FREEMAN, SULLIVAN & CO.

A MEMBER OF THE FSC GROUP

2011 Residential and Small Commercial peaksaver®

September 2012

Prepared for:
The Ontario Power Authority

Prepared by:
Dries Berghman
Michael Perry
Freeman, Sullivan & Co.

Freeman, Sullivan & Co.
101 Montgomery St., 15th Floor
San Francisco, CA 94104
fscgroup.com

Message from the Evaluator

I fully endorse the process used to generate the evaluation results for OPA's *peaksaver*® program, as documented in this report.

Stephen S. George
Chief Executive Officer and Principal Consultant
The FSC Group

September 13, 2012

Acknowledgements

We would like to thank Nik Schruder and Jessei Kanagarajan at OPA, as well as Stephen George at FSC, for their helpful comments and insight.

Table of Contents

1 Executive Summary

1.1 Summary of Evaluation Goals and Objectives

This report documents methods and results for the 2011 Load Impact Evaluation of Ontario Power Authority (OPA) residential and small commercial demand response initiative (*peaksaver*®). The *peaksaver*® program involves the installation of programmable communicating thermostats (PCTs) and/or direct load control switches (DLC switches, or simply "switches") in households and small to medium businesses (SMB) with central air conditioning (CAC). The control devices allow CACs to be cycled or thermostats to be adjusted when an event is triggered, thereby reducing energy demand associated with CAC load.

The goal of this evaluation is to estimate the *ex post* load impacts of the July 21, 2011 province-wide *peaksaver*® event and to estimate *ex ante* load impacts that can be expected under specified conditions in the future.

1.2 Summary of Impact Evaluation Results

One province-wide *peaksaver*® event took place on July 21, 2011 from 4 PM to 8 PM, which was a very hot day in the province. In Toronto and Ottawa, high temperatures exceeded 36°C and, just as importantly, overnight lows never dipped below 23°C in Ottawa and 26°C in Toronto. This event provides an excellent example of program operation under the conditions for which the *peaksaver*® program was designed. Impacts are estimated for this event day using an updated version of the CAC logger model used in the 2009 and 2010 evaluations. The models use logger data from 2009 (for residential customers) and 2010 (for SMB customers) to estimate regression coefficients for the event impacts, which are then used to estimate event impacts for the July 21 event using weather data from that day.

Although there was no EM&V plan in place for 2011, many Local Distribution Companies (LDCs) were able to provide smart meter interval data as a source of information for event impact estimation. Due to the availability of this data, this is the first evaluation of the program that includes impact estimates based on load data collected outside the Toronto Hydro or Hydro Ottawa LDCs.

At the time of this writing, smart meter interval data was provided by the Hydro One and Veridian LDCs. This is the first time that load impacts have been estimated directly for those LDCs. These load impact estimates were found to be broadly in line with those previously estimated based on data collected from Toronto Hydro and Hydro Ottawa. This conclusion was made with a fair amount of uncertainty because the only source of reference loads for the 2011 event was loads observed under much cooler conditions. The 2012 evaluation will address this issue through the use of a treatment-control methodology that provides more accurate reference loads for every event day.

Due to the non-representative sample of LDCs providing data for the July 21 event, and due to the lack of comparable non-event days to provide reference load, primary *ex post* estimates for the province and for each settlement zone are based on the logger data and models used in the 2010 evaluation. *Ex post* estimates for each LDC that provided smart meter data are provided because they are the first measurement of impacts for those LDCs and therefore it is useful to compare their observed results to the estimates made based on the logger data and 2010 models.

Table 1-1 shows a broad overview of *peaksaver®* impacts.

Table 1-1: Summary of Impact Evaluation

Program Metric	Residential	Commercial	Total
Total Number of Activated Devices[1]	173,741	4,211	177,952
Aggregate *Ex post* Load Impact Estimate (MW)[2]	128.0	3.0	131.0
Aggregate *Ex post* Energy Savings (MWh)[3]	252.4	9.9	262.3
Per Customer *Ex ante* Load Impact Estimate (kW)[4]	0.56	0.64	n/a

Table 1-2 shows residential *ex post* load impact estimates for the July 21 system event and Hydro One's EM&V test event. Using the CAC logger model, the event impact for the average residential *peaksaver®* customer in Ontario is estimated to be 0.71 kW. For the average Veridian customer, the impact was 0.85 kW. Hydro One did not participate in the system-wide event; instead, it had a test event, for which the average impact was 0.50 kW.

Table 1-2: Average Residential per CAC Unit Reference Loads, Impacts and Temperatures during Event Hours

Event Type	LDC	Event Date	Event Hours	Average Reference Load (kW)	Average Event Impact (kW)	Percent Impact	Aggregate Energy Savings (MWh)[5]	Average Event Temp. (℃)
System	All	7/21/11	4 PM – 8 PM	1.89	0.71	37%	252.4	34
Test	Hydro One	7/21/11	1 PM – 3 PM	2.91	0.50	17%[6]	-	36
System	Veridian	7/21/11	4 PM – 8 PM	3.27	0.85	26%	-	34

Italics indicate that the number is calculated using smart meter data and that reference load and percent impacts are relative to whole-house load.

