Innovative Approaches to Verifying Demand Response of Water Heater Load Control

Grayson C. Heffner, *Senior Member, IEEE*, Charles A. Goldman, and Mithra M. Moezzi

***Abstract*—This study describes a pilot effort to measure load reductions from a residential electric water heater (EWH) load control program using low-cost statistically based measurement and verification (M&V) approaches. This field experiment is described within the larger framework of overcoming barriers to participation of noninterval metered customers in Demand Response (DR) Programs. We worked with PJM Interconnection and a Curtailment Service Provider (CSP) to collect hourly load data for two substations and several hundred households over six weeks of load control testing. The experimental design reflected constraints imposed by limited funding, manpower, equipment, and the routine operation of the load control system by the CSP. We analyzed substation- and premise-level data from these tests in an attempt to verify several "point estimates" taken from the hourly diversified demand curves used by the CSP to establish aggregate load reductions from their control program. Analysis of premise-level data allowed for provisional verification that the actual electric water heater load control impacts were within a -60 to $+10\%$ band of the estimated values. For sub-station level data, measured values of per-unit load impacts were generally lower than the CSP estimated values for Electric Cooperative #2, after accounting for confounding influences and operational test problems. Based on this experience we offer recommendations to ISO and utility DR program managers to consider before undertaking further development of alternatives to the conventional but costly program-wide load research approach.**

***Index Terms*—Demand management, electricity markets, load control, load management, load shedding, statistics, substation measurements, water heating.**

I. Nomenclature

ADDF	Adjusted Diversified Demand Factor.
ALM	Active Load Management.
AMR	Automatic Meter Reading.
CSP	Curtailment Service Provider.
DR	Demand Response.
ELRP	Economic and Emergency Load Response Programs.
ETS	Electric Thermal Storage.
EWH	Electric Water Heater.
FERC	Federal Energy Regulatory Commission.
ISO	Independent System Operator.
M&V	Measurement and Verification.
PURPA	Public Utility Regulatory Policy Act.
SCADA	Substation Control and Data Acquisition

Manuscript received September 8, 2004; revised January 6, 2005. This work was supported by the Assistant Secretary of Electric Transmission and Distribution, Electric Markets Technical Assistance Program of the U.S. Department of Energy, under Contract DE-AC03-76SF00098. Paper no. TPWRD-00413-2004.

G. C. Heffner is with the Global Energy Associates, Inc, North Potomac, MD 20878 USA (e-mail: gcheffner@lbl.gov).

C. A. Goldman and M. M. Moezzi are with the Lawrence Berkeley National Laboratory, Berkeley, CA 94720,USA (e-mail: cagoldman@lbl.gov; mmmoezzi@lbl.gov).

Digital Object Identifier 10.1109/TPWRD.2005.852374

II. Introduction

ISOs that oversee and administer wholesale electricity markets are making efforts to ensure that these markets provide comparable opportunities for participation of both supply-side and demand-side resources, consistent with FERC policy direction. Demand Response (DR) programs allow for customer-operated resources, including on-site generators and end-use loads, to offer and receive payments for measurable reductions in their power demands during emergency events or in response to high real-time or forward market prices.[1]

Experience thus far has shown that larger and interval-metered customers can effectively participate in ISO-operated emergency, real-time, and day-ahead wholesale energy markets [1]. Such mobilization of end-use loads in response to price or system emergency signals can be shown to have very significant benefits both in terms of system reliability and dampening of volatile market clearing prices [2].

A difficult technical and program design issue is posed by small customers without interval meters, who are often excluded from participation by requirements for measurement and verification (M&V) of their hourly load reductions. FERC has encouraged ISOs to take steps to develop solutions to this problem, and PJM Interconnection has responded with a pilot DR program targeted at smaller, noninterval metered customers [3].

ISO-New England and New York ISO have adopted measurement and verification protocols that allow statistical sampling of small customer loads as a substitute for direct interval metering for certain DR programs [4]. However, the guidelines for statistically based M&V remain general, and there are very few documented examples of its successful application for ISO-based small customer DR programs [5]. The participation barriers faced by small, noninterval metered customers are likely to persist until there are standardized M&V protocols in place that are acceptable to ISO administrators (and system dispatchers) and have been empirically validated.

One component of PJM's small customer pilot program seeks to identify and test new approaches for M&V of these customers. PJM approached the US Department of Energy (DOE) and Lawrence Berkeley National Laboratory (LBNL) regarding a collaborative effort in this area. In this paper, we present an overall perspective on measuring and verifying demand reductions for noninterval metered customers together

[1]Regional transmission organizations offering DR programs include ISO-New England, New York ISO, PJM Interconnection, ERCOT, and CAISO.

with initial results for one (of many possible) novel approaches to the problem.

III. M&V Methods for Small Customer Load Control

Estimating the impacts of residential load control has been of interest to utility planners and engineers since it became a common industry practice [6]. Much effort has been devoted to thoroughly understanding the natural diversity of end-use loads and their behavior under various load control regimens (7, 8, 9). Much of this foundational work was accomplished using load research methods, in particular end-use load research, and standard experimental frames including comparison of average load curves for a population on "control" (test) and "normal" (non-control) days [10], [13], [15]. In addition to premise- or appliance-level load measurement, some analysts have estimated the impacts of end-use load control at more aggregate levels, notably feeder circuits [11], substations [12], [13], municipalities [14], and the utility system [15], [16]. These studies often used "notch" tests, "nick" tests, or "SCRAM" tests that are designed to elicit the maximum possible instantaneous demand drop from end-use load control strategies.

Each of these measurement approaches has potential drawbacks and implementation requirements that creates challenges in different settings. For example, [13] notes the difficulty in performing side-by-side comparisons of circuit-level data for "test" and "baseline" days, especially for temperature-sensitive loads, as the circuit-level loads on a given day can be affected by the temperature trends of several preceding days. Additionally, Reed *et al.* [11] discuss the need for frequent data scanning (short averaging intervals), as low as 50 scans per second, when attempting to detect the impacts of load control at the circuit level. Several analysts suggest the simultaneous application of load impact measurement using different methods applied at different estimation levels [11], [12], [15], as is attempted here.

