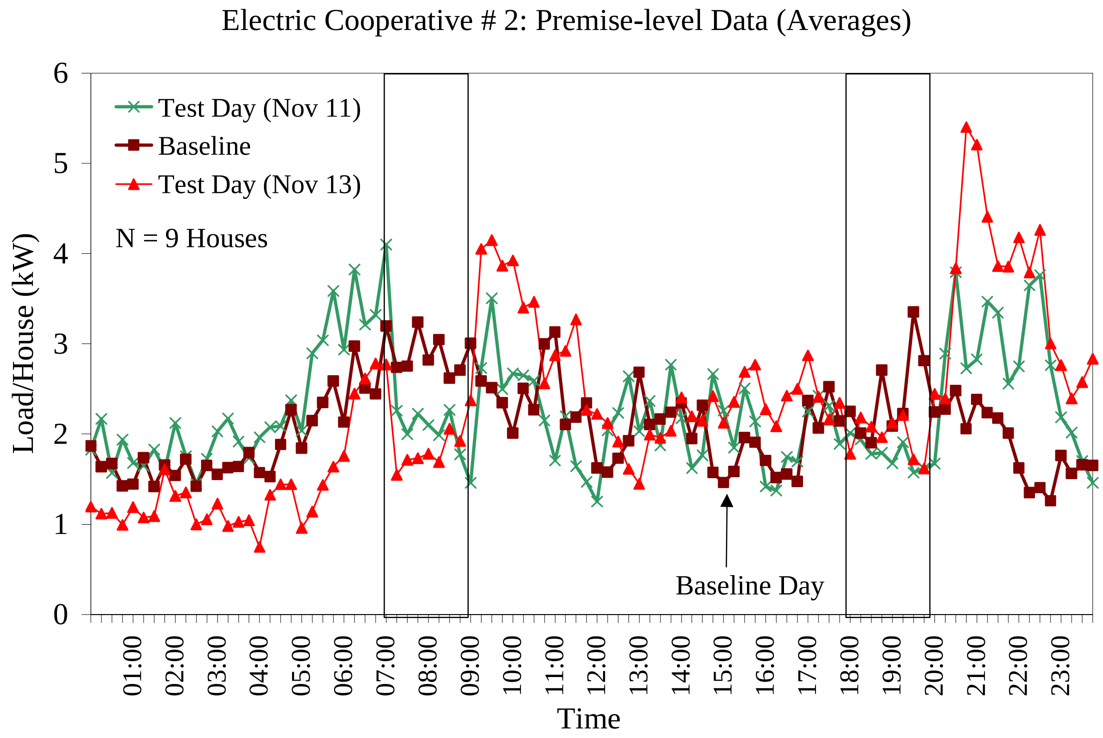

Figure 5: Premise-Level Load Shape for Electric Cooperative # 2

The contrast in behavior of the MV-level vs. premise-level data is apparent in comparing Figures 3 and 5. The premise level data for both load control test days (i.e., Nov. 11 and 13) and test periods (i.e., 7-9 am and 6-8 pm) is consistently below the corresponding baseline load shape, yielding much more consistent and hence credible load control results. The extent of the load reduction and load rebound effects (note especially the comparison between the Nov. 13 test day and the baseline) are also more apparent and thus much easier to quantify. In contrast, in analyzing the aggregate MV-substation level data, we observe that the difference in substation load is positive in the morning of Nov. 11 and evening of Nov. 13 (compared to the baseline period), a result that is inconsistent with the underlying load control strategies being implemented for that portion of the load that is "controlled".

We plan further analysis of these data that will allow some statistical inference into whether the observed differences between the Load Control Test and Baseline load shapes are correlated with the known test periods or whether the differences are randomly dispersed as in the other pairs of hours compared.

6. Conclusion

The network-level aggregate results for Electric Cooperative # 2 and premise-level customer results for both cooperatives yield three sets of results that vary considerably. The substation-level data for Electric Cooperative # 2 provide load reduction measures that are much lower than the ADDF values provided by the CSP, especially since LBNL has not adjusted for ETS operations. Using substation-level load reduction measurements alone would suggest that the ADDF values are uniformly too high by a factor of three (see Table 5).

The premise level data results suggest that observed load reductions — that is, the difference in loads between comparison day and test day during the test control period — are on average lower than the ADDF values, but not uniformly lower. Depending on the load control test period, the electric cooperative, and whether observed or normalized results are considered, the measured load reductions are either higher than, considerably lower than, or quite close to the ADDF values provided by the CSP. The overall load reduction measurement from premise level data averaged over all three load control tests is 40% lower than the average ADDF value for Electric Cooperative #1 (60% lower if we consider the normalized results) and 10% higher than the average ADDF value for Electric Cooperative # 2 (and spot on if we consider the normalized results).