The load impact estimate for the province shown in Table 1-2 is larger than any of the *ex ante* values predicted for the program, which might at first appear to imply that the *ex ante* estimates are inaccurate. This is not the case. The load impact estimate is higher than any of the *ex ante* impact estimates because the average temperature on July 21 exceeded that of any of the *ex ante* weather

[1] Devices installed on central AC units only; electric water heaters are not included. The total number of devices that can be activated will likely be higher in future events because some LDCs did not participate in the system-wide event due to concerns about expiring participant agreements.

[2] During event hours.

[3] During entire event day.

[4] Impacts are shown at the per customer level because there is no program population forecast. Impacts are given for a 1-in-10 August Peak Day event.

[5] Energy savings in Table 1-1 are calculated by multiplying the average per CAC unit impact by the number of available control devices for each hour of each event, and also including the hours after the event to capture post-event snap-back. These values are then added up over each event day to get a total energy value.

[6] For Hydro One and Veridian, percent impact indicates impact on whole house load. Likewise, the average reference load is the average whole house reference load.

conditions for the province. Moreover, the maximum and minimum temperatures measured on July 21 were so extreme that they exceeded 99% of all summer time maximum and minimum temperatures measured in Ottawa and Toronto over the last century. This was true whether the July 21 weather was compared only to other July conditions or to all summer conditions. The *ex ante* weather conditions are not meant to represent the most extreme conditions that might ever be observed; they are only meant to represent particular points in a distribution of possible conditions that might be observed. In fact, the models used to estimate the *ex ante* load impacts agree quite well with load impact estimates developed using data observed in 2011, as shown later in this report. Therefore, the fact that an observed load impact exceeded the *ex ante* load impact estimate is not evidence that there is something wrong with the *ex ante* estimate. These issues are discussed more fully in Section 3.2 of this report.

Table 1-3 shows SMB *ex post* load impacts for the July 21 event. For the average SMB *peaksaver*® customer, the impact was 0.67 kW. These values are calculated using logger data gathered and models developed in the 2010 evaluation and updated based on 2011 weather. The impact for the average SMB customer is smaller than the impact for the average residential customer because most of the event took place after 5 PM, when most businesses are closed. Closed businesses tend to use their air conditioning less than open businesses. In addition, there were relatively fewer SMB customers in Ottawa than in Toronto; because Ottawa SMB customers have higher impacts than Toronto SMB customers, this causes the overall impact to be smaller.

Table 1-3: Average SMB per CAC Unit Reference Loads, Impacts and Temperatures during Event Hours

Event Date	Event Hours	Average Reference Load (kW)	Average Event Impact (kW)	Percent Impact	Aggregate Energy Savings (MWh)	Average Event Temp. (℃)
7/21/11	4 PM – 8 PM	2.18	0.67	31%	9.9	34

Ex ante load impact estimates for 2011 are identical to those produced for 2010, due to lack of new data and the fact that there was only one event in 2011. *Ex ante* estimates are reproduced for residential customers in Table 1-4. More details about these estimates can be found in last year's evaluation. The highest average impact is 0.56 kW, occurring on a 1-in-10 August Peak Day. The highest aggregate impact from residential customers is 103 MW, also on a 1-in-10 August peak day.

Table 1-4: Residential *peaksaver*® *Ex Ante* Load Impact Estimates by Weather Year and Day Type (Event Period 2-6 PM)

Type of Estimate	Day Type	Extreme Year (1-in-10)	Normal Year (1-in-2)
Per CAC unit (kW)	May Peak Day	0.29	0.22
	June Peak Day	0.42	0.29
	July Peak Day	0.54	0.33
	August Peak Day	0.56	0.37
	September Peak Day	0.28	0.24
Whole-Province Aggregate (MW)	May Peak Day	53	40
	June Peak Day	77	54
	July Peak Day	99	61
	August Peak Day	103	68
	September Peak Day	52	44

Table 1-5 shows *ex ante* load impact estimates for SMB customers for the province. The highest per CAC unit load impact occurs on a 1-in-10 August Peak Day, with an average impact of 0.64 kW. This leads to a 2.8 MW aggregate load impact from SMB customers given current customer characteristics and participation levels.

Table 1-5: SMB *peaksaver*® *Ex Ante* Load Impact Estimates by Weather Year and Day Type (Event Period 2-6 PM)

Type of Estimate	Day Type	Extreme Year (1-in-10)	Normal Year (1-in-2)
Per CAC Unit (kW)	May Peak Day	0.39	0.32
	June Peak Day	0.52	0.42
	July Peak Day	0.62	0.42
	August Peak Day	0.64	0.48
	September Peak Day	0.36	0.33
Whole-province Aggregate (MW)	May Peak Day	1.7	1.4
	June Peak Day	2.3	1.9
	July Peak Day	2.7	1.9
	August Peak Day	2.8	2.1
	September Peak Day	1.6	1.5

1.3 Recommendations

From the perspective of program evaluation, it makes sense in the future to take advantage of the presence of smart meters to enable cheaper and easier data collection for estimating load impacts. This also allows for a more precise and accurate evaluation strategy based on the use of control groups.

For the program itself, there are three main recommendations for improving performance:

- Gather data on device failure and signal reception failure and work with aggregators and Cooper (the primary technical implementer of the program) to fix those issues;
- Target high usage customers who are likely to provide large load impacts; and
- Explore other control strategies that might produce higher impacts without making customers uncomfortable.

2 Introduction

2.1 Evaluation Goals and Objectives

The objectives of the main analytical section of this report are to measure 2011 *ex post* impacts for residential and small to medium business (SMB) *peaksaver*® customers. This is accomplished primarily using CAC load data collected during the 2009 evaluation for residential customers and using CAC load data collected during the 2010 evaluation for SMB customers. An additional objective is to investigate the program's cost effectiveness.