Table I provides a framework for considering conventional and alternative M&V approaches for DR programs targeting smaller customers without interval meters. Table I highlights unconventional sampling or experimental design strategies (shown in italics) that might be utilized given the availability of data and the level at which impacts are being estimated (e.g., end use, premise, utility system). For example, the conventional approach used to measure and verify load impacts of noninterval metered customers in DR programs typically relies on a sample of interval-metered end-use loads or premises. Average load reductions are calculated for the sample; results are then used to represent the entire program population (see Table I—Grid Locations A4 and B4 [1], [4], [5], [15].[2] A drawback of this conventional M&V approach is its costliness and the time required for implementation, factors which are particularly limiting for load aggregators participating in pilot programs that may be short-lived or whose economics are sensitive to fixed costs.[3]

TABLE I
M&V Protocol Framework for Customers Participating in DR Programs

Data Source \ Impact Estimation Level	1. End-use	2. Premise/ Facility	3. Substation	4. Program/ System
A. End-use **Conventional**	**Interval Metering [7, 8, 9, 10]**			**End-use Load Research Sample [15]**
Unconventional			*End-use Load Research Sample*	*Extrapolate from substation results*
B. Premise/ Facility **Conventional**		**Interval Metering**		**Class Load Research Sample [1, 4, 5]**
Unconventional			*Class Load Research Sample*	*Extrapolate from substation results*
C. Substation **Conventional**			**SCADA [11, 12, 13, 14]**	
Unconventional			*SCADA*	*Extrapolate from substation data [12]*
D. Program/ System *Unconventional*				*System Data [16]*

Note: **Bold lettering signifies the "conventional" M&V Approach;** *italic lettering signifies unconventional and possibly innovative approaches.*

This framework of possible M&V protocols suggests there may be opportunities for cost-saving innovation in M&V approaches. Several have been tried, notably extrapolation of interval metering results at substation [Grid Location C4] and system levels [Grid Location D4]. The work reported here explores two unconventional and possibly innovative M&V approaches: interval metering of a sample of premises drawn at the substation level and used to represent the entire program population [Grid Location B4]; and substation-level interval metering used to measure the average per-unit diversified demand of the entire program [Grid Location C4].

The venue for this M&V experimentation was a PJM-sponsored Small Customer Pilot Program. The objective was to explore whether unconventional and less costly M&V approaches implemented at the substation level might provide lower-cost, quicker-turn-around estimates of the diversified demand impacts of a water heater load control program. The experimental M&V approach utilized interval metering to simultaneously measure aggregate load impacts at the Medium-Voltage (MV) substation transformer bank and per-end use load impacts at the premise-level.[4] The intent was to develop some field experience with these two unconventional but promising approaches to

[2] The extrapolation algorithm depends on sample design, especially whether the sample was stratified according to a variables thought to significantly influence load impacts (e.g., kWh monthly usage).

[3] A PURPA-compliant load research survey for a small customer load control program comprising 50 000 participants can cost $50 000–$75 000.

[4] As the paper describes, 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 aggregate effect of many small load impacts is observable.

measuring the load impacts of small customer DR programs, subject to the practical constraints of working with distribution utilities and load aggregators that have limited manpower and equipment resources and need to maintain normal operations of their load control system.

IV. Small Customer Pilot Program Characteristics

The CSP participating in PJM's small customer pilot program operates an integrated load management system that serves the needs of rural electric cooperatives located throughout the PJM control area. This integrated system comprises 45 000 load control switches and delivers an estimated 35 MW of load reduction in summer (50 MW in winter) through control of residential electric water heaters, water pumps, and electric thermal storage space heaters. Communications and dispatch is coordinated via a head-end computer system and an extensive telephone/radio communications system. The system has been in place for ~ 20 years and is operated on a daily basis to manage the daily maximum demand presented by rural cooperatives to wholesale suppliers.

LBNL worked 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 met the following operational criteria:

- did not interfere with ongoing system operations;
- were not in the busiest seasons (summer and winter); and
- minimized the potential intervening effects of seasonal nonresidential loads.

The two substations selected are both important delivery points (and therefore metering points) for wholesale service from PJM-member generators. Table II[5] provides summary information on the two substations, including (1) the customer mix and the number of households 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 substation; and (4) the CSP's estimated load reduction per electric water heater control point for each of the intended load control test periods, expressed as an Adjusted Diversified Demand Factor (ADDF).

The CSP uses these hourly ADDF values along with water heater address groupings stratified by size and customer usage characteristics to form load control strategy tables that locally optimize dispatch according to network needs, time of day, season, and other variables. The ADDF values together with the estimated restore demand function constitute the core assumptions underlying the estimated diversified water heater load control impacts, and were drawn from [9], [13] and [17].

[5]Note that the CSP-provided ADDF values can vary between substations for the same hourly period. This reflects adjustment by the CSP to reflect household demographics (EWH size and vintage, household size, etc.) particular to each cooperative.

TABLE II
Characteristics of 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 -Res. -EWH LC group	 4,400 4,350 631	 6,500 6,250	215 premises 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 -Res. -EWH LC group	 895 895 243	 1,990 1,990	9 premises 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

In addition to exploring applications of unconventional, lower cost M&V approaches, the specific experimental objective of this study was to measure and verify a selection of "point estimates" taken from the ADDF curve, shown in Fig. 1.[6] This choice of M&V approach—measuring individual point estimates from a daily seasonal end-use diversified demand curve—verifies the overall estimation approach without measuring a particular (peak) summer or winter hour.