There is some basis to suggest that the estimated ADDF values may be somewhat higher than the observed load reduction estimates, given differences in sample size and the statistical robustness of results (e.g., compare standard errors for the average measured load reductions in Table 5). The average observed load reduction for Electric Cooperative # 1 is accompanied by a relatively small standard error (0.041, or less than 10% of the average value for the observed results and about the same for the normalized results). By contrast, the average result for Electric Cooperative # 2 has a larger standard error (0.195, about 20% of the average value). Thus, we are relatively confident that the observed reductions in premise-level data are less than the ADDF estimates for Electric Cooperative #1 and comparable for Electric Cooperative #2. Given the significant disparity in results across the two cooperatives, LBNL believes that it would be difficult to justify a unilateral adjustment in CSP-provided values at the present time.

Table 5. Comparison of Estimated vs. Observed and Normalized Load Reduction Values

Load Control Test	Electric Cooperative # 1 – Per unit Load Reduction (kW/participant)			Electric Cooperative # 2 – Per unit Load Reduction (kW/participant)			
	ADDF Values reported by CSP	Observed Premise Level Data	Normalized Premise Level Data	ADDF Values reported by CSP	Substation Data	Observed Premise Level Data	Normalized Premise Level Data
A: Oct 2-4 pm	-0.663	-0.738	-0.391	-0.825	-0.284	-0.208	-0.62
B: Nov 7-9 am	-0.89	-0.665	-0.488	-0.965	-0.368	-1.73	-1.303
C: Nov 5-7 pm	-1.175	-0.24	-0.337	-1.05	-0.325	-1.32	-0.723
Average over all tests	-0.91	-0.55 Std Error: 0.041 90 % CI: (-0.618, -0.482)	-0.41 Std Error: 0.05 90% CI: (-0.523,-0.372)	-0.947	-0.326	-1.08 Std Error: 0.195 90% CI: (-1.40, -0.76)	-0.88 Std Error: 0.224 90% CI: (-1.26,-0.46)

LBNL would like to emphasize that these results are based on a novel measurement approach that has not been approved for measurement and verification applications, especially where financial settlements are involved. Additionally, the operational problems encountered in implementing the load control test program (e.g. ETS load control) means that sub-station level data results can not be relied upon. This work and the results obtained should instead be used to help both the CSP and PJM better assess how EWH load control programs that have operated for many years may provide customer load resources for wholesale electricity markets operated by ISOs. Based on our analysis, it is unclear whether there are any shortcuts to a full-fledged load research study as the basis for measuring and verifying the load impacts of non-interval metered customers.

7. Recommended Next Steps

We recommend PJM consider the following two options for how to proceed:

- Based on the shortfalls of this method in definitively verifying the CSP's load reduction assumptions, PJM could require use of a formal, load research-based approach to measurement and verification for load aggregators that participate in its Small Customer Pilot Program in the future.[14]
- Based on some promising results from this effort, and the prospect for improved results from a fine-tuned M&V approach, PJM could continue some experimentation with this and other new (and lower cost) measurement and verification approaches.

Should PJM wish to go forward with additional experimentation with the approach described in this study, we recommend the following specific improvements and next steps:

1. **Intervening Operations.** The present analysis was hampered by both inadvertent load control and normal CSP load control operations (a potential problem recognized during development of the measurement program) that interfered with the planned load control tests. In the case of Electric Cooperative # 1 these operations essentially washed out the effect of the electric water heater operations, and thus detailed analysis of the sub-station level data was abandoned. This problem could no doubt be averted with additional experience by the CSP and electric cooperatives in performing special load control tests such as called for in this type of measurement approach. Any additional efforts to use the present measurement approach must include a scrupulous attention to controlling and documenting all load control operations during the period when data is collected.

2. **More Observations.** The load control testing regimen was kept to a minimum to avoid placing undue burdens on either the CSP or electric cooperative staff. This resulted in only six load control test-baseline load shape "pairs" to estimate each of the three load control strategies, which analysis now shows to be insufficient – especially if any "pairs" are dropped for whatever reason. Any subsequent effort should be designed to provide additional control-baseline observations.

[14] The scope and requirements of such a formal approach are well documented, most recently in ISO-New England filings with the FERC on its small customer pilot program.

3. **Longer Duration Load Control Tests.** Another time-consuming aspect of assembling the data and conducting the analysis was time stamping issues. Longer control durations – minimum three hours – would allow for these boundary conditions to average out and also provide a clearer "rebound effect" at the end of the longer control period.