2.2 Program Purpose and Goals

OPA's *peaksaver*® program involves the installation of programmable communicating thermostats (PCTs) and/or direct load control switches (switches) in households and small/medium businesses with central air conditioning (CAC).[7] The control devices allow CAC equipment to be cycled or thermostats to be adjusted when an event is triggered, thereby reducing energy demand associated with CAC load. Load control is administered by a set of aggregators that each use a different load control strategy. For customers with switches, 50% standard cycling is used. Customers with PCTs may be subject to several different strategies, such as:

- 0.5°C per hour temperature setback for each hour of the event;
- 2°C setback for the whole event; or
- 50% cycling.

Events can be triggered in two ways. First, the Independent Electricity System Operator (IESO) can declare an "Energy Emergency Alert."[8] Second, an event can be called when the external temperature is above 30°C and primary demand is greater than 23,000 MW. There was one such event on July 21, from 4 PM to 8 PM; during the event window, the average temperature was 34°C, and on the event day, temperatures in Toronto and Ottawa exceeded 36°C. The event was officially triggered because system demand exceeded 23,000 MW on the event day and because temperature exceeded 30°C during event hours.

Individual LDCs can also call their own test events. These types of events are generally called to test the *peaksaver*® load control system. Hydro One customers had one such event, also on July 21, from 1 PM to 3 PM. Hydro One customers did not participate in the system-wide event on July 21.

Table 2-1 shows the number of active devices as of July 2011 by customer type, device type and settlement zone. As seen in Table 2-1, the majority of *peaksaver*® customers and devices are associated with residential households. The residential segment comprises 99% of all *peaksaver*® PCTs and 95% of all load control switches. The values in Table 2-1 are taken from a list of continuing participants provided by the LDCs participating in the program.

---

7 The program also includes some load control devices on electric water heaters, but no evaluation of those load impacts is reported here.

8 Conditions for such an alert being called are referenced in the IESO's Systems Operations Manual.

The distinction between CAC units and *peaksaver*® customers is important for SMB customers, but much less so for residential customers. As of early 2012, there are about 1.01 devices per residential participant. In contrast, in the Toronto Hydro SMB population, there is an average of 1.7 devices per participant and in the Hydro Ottawa SMB population there is an average of 1.6 devices per participant. In this report, results are reported at the per device (or per CAC unit) level, and results are aggregated by multiplying by the total number of devices installed in the population. Load impact analysis using logger data is done at the level of the CAC.

The table excludes devices located in LDCs that did not participate in the event (certain LDCs aggregated by Rodan and Hydro One). In addition, it excludes nearly 1,800 load control devices installed on electric water heaters.

Table 2-1: *peaksaver*® Available Control Devices for the July 21, 2011 event

Settlement Zone	SMB			Residential			Total
	Switch	PCT	Total	Switch	PCT	Total	
Bruce	0	0	**0**	0	4	**4**	**4**
East	0	23	**23**	126	10,090	**10,216**	**10,239**
Essa	0	15	**15**	9	6,792	**6,801**	**6,816**
Georgian Bay	0	0	**0**	0	190	**190**	**190**
Long Point	1	41	**42**	20	1,071	**1,091**	**1,133**
Niagara	0	51	**51**	26	4,141	**4,167**	**4,218**
Northeast	0	1	**1**	0	178	**178**	**179**
Ottawa	4	918	**922**	6	26,377	**26,383**	**27,305**
South Central	2	46	**48**	46	17,427	**17,473**	**17,521**
Southwest	0	16	**16**	33	2,742	**2,775**	**2,791**
Toronto	2,969	53	**3,022**	63,010	35,783	**98,793**	**101,815**
West	0	71	**71**	428	5,242	**5,670**	**5,741**
Grand Total	**2,976**	**1,235**	**4,211**	**63,704**	**110,037**	**173,741**	**177,952**

2.3 Report Overview

The remainder of this report is organized as follows. Two impact evaluation sections are included, one for residential customers and one for SMB customers. Each section contains a methodology relevant to that program segment, and presents program impacts for that segment. Recommendations for both program segments are presented in Section 4.3. The final section of the report presents a cost effectiveness analysis.

Detailed tables presenting *ex post* and *ex ante* impact estimates that conform to the requirements of the Ontario Load Impact Protocols have been provided to OPA.

3 Residential Impact Evaluation

This section discusses the methods and results of the residential impact evaluation. Primary *ex post* and *ex ante* results are based on models developed in the 2010 evaluation using data collected in 2009. Supplemental *ex post* estimates are developed based on smart meter interval data provided by some LDCs.

3.1 Methodology

The objectives of the main analytical section of this report are to measure 2011 *ex post* impacts for residential *peaksaver*® customers. This is accomplished primarily using CAC load data collected during the 2009 evaluation. The following subsections briefly describe the sample and also describe the analytical approach used for the evaluation. For greater detail on these approaches, see the 2009 and 2010 *peaksaver*® evaluations.[9,10]

The CAC load data is complemented by hourly whole-building smart meter data collected during the 2011 event. Although several LDCs were contacted to obtain this data, at the time of this writing only Hydro One and Veridian provided it. Due to the limited data from 2011, both in terms of territory and the number of events, *ex post* results from 2011 are not used to develop *ex ante* estimates. *Ex ante* estimates are developed based on the data collected during 2009.