V. Measurement Approach

The experimental design and schedule for the load control tests was developed within the operational constraints imposed by the CSP. A six week period from early October to mid-November 2003 was selected to avoid the summer and winter peak demands and take advantage of utility staff that were available during the shoulder season. Three separate load control "notch" tests, defined by start time and duration, were undertaken to provide "point estimates" of load reductions for comparison to the CSP ADDF values for electric water heater (see Table II). The load control tests were:

[6]The ADDF values used by the CSP in formulating load control strategies applied Summer Season diversified demand data for May–October and Winter Season diversified demand data for November through April.

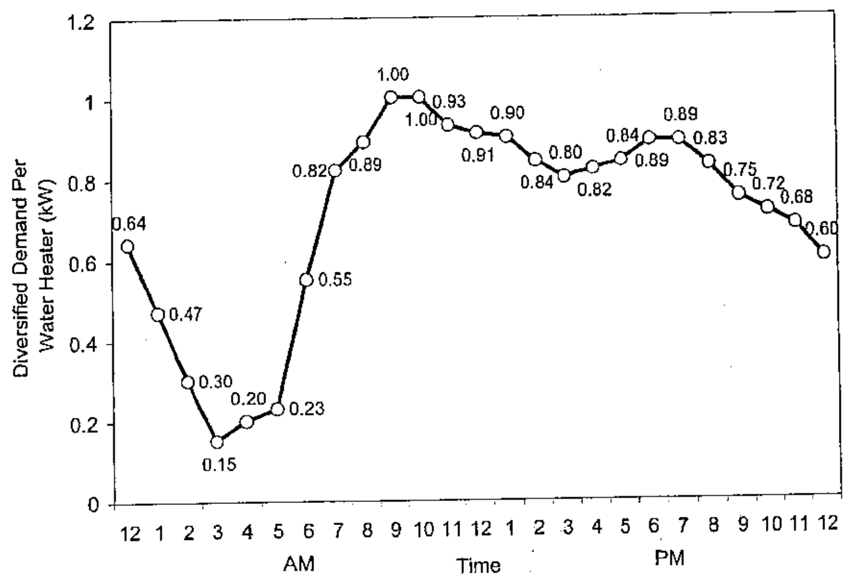

Fig. 1. Summer season hourly diversified demand curve.

- Test A: 2–4 pm, Tuesdays and Thursdays in October[7]
- Test B: 7–9 am, Tuesdays and Thursdays in November[8]
- Test C: 6–8 pm, Tuesdays and Thursdays in November.

Wednesdays were designated as "baseline" days, with no load control scheduled.[9] The CSP and two cooperatives collected data for the test period, which was forwarded to LBNL for analysis. Operating difficulties were encountered during the tests, as the CSP reported that parallel Electric Thermal Storage (ETS) load control and Voltage Control (VC) devices had been inadvertently dispatched during many of the EWH load control tests. This seriously confounded efforts to analyze the substation level data, as the estimated demand impact of the ETS and VC programs was comparable to or larger than the EWH program in some cases. Ultimately, LBNL was able to estimate load control impacts only for Electric Cooperative #2 after making some adjustments to the sub-station level data.

VI. Results

For the premise-level data LBNL estimated the load reduction due to EWH load control by measuring the differences in hourly usage patterns between the "Load Control Test" and "Baseline" days. For substation-level data, "Load Control Test" day data was examined to determine the demand drop at the time of control and the demand recovery when control was relinquished.

A. Premise-Level Results

LBNL applied a normalization technique to the raw premise-level observations in order to account for load shape or magnitude differences between the Test and Baseline day [6]. Normalizing 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. Normalization reduces the bias and variability in load reduction estimates, without resorting to complex models adjusting for weather or other variables, and helps adjust for any skew in Baseline Day load profile or load level. The observed and normalized results provide a good estimate of the effect of load control while attempting to account for some intervening influences and effects.

TABLE III
Load Control Test Results for Electric Cooperative #1

(1) Date – Test Type	(2) Observed Average Load Reduction (kW)	(3) Observed *p*-value	(4) "Normalized" Average Load Reduction (kW)	(5) "Normalized" *p*-value
10/7/2003 - A	-0.16	0.23	-0.21	0.11
10/9/2003 – A	-0.56	<0.001	-0.46	<0.001
10/14/2003 – A	-0.17	0.14	-0.16	0.22
10/16/2003 – A	0.16	0.23	-0.06	0.69
10/21/2003 – A	-1.08	<0.001	-0.69	<0.0001
10/23/2003 – A	-2.61	<0.001	-1.37	<0.0001
11/4/2003 - B	-1.09	-0.04	-0.75	<0.0001
11/4/2003 - C	-0.48	0.01	-0.43	0.01
11/6/2003 – B	-0.93	<0.001	-0.58	0.00
11/6/2003 – C	-0.31	0.06	-0.03	0.88
11/11/2003 – B	-0.17	0.35	-0.49	0.01
11/11/2003 – C	-0.32	0.10	-0.40	0.03
11/13/2003 – B	-1.02	<0.001	-0.87	<0.0001
11/13/2003 – C	-0.53	0.01	-0.59	0.00
11/18/2003 – B	-0.16	0.36	-0.27	0.12
11/18/2003 – C	0.00	1.00	-0.39	0.06
11/20/2003 – B	-0.63	0.00	-0.51	0.00
11/20/2003 - C	0.19	0.35	-0.19	0.36

Note: A = 2 pm – 4 pm; B = 7 am – 9 am; C = 6 pm – 8 pm

[7]October is considered part of the Summer Season

[8]November is considered part of the Winter Season

[9]Mondays and Fridays were avoided, as the pre- and post-weekend weekdays often have a different load shape and/or magnitude.