4. **Additional 15-minute interval meter load recorders.** Electric Cooperative # 2 fielded nine 15-minute interval load recorders. While not statistically representative, it has been a very useful data source. Electric Cooperative # 1 provided interval metering on a much-larger number (215) of customers, and despite the longer recording intervals (one-hour instead of 15 minutes) provided enough data for worthwhile analysis. The most desirable experimental design would be a formal sample design designed to provide known precision and confidence intervals given the variability within the participant population served by each cooperative or each MV network.

5. **Additional post-test analysis, including rudimentary model development.** The variability of hourly loads between test and baseline days is the real challenge to this type of analysis. Some (but not nearly all) of this variability should be possible to control for with a model-based approach incorporating key variables influencing day-to-day load levels, notably temperature, day of the week, and a "community activity" index of some type. One promising alternative would be to normalize baseline-test comparisons for a load control period on the difference (or other index) of the difference between baseline and test load for the hour immediately before the load control period. This empirical approach would be an alternative to a more complicated model-based approach to controlling variability.

References

Heffner, G."Distribution Substation Load Impacts of Residential Air Conditioner Load Control", *Transactions on Power Apparatus and Systems (*IEEE Power Engineering Society), Spring 1985.

Goldberg, M.L. and G.K. Agnew. "Protocol Development for Demand Response Calculation – Findings and Recommendations," KEMA-Xenergy for California Energy Commission, Report no. 400-02-017F, February 2003.

Appendix A: Normalization Methods for Treating Load Control Test Data

In analyzing load control impacts, we used normalization methods to "correct" for potential differences between the Test and Baseline days due to temperature differences and other non-random as well as "random" variations between load shapes.

Normalization is an alternative to calculating the load reduction as a simple difference between the load on test day and load on baseline day. Normalizing the load shape is a straightforward process of shifting the load shape in the hour just before the test period so that the test day and baseline day load shapes coincide exactly in that hour. This simple approach can potentially reduce the bias and variability in the calculation of load reduction without having to resort to complex models for adjusting for the impact of weather or other variables.[15]

The exact adjustment differs between load control tests. In the case of the November tests, where two tests were conducted per day, the load curve is shifted differently depending on whether we consider the morning or the afternoon test; and likewise the Wednesday baseline day is shifted differently for Tuesday than for Thursday load control tests.[16]

The normalization approach comprises the following steps:

1. Average load during test-period is calculated as,
   $\text{Load}_{\text{Test-day, Test-period}} = \text{Average}(\text{Load}_{\text{Test-day, Hour 1}}, \text{Load}_{\text{Test-day, Hour 2}})$
   $\text{Load}_{\text{Baseline-day, Test-period}} = \text{Average}(\text{Load}_{\text{Baseline-day, Hour 1}}, \text{Load}_{\text{Baseline-day, Hour 2}})$

2. Then the non-normalized substation level difference is given as,
   $\text{Non-normalized difference} = \text{Load}_{\text{Test-day, Test-period}} - \text{Load}_{\text{Baseline-day, Test-period}}$

3. The normalization factor is given by,
   $\text{Load}_{\text{Difference, Hour 0}} = \text{Load}_{\text{Baseline-day, Hour 0}} - \text{Load}_{\text{Test-day, Hour 0}}$

4. The normalized substation level difference is given as,
   $\text{Normalized difference} = \text{Non-normalized difference} - \text{Load}_{\text{Difference, Hour 0}}$

where,
If the Test-period is considered to be the period from 2 pm to 4 pm, then
Hour 0 = Period from 1 pm to 2 pm (or the hour just before the test-period),
Hour 1 = Period from 2 pm to 3 pm, and
Hour 2 = Period from 3 pm to 4 pm.

15 Goldberg, M.L., and G.K. Agnew (2003). Protocol Development for Demand Response Calculation – Findings and Recommendations, Prepared by KEMA-XENERGY for California Energy Commission, Report no. 400-02-017F, February 2003.

16 There are other ways of normalizing. The method used here was worked out in coordination with the Curtailment Service Provider and is particularly suitable for cases where weather is the largest intervening variable, as opposed to customer gaming or other influences.

If the load in the period from 1 pm to 2 pm matched perfectly between Wednesday and Tuesday then, the Normalized difference is equal to the Non-normalized difference.

The normalization method described above can be applied to either the premise-level (meter) data or the substation data. Given the confounding issues of unscheduled control program operations, we have only applied the extra normalization step to the premise-level data.