3.1.1 Dataset

The EM&V sample dataset contains hourly average CAC load data for July through September 2009 for 407 residential customers. The exact dates covered by each logger vary due to installation and retrieval schedules, but all loggers cover at least July 15 through September 30. The dataset includes weather data, but not load data, for the only 2011 *peaksaver*® event on July 21.

The logger and weather data are complemented by smart meter data from Hydro One and Veridian. Hydro One did not participate in the system-wide event on July 21, but did call an EM&V test event. Smart meter data includes load from the entire building, rather than just the CAC unit, which means that it is more difficult to estimate the drop in CAC load because it could be confounded with other activity in the house. However, smart meter data is helpful because it can include many more customers than logger data and can help to verify on a larger scale that the program runs as expected.

3.1.2 Logger Model Description

The regression models from the 2010 *peaksaver*® evaluation were used with weather data from July 21 to predict load impacts for the July 21 event. The following paragraphs only provide a brief overview of the regressions used in this evaluation. They were fully described in the 2010 *peaksaver*® evaluation.

9 Ontario Power Authority 2009 *peaksaver*® Residential Air Conditioner Measurement and Verification Study. Produced by KEMA, May 17, 2010.

10 2010 Residential and Small Commercial Demand Response Initiative and Hydro Ottawa *peaksaver*® Small Commercial Pilot Program Evaluation. Produced by the FSC Group, June 20, 2011.

In the residential sample, data from the loggers installed on the EM&V sample was analyzed using linear regressions done individually for each CAC unit. Each customer has a different usage pattern over time, and each customer's usage responds differently to changes in weather. This led us to estimate separate regressions for each CAC unit using the same regression specification. For all CAC units, the factors used to estimate usage patterns were weather variables interacted with time indicators. These allow the model to take into account different reactions to weather conditions at different times of day, times of the week and times of year. For example, a residential customer's energy usage might respond strongly to high temperatures on a Saturday afternoon when they are at home, while it might not respond at all on a Wednesday afternoon when they are at work. All analysis and results are at the per CAC unit level, with the exception of smart meter results, which are at the premise level.

The regression specification was:

$$kWh_t = a + \sum_{h=1}^{24} \mathrm{b}_h \cdot \mathrm{wacdh_t} \cdot I_h + \sum_{h=1}^{24} \mathrm{c}_h \cdot \mathrm{wacdh_t^2} \cdot I_h \cdot I_c + d \cdot wacdh_t \cdot \boldsymbol{I_e} + \boldsymbol{e} \cdot wacdh_t \cdot \boldsymbol{I_{pe}} + \sum_{h=15}^{20} \mathrm{f}_h \cdot I_h \cdot I_T + \varepsilon_t$$

Variable	Description
a	Estimated constant.
$b - i$	Estimated parameter coefficients.
$wacdh$	A weighted-average of the past 12 cooling degree hours using a base of 21°C. Weights decrease by 10% per hour so that more recent hours have stronger effects.11
I_c	Indicator for commercial customers. Interacted with hourly effects so that commercial, but not residential, customers have the effect of weighted-average CDH-squared.12
I_h	Indicator variables representing the hours of the day, designed to estimate the effect of daily schedule on usage behavior and event impacts.
I_e	Indicator variables designed to model the effects of events.
I_{pe}	Indicator variable to model the effects of post-event periods. Decreases in magnitude over the 4 hours after the event at a rate of 67% per hour to reflect that snap-back should decrease fairly quickly.
I_T	Indicator variable that models the effects of the thunderstorm in Ottawa during the August 4, 2010 event.
ε_t	The error term, assumed to be a mean zero and uncorrelated with any of the independent variables.

[11] The precise specification of temperature in the model leaves a large amount of leeway for analyst judgment. A 12-hour interval was chosen because it captures the vast majority of important variation in historical temperature, but still allows for *ex ante* predictions of afternoon event impacts without requiring assumptions about what the weather was on the day before. Whether cooling-degree hours or raw temperature is used for modeling has little impact on results.

[12] Weighted-average CDH-squared is included in the model for commercial customers, but not for residential. It does not improve the model fit for residential customers. This is not likely due to a difference between commercial and residential customers, but due to the fact that commercial customers experienced higher temperatures during the 2010 logger period than residential customers did during 2009. A squared CDH term primarily helps model fit in times of very high temperatures.

The subscript *t* indicates hour of the summer. Only non-holiday weekdays were modeled because no events were called on the weekend (weekend usage behavior is different from weekday usage). The table defines the variables and describes the effects they seek to identify.

The model described above was fit to residential CAC logger data collected in 2009. To estimate load impacts for 2011, individual regressions were first run using 2009 load and temperature data. Those individual regression estimates were then stored in memory. This step provides a function that can be used to estimate load and load impacts for any given set of weather conditions. Individual customer load was estimated using the stored estimates but applying 2011 weather data. This approach assumes that the 2009 regressions do a good job of explaining the variation in CAC load and that they are capable of producing accurate out-of-sample results. Further details about the ramifications of model selection can be found in the 2010 *peaksaver®* evaluation.

3.1.3 Logger Model Validation

In order for a model to be useful in the context of *peaksaver®*, it must make accurate predictions of CAC loads, primarily at high temperatures. This section presents an out-of-sample validation of the model's capability to do this accurately, which is especially useful to validate whether the model can predict load for days where only temperature is known. The 2010 evaluation also presents in-sample testing validation and whole-building energy regressions to further test the model presented above.