Tables III and IV summarize the premise-level results for Electric Cooperatives # 1 and 2, respectively. Each row represents a comparison of a Load Control Test Day and a Baseline Day in sequential order of testing. The first eight rows summarize measurement results for the 2–4 pm summer period. The subsequent twelve rows show measurement results for the 7–9 am and 6–8 pm winter period tests. Column (2) shows the observed mean kilowatt value of the difference between each pair of Load Control Test and Baseline Days and is calculated for the entire population (N) of premise-level interval meters available at each electric cooperative. A negative value signifies a load reduction. Column (4) provides the corresponding "normalized" mean difference for each test period. LBNL calculated a p-value (using a paired t-test) for both the observed and normalized differences in load between each test and baseline period (see Columns 3 and 5). Premise-level data comparisons considered statistically significant (p-value less than 0.1) are shown in italics.[10]

B. *Substation-Level Results*

Substation-level load impact analysis can be based on either a simple time series analysis of loads over the test period or by a comparison of load values between "baseline" and "test" days. Because of inadvertent intervening load control operations (e.g., ETS), we opted for the simplest possible analysis scheme—using the load values immediately preceding the load control as the baseline against which the effects of load control are measured. This simple method yields estimates of the instantaneous Diversified Demand Drop as well as the subsequent Demand Recovery, from which a measure of diversified duty cycle for the end use can be derived [12].

Interval (15 min.) metered data was available at the substation level for both Electric Cooperatives 1 and 2. We discarded results for Electric Cooperative # 1 because of the inadvertent operation of an electric thermal storage (ETS) space heater load block during all of the test periods which was in aggregate larger than the corresponding electric water heater load block (755 versus 631 units).

The magnitude of the confounding effect of other load control operations was much smaller at Electric Cooperative #2. At this coop, two load blocks besides the EWH load block were dispatched during the October and November load control tests—a load block comprising a 2.5% conservation voltage reduction (CVR), and a load block comprising 15 ETS units. We adjusted for the CVR operation by subtracting 2.5% of the substation voltage values at the start and the end of the tests from the calculations of Diversified Demand Drop and Demand Recovery. The ETS load block was small and considered negligible relative to the EWH load block (i.e., 15 ETS units versus 243 EWH units). We included the ETS units along with the EWH units in calculating per-unit demand reductions.[11] The results of this analysis of adjusted substation load data for Electric Cooperative # 2 are shown in Table V.

TABLE IV
LOAD CONTROL TEST RESULTS FOR ELECTRIC COOPERATIVE #2

(1) Date – Test Period	(2) Observed Average Load Reduction (kW)	(3) Observed p-value	(4) "Normalized" Average Load Reduction (kW)	(5) "Normalized" p-value
10/7/2003 - A	0.14	0.78	-0.45	0.38
10/9/2003 – A	-0.42	0.47	-0.37	0.46
10/14/2003 – A	0.19	0.55	0.42	0.17
10/16/2003 – A	0.14	0.81	-0.03	0.96
10/21/2003 – A	*-0.97*	*0.08*	-0.48	0.17
10/23/2003 – A	-0.33	0.63	-0.31	0.49
11/4/2003 – B	*-2.76*	*0.06*	-2.00	0.28
11/4/2003 - C	*-1.97*	*0.05*	*0.02*	0.98
11/6/2003 – B	-3.73	*0.05*	-2.22	0.14
11/6/2003 – C	*-1.82*	*0.07*	-0.79	0.43
11/11/2003 – B	-0.71	0.40	-1.54	0.14
11/11/2003 – C	-1.26	0.16	-1.20	0.22
11/13/2003 – B	-1.88	0.13	-1.17	0.28
11/13/2003 – C	-0.92	0.25	-1.09	0.34
11/18/2003 – B	0.13	0.90	*0.05*	0.96
11/18/2003 – C	-0.75	0.40	-0.15	0.83
11/20/2003 – B	-1.42	0.07	-0.94	0.18
11/20/2003 - C	*-1.22*	*0.06*	*-1.13*	*0.09*

Note: A = 2 pm – 4 pm; B = 7 am – 9 am; C = 6 pm – 8 pm

VII. DISCUSSION

A. *Electric Cooperative #1: Premise-Level Results*

Electric Cooperative # 1 was selected for load control testing because of the large number of EWH program participants (215) fitted with hourly interval meters. This large sample produced statistically robust results that were not tainted by the inadvertent operations problem, as none of the interval-metered EWH customers were part of the ETS load block.

As shown in Table III, 15 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 of twelve mean difference comparisons were statistically significant at a 90% confidence interval based on the normalized results. Fig. 2 compares the average load reductions with the CSP-estimated ADDF values for each period; average load impacts in each period that are statistically significant are shown with a dashed bar. Fig. 2 shows the underlying variability in the observed usage data across comparison periods, as well as the effect of normalization on this vari-

[10] The pairwise 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 pairwise 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. 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.

[11] Our rationale was that electric water heaters and thermal storage space heaters are intrinsically energy storage devices and have comparable connected loads.

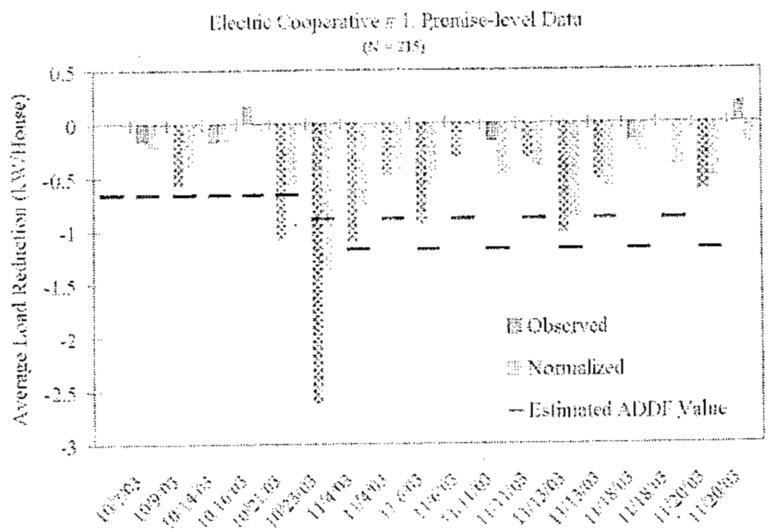

Fig. 2. Observed load reduction vs estimated ADDF value (electric cooperative # 1).

TABLE V
SUBSTATION LOAD CONTROL TEST RESULTS FOR ELECTRIC COOPERATIVE #2

Date	Test Type	Div Demand Drop (kW)	Recovery (kW)	Div Duty Cycle
7-Oct	A	77.7	184.9	0.42
9-Oct	A	58.6	286.9	0.20
14-Oct	A	101.2	341.3	0.30
16-Oct	A	135.7	371.9	0.36
21-Oct	A	93.9	294.2	0.32
23-Oct	A	84.9	405.9	0.21
Average 2-4 pm Summer		**92.0**	**314.2**	**0.29**
Unit Demand Drop:		**0.4**		
Unit Demand Recovery:			**1.2**	
4-Nov	B	254.9	341.7	0.75
6-Nov	B	313.9	418.3	0.75
11-Nov	B.	306.9	434.2	0.71
13-Nov	B	250.0	330.9	0.76
18-Nov	B	241.4	383.2	0.63
20-Nov	B	251.1	398.5	0.63
Average 7-9 am Winter		**269.7**	**384.5**	**0.70**
Unit Demand Drop:		**1.1**		**0.82**
Unit Demand Recovery:			**1.5**	
4-Nov	C	161.9	418.7	0.39
6-Nov	C	-3.7	460.7	N/A
11-Nov	C	-3.1	332.1	N/A
13-Nov	C	82.7	26.5	1.00
18-Nov	C	156.5	576.3	0.27
20-Nov	C	197.7	-7.4	N/A
Average 6-8 pm Winter		**98.7**	**301.1**	**0.33**
Unit Demand Drop		**0.4**		
Unit Demand Recovery			**1.2**	

ability. 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 load impact values, and 2) produce average load impacts that were lower than the estimated (ADDF) values. The average observed load reduction over the 18 control-versus-baseline comparison periods is -0.55 kW/participant with a standard error of 0.041, while the average normalized load reduction is -0.41 kW/participant, with a standard error of 0.05. These values are much lower than the ADDF values provided by the CSP.

B. Electric Cooperative #2: Premise-Level Results

Only nine interval load recorders were available for placement on EWH program participants served by this substation; none of these participated in the ETS program. The observed and normalized premise-level results of the load control testing are shown in tabular format in Table IV, while Fig. 3 compares the observed and normalized average load reductions for these nine premises with the estimated ADDF values. Only six observed load impact results and one normalized load impact result were statistically significant, which is primarily a result of the small sample size. The normalized mean difference results are signed correctly in 15 out of 18 load control tests. 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 estimated (ADDF) values, both for the individual load control strategies and for the overall average. The average observed load impact over the 18 control-versus-baseline comparison periods is -1.08 kW/participant, with a standard error of 0.195 while the average normalized load reduction is -0.88 kW/participant with a standard error of 0.22. On average, the observed and normalized results tend to support the ADDF values provided by the CSP for each load control period.

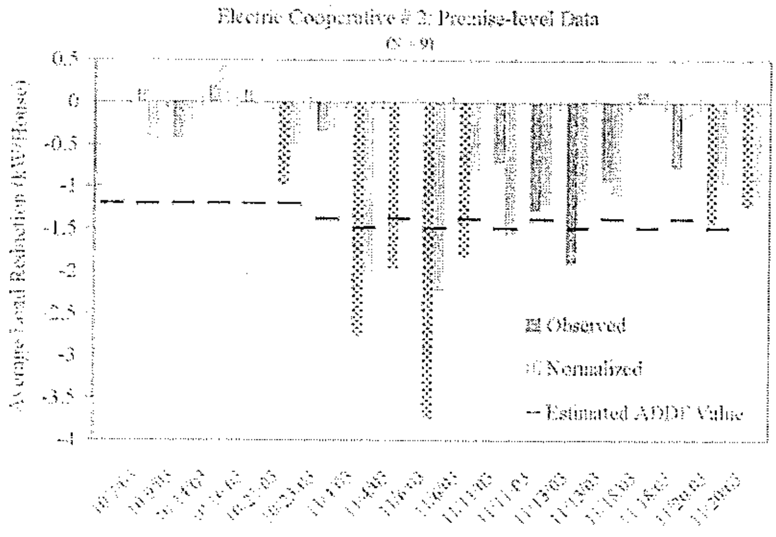

Fig. 3. Observed load reduction vs estimated ADDF value (electric cooperative # 2).

C. *Electric Cooperative #2: Substation-Level Results*

Table V summarizes the adjusted substation level measurements of Diversified Demand Drop for the 18 load control tests. Dividing the Demand Drop by the units controlled (including both EWH and ETS devices) yields a per-unit demand impact, which can be compared to the ADDF load curves (see Fig. 1) and the normalized premise-level mean difference results (see Fig. 3). Following Heffner and Kaufman [12], the Demand Recovery, or Rebound Effect, can be calculated over the period immediately before and after the time when control is relinquished. The quotient of these two terms provides a rough estimate of the diversified device duty cycle for that particular load control test period and is a proxy for the percentage of devices that are active over a given test period. For water heaters, this calculation of duty cycle will vary according to time of day, length of the control test period, and other variables, such as inlet water and possibly ambient temperature.[12]

Table VI compares the load impacts from the CSP's engineering estimates (ADDF) with normalized premise-level and adjusted substation level data for the three load control test strategies. The measured values of per-unit load impact from sub-station level data are generally lower than the normalized load impacts from premise-level data for Electric Cooperative #2, after adjusting for the intervening effects of CVR and ETS operations. The measured load impacts for Test Period B from the two methods are reasonably in line with one another. For Test Period A, the premise-level and substation-level results are both much lower than the ADDF value, with the substation level results less than half of the ADDF value. The sub-station level results for Test Period C are skewed by conditions in the evening hours of 4 of 6 November test days when either the Demand Drop, Demand Recovery, or Duty Cycle is signed incorrectly or cannot be calculated.

[12]Additional data from load control tests of differing duration and conducted during other times of the day would make it possible to fully characterize the aggregate characteristics of this EWH program, including hourly diversified demand and the net restore demand function [9].