The procedure for out-of-sample testing consisted of running the regression models multiple times, each time holding back one of the hot non-event days of the summer from the estimation. Then, predicted loads were compared to actual loads on the days held back. This is a true test of the regression model's predictive power for weather conditions actually observed during 2009.[13]

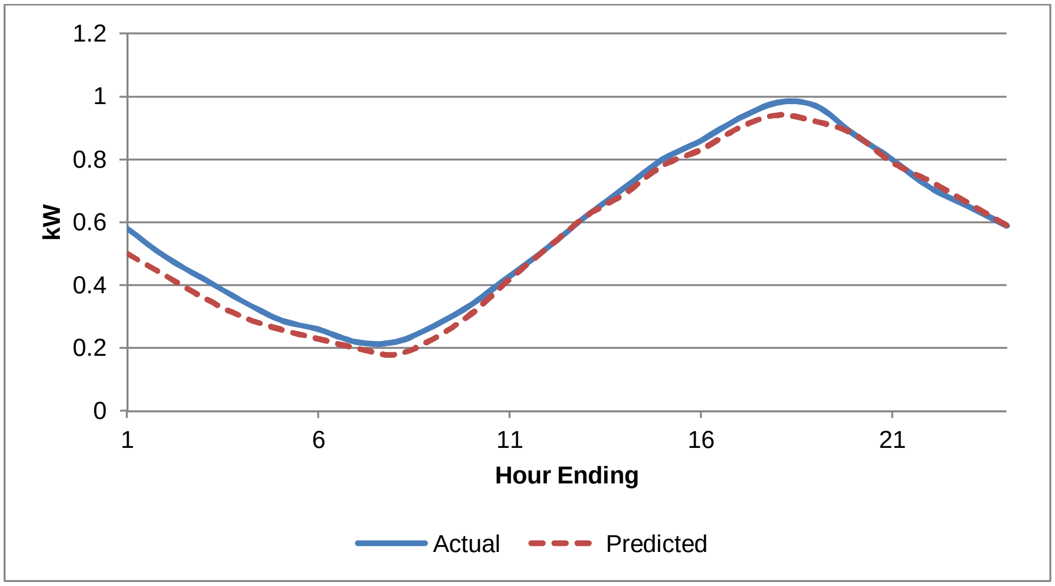
Figure 3-1 shows the actual average hourly energy use of residential CAC units on five hot out-of-sample days compared to the regression-predicted average energy use. The out-of-sample days were chosen randomly among the hottest 20 non-event weekdays of the summer of 2009.[14] The close match between predicted values and actual values reflects the ability of the regressions to predict accurately. The predicted load is very close to the actual load. For residential customers, the actual load is, on average, about 4% higher than predicted load during the hours of 1 PM to 7 PM.

However, it should be noted that the average event temperature of 34°C and the average daily maximum temperature of 36°C on July 21, 2011 was much warmer than any of the 2009 days used to do the out-of-sample testing. Thus, it is impossible to determine whether or not the logger model, which was estimated using temperature data in which less than 1% of observations exceeded 32°C, can accurately estimate impacts on the 2011 event day. This introduces uncertainty into the impact estimates, although it is not certain whether this skews impacts up or down.

[13] Additional information about the predictive power of the regression model can be found in the 2010 *peaksaver®* evaluation.

[14] For residential customers, the five 2009 days are July 11, August 1, August 13, August 15 and August 18.

Figure 3-1: Average Residential CAC Unit Actual and Predicted Load for Out-of-sample Days



3.1.4 Smart Meter Model Description

At the time of this writing, the Hydro One and Veridian LDCs had provided smart meter data. They provided data for the July 21 event for a sample of 3,833 and 6,735 *peaksaver*® customers, respectively.[15] Hydro One called an EM&V event from 1 PM to 3 PM, while Veridian participated in the system-wide OPA event. Both LDCs also provided data for the same customers for July 20 and June 8, which were the next two hottest weekdays in Toronto and Ottawa during 2011.[16]

Because the event day was so much hotter than the non-event days, a regression model such as the one above would not be particularly accurate here. The regression approach relies on the existence of non-event days that have comparable temperatures to the event days. Instead, participants' average load on the July 21 event day was compared to their average load on the next hottest day, July 20. Pre-event loads on July 21 were much higher than loads during the same hours on July 20 because July 21 was much hotter. To account for this difference, loads on July 20 were adjusted upward using the following ratio for Hydro One:

$$\frac{Average\ load\ on\ July\ 21, 11\ AM - 1PM}{Average\ load\ on\ July\ 20, 11\ AM - 1\ PM} = \frac{2.51\ kW}{2.00\ kW} = 1.26$$