TABLE VI
COMPARISON OF LOAD IMPACT MEASUREMENTS AT THE SUBSTATION AND PREMISE LEVEL FOR ELECTRIC COOPERATIVE #2 (IN kW PER UNIT)

	Test A (Summer 2 pm - 4 pm)	Test B (Winter 7 am - 9 am)	Test C (Winter 6 pm - 8 pm)
CSP-provided ADDF Values	-0.83	-0.97	-1.05
Coop # 2 Premise Level (normalized)	-0.62	-1.30	-0.72
Coop # 2 Substation Level – Demand Drop	-0.36	-1.05	-0.38
Coop # 2 Substation Level – Demand Recovery	1.22	1.49	1.17

D. *Possible Causes of Discrepancies Between Engineering Estimates (ADDF Values) and Measured Results*

There are a number of possible explanations for the discrepancies between the CSP's ADDF estimates of hourly diversified demand and the corresponding field measurements. First, since no end-use load meters were available, both the interval meter data from premises and substation are susceptible to intervening factors that affect the level and shape of the measured loads. These electric cooperatives serve rural households with other significant end-uses at both the premise and network levels whose on- and off-cycles could mask or distort the measured EWH load impact. Some measured diversified demand impacts are considerably closer to the ADDF estimates (see Test Period B, which are winter morning results) than others (see Test Period A which is the summer afternoon). These larger discrepancies may be caused by the relatively small number of load control test data points, especially when one or two tests are strongly affected by exogenous factors such as unseasonably cold or warm weather. Finally, there is the possibility that the engineering estimates are indeed overstated in at least a few cases. Other utilities use much lower diversified demand values for their summertime

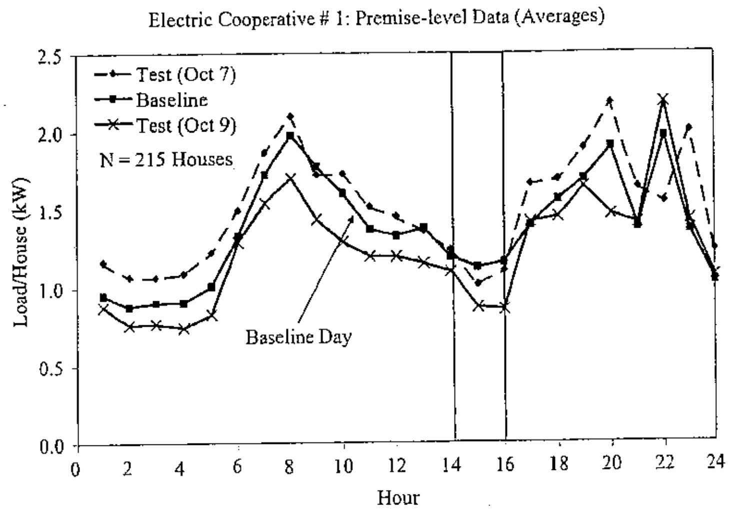

Fig. 4. Premise-level load shape—electric cooperative # 1.

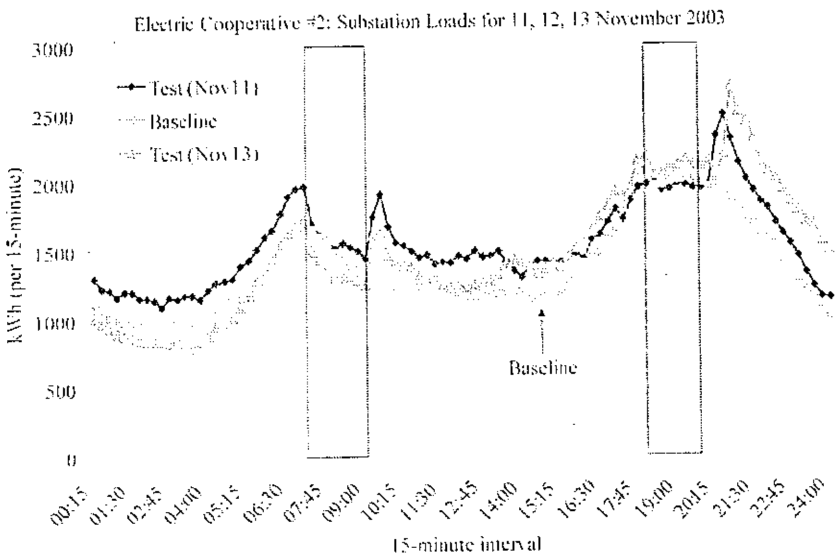

Fig. 5. MV-level load shape—electric cooperative # 2.

electric water heater loads, reflecting the warmer piped water supply during summer. In contrast, the Summer and Winter diversified demand curves used by the CSP are not very much different, based on an assumed prevalence of well water versus piped water supply in the rural areas served by Electric Cooperatives # 1 and 2.

E. *Qualitative Load Shape Analysis*

Fig. 4 shows a typical average load curve derived from the premise-level data for Electric Cooperative # 1. Comparing the load curves of the two Test Days (October 7 and 9) with the Baseline Day (October 8) 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.

Fig. 5 illustrates some of the practical issues associated with load impact measurements taken at the substation level. During the 7–9 am load control tests on November 11 and 13 we can clearly see the impacts of electric water heater load control—rapidly reduced load at the time control is instituted followed by a rebound effect after control is relinquished. This expected pattern is reflected in the values of Table V. The relationship between EWH/ETS control and substation load is much less clear for the evening load control test—a large rebound effect is apparent but there is little or no demand drop at the time that EWH/ETS control is initiated. Accounting for the intervening factors that distort or obscure the expected load curve behavior is very difficult short of extensive time-series analysis and multivariate regression analysis—both of which are outside the scope of this modest experimental effort.

VIII. Limitations of the Pilot M&V Approaches

The adjusted substation-level data from Coop # 2 yields load impact measurements much lower than the ADDF values provided by the CSP or the premise-level data, except for Test B (Winter Morning). Using substation-level load reduction measurements alone would suggest that the ADDF values are too

TABLE VII
ESTIMATED, 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	Obs. Premise Level Data	Norm. Premise Level Data	ADDF Values reported by CSP	Obs. Premise Level Data	Norm. Premise Level Data
A: Oct 2-4 pm	-0.66	-0.74	-0.39	-0.82	-0.21	-0.62
B: Nov 7-9 am	-0.89	-0.67	-0.49	-0.97	-1.73	-1.30
C: Nov 6-8 pm	-1.17	-0.24	-0.34	-1.05	-1.32	-0.72
Average over all tests	-0.91	-0.55	-0.41	-0.95	-1.08	-0.88
Std Error		0.04	0.05		0.19	0.22
90 % CI		(-0.62, -0.48)	(-0.52, -0.37)		(-1.40, -0.76)	(-1.26, -0.46)

high for Summer Afternoon and Winter Evening but about right for Winter Morning (see Table V).

The premise-level measurements indicate that load impacts are on average lower than the ADDF values, but not uniformly lower (see Table VII). The overall load impact measurement from normalized premise-level data is $\sim 60\%$ lower than the average ADDF value for Electric Cooperative #1, if we consider the normalized results, but is within 10% of the average ADDF value for Electric Coop # 2.

The M&V results reported here use a nonconventional approach that has not yet been approved or adopted for measurement and verification applications by PJM. A primary purpose of this study is to help the CSP, PJM, and other ISOs consider how "legacy" EWH load control programs (i.e., that have operated for many years but can't provide recent PURPA-compliant M&V studies) may provide verifiable, customer load resources for wholesale electricity markets in a manner acceptable to ISOs and other stakeholders. Based on our analysis, it is not yet clear whether there are any shortcuts to a conventional PURPA-compliant load research approach as the basis for measuring and verifying the load impacts of noninterval metered customers.

IX. Recommended Next Steps

Despite limitations on the usefulness of the pilot M&V results, they form a basis for considering which novel approaches may have relatively more promise, and the types of technical issues that should be accounted for in future efforts. Should PJM or another ISO wish to go forward with additional M&V experimentation along the lines described in this study, we recommend the following improvements and suggestions:

1) **Intervening Operations**: The present analysis was hampered by both inadvertent load control and normal CSP load control operations that interfered with the planned load control tests. This problem could be averted in future if the CSP set up and documented customized load control test strategies that would reside on the CSP's head end computer and could be implemented as required in order to verify load impacts.
2) **More Observations**: The load control test regimen was kept to a minimum to avoid placing undue burdens on either the CSP or electric cooperative staff. Any subsequent effort should be designed to provide additional control-baseline observations.
3) **Include Shorter and Longer Duration Load Control Tests**: A mix of load control test durations will allow for better characterization of the Demand Drop, Demand Recovery, and Diversified Duty Cycle for the EWH program on an aggregate and premise-level basis. The net restore demand function for the program could also be explicitly measured.
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. 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. Time-series analysis of substation level data is another promising possibility.

References

[1] B. Neenan *et al.*, "How and Why Customers Respond to Electricity Prices: A Study of NYISO and NYSERDA 2002 PRL Program Performance," Lawrence Berkeley Nat. Lab., Berkeley, CA, LBNL-52 209, 2003.

[2] J. W. Black, J. M. Ilic, and J. L. Watz, "Potential benefits of implementing load control," in *Proc. IEEE Power Eng. Soc. Winter Meet.*, vol. 1, Jan. 2002, pp. 177–182.

[3] "PJM Interconnection,", Norristown, PA, 99 FERC ¶ 61 139 and 99 FERC ¶ 61 227, 2002.

[4] *ISO-NE Load Response Program Manual, Appendix E: Developing Measurement and Verification Plan*, ISO New England, 2004.

[5] *NY ISO Day Ahead Demand Response Program Manual, Revised: 07. 25. 2003, Para. 2–16: Small Customers.*

[6] *Fundamentals of Load Management. IEEE Tutorial Course.* Palo Alto, CA: Electric Power Res. Inst., 1988.

[7] W. E. Mekolites, R. M. Murphy, and J. L. Laine, "AEP water heater load management test," in *Proc. Amer. Power Conf.*, 1981.

[8] J. H. Reed, J. C. Thompson, R. P. Broadwater, and A. Chandrasekaran, "Analysis of water heater data from athens load control experiment," *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1232–1238, Apr. 1989.

AUTHOR: Please supply inclusive page numbers for [9].

[9] S. H. Lee and C. E. Wilkins, "A practical approach to appliance load control analysis: A water heater case study," *IEEE Trans. Power App. Syst.*, vol. PAS-102, Apr. 1983.

[10] J. H. Reed and R. P. Broadwater, "Analysis of energy usage data for water heaters under load control," in *Proc. Southeastern IEEE Conf.*, Apr. 1988.

[11] J. H. Reed, W. R. Nelson, G. R. Weatherington, and E. R. Broadaway, "Monitoring load control at the feeder level using high speed monitoring equipment," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 694–703, Jan. 1989.

[12] G. C. Heffner and D. A. Kaufman, "Distribution substation load impacts of residential air conditioner load control," in *Proc. IEEE PES Winter Power Meet.*, Jan. 1985.

[13] M. W. Davis, T. J. Krupa, and M. J. Diedzic, "The economics of direct control of residential loads on the design and operation of the distribution system," in *Proc. IEEE PES Summer Power Meet.*, Jul. 1982.

[14] C. Forlee. Water Heating in South Africa: Facts and Figures From the 1997 "Notch Testing" Program. [Online]. Available: http://www.ctech.ac.za/conf/due/SOURCE/Web/Forlee/Forlee.html

[15] R. E. Stephens, "Load control demand reduction estimation," *IEEE Trans. Power Syst.*, vol. 3, no. 1, pp. 59–66, Feb. 1988.