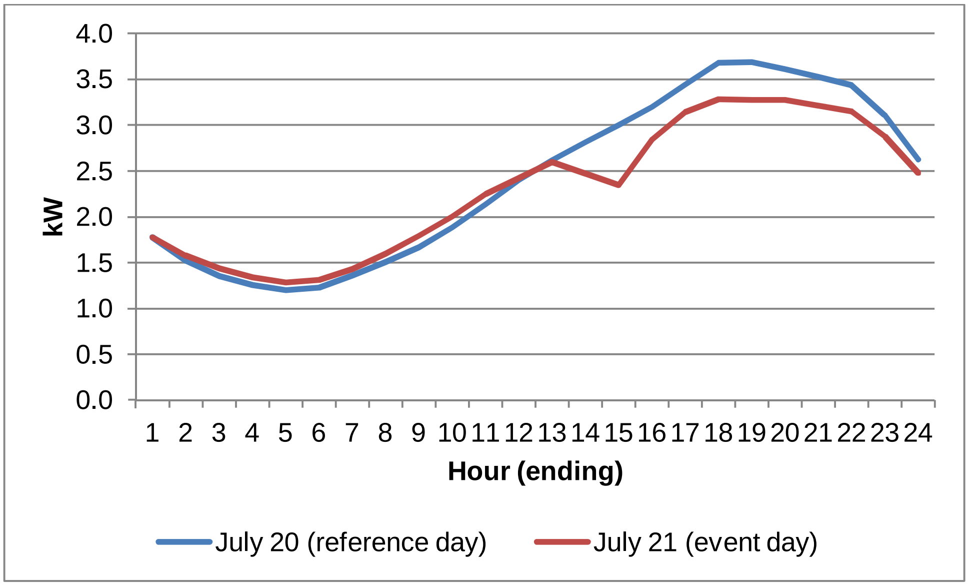
This same-day adjustment is a standard method for adjusting baseline calculations. It relies on the high correlation in loads across hours of the same day to increase accuracy in load impact estimates. The adjustment yields loads on July 20 that are more comparable to loads on July 21. After adjusting July 20 loads, event impacts can be calculated simply by subtracting July 20 loads from July 21 loads. Figure 3-2 shows the load on July 21 and the adjusted load for July 20 for Hydro One. Between the

[15] When it provided the data, Hydro One stated that it included both residential customers and farm customers classified as residential customers. It was assumed that customers with peak hourly event-day loads in excess of 6 kW were farm customers; they were excluded. This leaves 3,833 residential customers.

[16] On July 21, the average maximum temperature was 36°C; on July 20, it was 33°C; and on June 8, it was also 33°C.

hours of 1 PM and 3 PM (hours ending 14 and 15), there is a clear notch in the load on July 21, indicating that customers are dropping CAC load as they participate in the event. Despite the same-day adjustment, the load on July 20 is not a perfect reference load because it is greater than the load on July 21, even after the event ended. This suggests that the event impact estimated using this method is over-estimated and can be viewed as an upper bound on the true event impact.

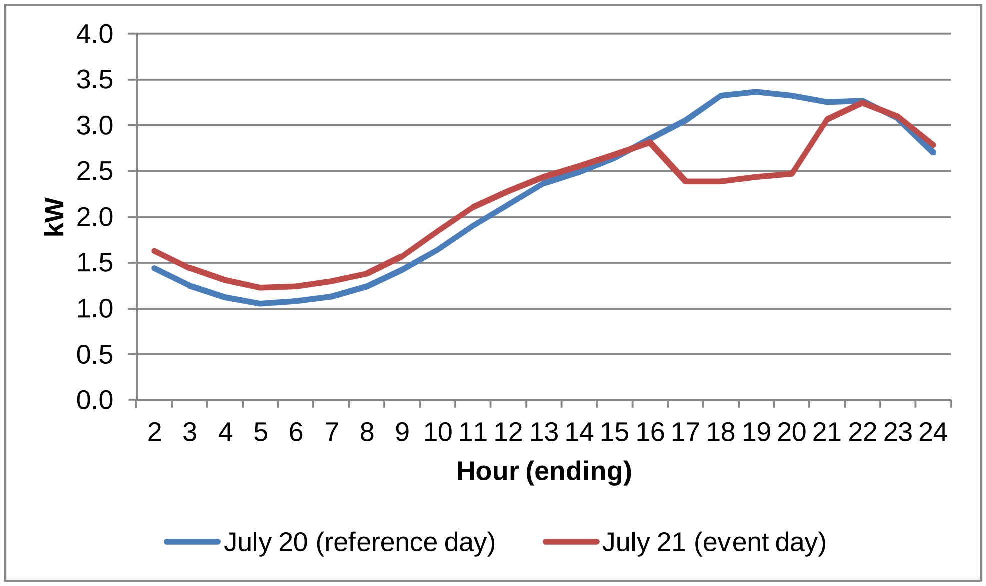
Figure 3-2: Average Load and Reference Load for Hydro One Test Event



The same method was used for Veridian. Loads on July 20 were adjusted upward by a ratio of 1.22 to serve as a reference load. Figure 3-3 shows the event load (July 21) and the reference load (adjusted load for July 20) for Veridian.[17] The adjusted July 20 load provides a very plausible reference load for July 21 for Veridian customers, with pre- and post-event loads that are similar between the two days. This single-day matching method of developing reference load always produces results with substantial uncertainty, but this is an example of a case where it appears to work quite well.

[17] Note that there are only 23 hours of data available in any one calendar day because Veridian's data was not initially adjusted for daylight savings time.

Figure 3-3: Average Load and Reference Load for 2011 Veridian Event



3.2 Results and Discussion

This section presents and discusses the load impacts for residential customers for the 2011 evaluation.

Table 3-1 shows the average impact per customer for each *peaksaver*® event along with total energy savings and average temperature over the event period. The impact on the only province-wide event day, July 21, was estimated to be 0.71 kW per customer. Unsurprisingly, temperatures on July 21 were by far the hottest experienced in 2011, exceeding 36°C in the hours just preceding the event.

The overall average system-wide impact of 0.71 kW represents a 37% load reduction. It is not surprising that the average load reduction is significantly less than 50% even though a 50% cycling strategy is used. Under a 50% cycling strategy, the maximum load reduction is 50% and that is only achieved when the CAC unit would normally be running at a 100% duty cycle. Even though July 21 was a particularly hot day, temperatures in Ontario rarely reach levels that would cause CACs to run at a 100% duty cycle. Additionally, even if some CACs are at 100% duty cycle, other CACs will be off or running at lower levels. This is to be expected even under very hot conditions.

Hydro One called a test event on the same day, lasting from 1 PM to 3 PM. The average impact for the event was 0.50 kW, which represents 17% of the whole-house reference load. In the table, the percentage impact for Hydro One's test event is calculated relative to the whole-house reference load, which includes other appliances in addition to the CAC unit. Thus, the percentage is smaller than the percent impact calculated for the system-wide event, which only includes the CAC unit. Nevertheless, the absolute (kW) impact of Hydro One's event is smaller than the impact of the system-wide event. This is because the Hydro One test event took place during a different set of hours than the system event. Residential CAC loads tend to peak in the early evening. Table 3-2 compares model and smart meter results for the same set of hours and shows that there is no significant difference between results calculated using the two methods.

The last row in Table 3-1 shows system event impacts calculated using Veridian smart meter data. Event impacts from 4 PM to 8 PM are 0.85 kW, or 26% of the whole-house load.

As was alluded to in the executive summary, the high load impacts on July 21 were the result of conditions that exceeded those in the *ex ante* weather conditions for this program. The load impact estimate for the province shown in Table 3-1 is larger than any of the *ex ante* values predicted for the program, which might at first appear to imply that the *ex ante* estimates are inaccurate. This is not the case. The load impact estimate is higher than any of the *ex ante* impact estimates because the average temperature on July 21 exceeded that of any of the July *ex ante* weather conditions for the province. The temperature on July 21 reached about 36°C, which is closer to the high temperature for a 1-in-10 August day in the *ex ante* weather conditions. However, on the 1-in-10 August day, the nighttime low reaches 21°C, while on July 21, the nighttime low only got down to 26°C. Overnight lows have a large impact on cooling load later in the day because they partially determine how much heat a building retains from the day before. When it does not cool off overnight, a building stays warm, more cooling load is needed, and more demand response is available.

The *ex ante* weather conditions are not meant to represent the most extreme conditions that might ever be observed; they are only meant to represent particular points in a distribution of possible conditions that might be observed. The maximum and minimum temperatures observed on July 21 in Toronto and Ottawa exceed 99% of the maximum and minimum temperatures measured in July between 1899 and 2003; they also exceed 99% of the maximum and minimum temperatures measured from July through September between 1899 and 2003. Because they exceeded 99% of the measured temperatures in a period exceeding 100 years, this implies that July 21 was closer to a 1-in-100 event than a 1-in-10 event; thus, it is not surprising that impacts on July 21 should exceed those expected on a 1-in-10 event day.

Table 3-1: Average Residential per CAC Unit Reference Loads, *Ex post* Impacts and Temperatures During Event Hours (Whole-province)

Event Type	LDC	Event Date	Event Hours	Average Reference Load (kW)	Average Event Impact (kW)	Percent Impact	Aggregate Energy Savings (MWh)	Average Event Temp. (℃)
System	All	7/21/11	4 PM – 8 PM	1.89	0.71	37%	252.4	34
Test	Hydro One	7/21/11	1 PM – 3 PM	2.91	0.50	17%[18]	-	36
System	Veridian	7/21/11	4 PM – 8 PM	3.27	0.85	26%	-	34

Italics indicate that the number is calculated using smart meter data and that reference load and percent impacts are relative to whole-house load.

The system-wide logger model impacts shown in the first row of Table 3-1 – which are weighted to be representative for the average *peaksaver*® customer in OPA's territory – are not comparable to the

[18] For Hydro One and Veridian, percent impact indicates impact on whole house load. Likewise, the average reference load is the average whole house reference load.

smart meter-based impacts for Hydro One and Veridian, which are shown in the second and third rows, respectively. Hydro One and Veridian's *peaksaver*® customers represent a small fraction of the entire population and are located in specific areas. Table 3-2 presents logger model results that are weighted to reflect the Hydro One and Veridian populations,[19] and compares them with the smart meter results. For Veridian, it presents results from 4 PM to 8 PM, which causes the smart meter model impacts shown in Table 3-2 to differ from the impacts shown in Table 3-1.

Table 3-2: Comparison Between Logger and Smart Meter Models

LDC	Event Date	Event Time	Impact Source	Average Event Impact (kW)
Hydro One	7/21/2011	1 PM - 3 PM	Logger Model	0.51
			Smart Meter	0.50
Veridian	7/21/2011	4 PM - 8 PM	Logger Model	0.77
			Smart Meter	0.85

Table 3-2 shows that for Hydro One, the logger and smart meter models do not differ significantly, and that despite the fact that no new logger data was collected, the logger model still performs well. For Hydro One, the results calculated using the two methods are nearly identical. For Veridian, the results differ by 0.08 kW, or 9%. This difference is well within confidence intervals of the model.

3.3 Recommendations

Recommendations for both the residential and commercial segments of the program are presented in section 4.3.