[16] G. H. Weller, "Managing the instantaneous load shape impacts caused by the operation of a large-scale direct load control system," *IEEE Trans. Power Syst.*, vol. 3, no. 1, pp. 197–199, Feb. 1988.

[17] B. F. Hastings, "Ten years of operating experience with a remote controlled water heater load management system at detroit edison," in *Proc. IEEE PES Summer Power Meet.*, Jul. 1979.

[18] M. L. Goldberg and G. K. Agnew, "Protocol Development for Demand Response Calculation—Findings and Recommendations," KEMA-Xenergy, Calif. Energy Comm., Burlington, MA, Rep. 400-02-017F, 2003.

Grayson C. Heffner (SM'04) received the B.S. degree in electrical engineering and the Master of Engineering degree in mechanical engineering from the University of California, Berkeley, and a Doctor of Public Administration degree from George Washington University, Fairfax, VA.

He specializes in energy technology R&D and project management, with an emphasis on power systems planning and customer end-use issues. He has extensive electric utility experience with the Electric Power Research Institute, Palo Alto, CA, and Pacific Gas and Electric Company (PG&E), San Francisco, CA.

Dr. Heffner is a past San Francisco Chapter Chair of the IEEE Power Engineering Society.

Charles A. Goldman received the M.S. degree from the Energy and Resources Group from the University of California, Berkeley.

He is Staff Scientist and Group Leader of the Electricity Markets and Policy Group at Lawrence Berkeley National Laboratory, Berkeley. He has published over 80 articles and reports on energy efficiency and demand response policy, programs, and technology analysis, utility integrated resource planning, retail energy services, and electric industry restructuring.

Mithra M. Moezzi received the Ph.D. degree from the University of California, Berkeley.

He is a Staff Research Associate at Lawrence Berkeley National Laboratory.

IEEE PAGE CHARGES & REPRINT ORDER FORM

ORDERING INFORMATION: Please read the enclosed information before completing this form. To order reprints of your article and/or pay page charges/overlength paper charges, complete this form, making a copy for your records, and return to:

IEEE Reprints, 445 Hoes Lane, Piscataway NJ 08855-1331, fax to: (732) 981-8062

Please direct questions about reprints and this form to the Reprint Department at (732) 562-3917 or (732) 562-3941.

Digital Object Identifier: **10.1109/TPWRD.2005.852374** [INCLUDE THIS IDENTIFIER IN ALL CORRESPONDENCE]
Author(s): **Grayson Heffner, Charles Goldman, et al.**
Transactions/Journal: *IEEE Transactions on Power Delivery*, Volume: 20, Issue: 4
Paper Title: *Innovative Approaches to Verifying Demand Response of Water Heater Load Control*

PAGE CHARGES

This publication needs two forms of financial support: See attached for more about page charges and overlength paper charges.

1) A page charge of $110 per page is requested for the first 8 pages.

☐ I/my organization will pay page charges of $110 per page. [IEEE expects payment once agreed to.] ______ pages x $110 = $ 0

☒ I/my organization will not pay page charges of $110 per page.

2) A mandatory overlength paper charge of $110 is required for each page in excess of 8 pages. 2 pages x $110 = $ 220

Total Page Charges and Overlength Paper Charges: $ 220

I understand that overlength paper charges are mandatory unless the Editor has notified IEEE Publishing Services in writing and I/my organization will be charged $110 for each page in excess of 8 pages.

Date 9/1/05 Authorized Signature Charles Goldman Title or position Staff Scientist

BILLING AND SHIPPING INFORMATION

Send Reprints to (per IEEE records):
Mithra Moezzi
Lawrence Berkeley National Lab
Berkeley CA 94720

Corrections to shipping address:
Charles Goldman
Lawrence Berkeley National Laboratory
1 Cyclotron Rd
MS 90-4000
Berkeley CA
94720

Billing address (if different): (Same as New Shipping Address)

REPRINTS

If you pay the page charges ($110 per page), you are entitled to 100 standard reprints of your article at no additional cost. Reprints are sold in lots of 50 only.

No of Reprints	1-4 pages	5-8 pages	9-12 pages	13-16 pages	17-20 pages	21-24 pages	25-28 pages	29-32 pages	33-36 pages	37-40 pages	41-44 pages	45-48 pages	Color Reprints
100	$209	$366	$418	$473	$564	$673	$735	$809	$876	$944	$1,011	$1,077	add $95 per 100
50	$109	$184	$209	$236	$282	$336	$367	$404	$435	$472	$505	$538	add $85 per 50

The TOTAL number of reprints I would like (including 100 for paying charges) is:
☐ 50 ☐ 100 ☐ 150 ☐ 200 ☐ 250 or ______ (in multiples of 50 reprints)

CD REPRINTS

Number of CDs: ______ ($54.99 each) CD Format: ☐ PDF or ☐ Postscript

Cost of basic reprints from table above .... $______
Cost of CD Reprints .... $______
Reprint Cover/Title Page ☐ $36 for 50 ☐ $71 for 100 .... $______
$16 must be added for all U.S. shipments to cover UPS shipping and handling .... $______
$55 Air freight charge must be added for all orders being shipped outside the U.S. .... $______
Articles with color figures MUST BE ORDERED WELL IN ADVANCE OF PRINTING. Add $95 per 100 or $85 per 50 reprints .... $______
Add applicable tax for shipments to CA, DC, FL, NJ, NY, OH and Canada (GST Registration No. 125634188) .... $______
Total Charges for Reprints: $ 0

Date 9/1/05 Authorized Signature Charles Goldman Title or Position Staff Scientist

PAYMENT

☐ Check Enclosed. Payable on a bank in the USA. ☐ Bill Me. Orders in excess of $1,000 must have a purchase order included.
☐ Charge My: ☐ VISA ☐ MasterCard ☐ American Express ☐ Diner's Club

Acct# ______ Name of card holder ______ Exp.Date ______

Signature ______ IEEE Member# ______ ☐ I have special instructions noted on reverse side.