[19] The logger model produces results at the settlement zone level. To get system-level results, settlement zone results are weighted by the number of system-wide *peaksaver*® customers in each settlement zone. To get model results for Hydro One, results are weighted by the number of Hydro One *peaksaver*® customers in each settlement zone. Model results for Veridian are simply results from the Toronto settlement zone.

4 Commercial Impact Evaluation

4.1 Methodology

The objectives of this analytical section of the report are to measure 2011 *ex post* impacts for small to medium business (SMB) *peaksaver*® customers. This is accomplished primarily using CAC load data collected during the 2010 evaluation. The following subsections briefly describe the sample and also describe the analytical approach used. For greater detail on these approaches, see the 2010 *peaksaver*® evaluation.[20]

In this case, 2010 CAC load data is not complemented by hourly whole-building smart meter data collected during the 2011 event. In addition, due to the limited data from 2011, both in terms of territory and the number of events, *ex post* results from 2011 are not used to develop *ex ante* estimates. *Ex ante* estimates are developed based on the data collected during 2010.

4.1.1 Sampling

Load impact estimates for SMB customers were derived from measurements of five-minute average CAC loads obtained from a sample of SMB *peaksaver*® customers by The FSC Group in 2010. Due to the low numbers of SMB devices associated with the other LDCs, these devices were only installed on businesses served by Toronto Hydro and Hydro Ottawa. To make the sample representative of the other LDCs, customers were assigned to represent settlement zones where no data was collected. To make the sample representative of SMB customers in Toronto, the recruitment sample for that LDC was stratified based on three industry classifications – office, retail and other – and the number of controlled CACs – 1 to 3 and 4+. For Hydro Ottawa, no stratification was used; since the recruited population contained almost 40% of the total customer population, it was likely to represent most population segments of interest. A total of 371 loggers were installed: 226 in Toronto and 145 in Ottawa.

For both LDCs, the sample is representative of the population in terms of the industry designations, but less so in terms of the number of devices per site. The sample under-represents sites with four or more control devices. Unfortunately, sample weights cannot solve this problem because the number of EM&V customers with four or more devices is so small that the weight assigned would be very large and would produce average results that were highly dependent on just a few customers. Therefore, reported results are un-weighted. Further details about the 2010 SMB logger data can be found in the 2010 *peaksaver*® evaluation.

4.1.2 Dataset

The EM&V sample dataset contains hourly average CAC load data for July through September 2010 for 371 SMB customers. The exact dates covered by each logger vary due to installation and retrieval schedules, but all loggers cover at least July 15 through September 30. The dataset includes weather data, but not load data, for the only 2011 *peaksaver*® event on July 21.

[20] 2010 Residential and Small Commercial Demand Response Initiative and Hydro Ottawa *peaksaver*® Small Commercial Pilot Program Evaluation. Produced by the FSC Group, June 20, 2011.

4.1.3 Logger Model Description

The logger model used for commercial CAC units was almost identical to that used for residential CAC units. See section 3.1.2 for a description.

One important modification to the model used in 2010 is that event impacts in later hours were adjusted slightly to account for the fact that the event took place when most businesses would have been closed. The events over which the model was estimated all took place primarily during normal business hours, so an unadjusted model using 2010 data would overestimate later-hour event impacts in 2011. This is because the model estimated impacts in terms of absolute demand rather than as a fraction of reference load and was originally estimated on data that included only events occurring earlier in the day, when most businesses are open and using their CACs. The adjustment consisted of applying the load impacts as a fraction of reference load for each hour from 4 to 6 PM to the reference load for each hour from 6 to 8 PM. For example, if the load impact from 4 to 5 PM was 25% and the impact from 5 to 6 PM was 20%, while the reference load from 6 to 7 PM was 2 kW and the reference load from 7 to 8 PM was 1.8 kW, the impact for 6 to 7 PM would have been 0.5 kW (2kW multiplied by 25%) and the impact for 7 to 8 PM would have been 0.36 kW (1.8kW multiplied by 20%).

4.1.4 Logger Model Validation

In order for a model to be useful in the context of *peaksaver®*, it must make accurate predictions of CAC loads, primarily at high temperatures. This section presents an out-of-sample validation of the model's capability to do this accurately, which is especially useful to validate whether the model can predict load for days where only temperature is known. The 2010 evaluation also presents in-sample testing validation and whole-building energy regressions to further test the model presented above.

The procedure for out-of-sample testing consisted of running the regression models multiple times, each time holding back one of the hot non-event days of the summer from the estimation. Then, predicted loads were compared to actual loads on the days held back. This is a true test of the regression model's predictive power for weather conditions actually observed during 2010.[21]

Figure 4-1 shows the actual average hourly energy use of commercial CAC units on five hot out-of-sample days compared to the regression-predicted average energy use. The out-of-sample days were chosen randomly among the hottest 20 non-event weekdays in 2010. The average high temperature for the five days was 28°C.[22] The close match between predicted values and actual values reflects the ability of the regressions to predict accurately. In both cases, the predicted load is very close to the actual load. For SMB customers, actual load is on average about 2% higher than predicted load during those hours.

However, it should be noted that the average event temperature of 34°C and the average daily maximum temperature of 36°C on July 21, 2011 was much warmer than any of the 2010 days used to do the out-of-sample testing. Thus, it is impossible to determine whether or not the logger model, which was estimated using temperature data in which less than 1% of observations exceeded 32°C,

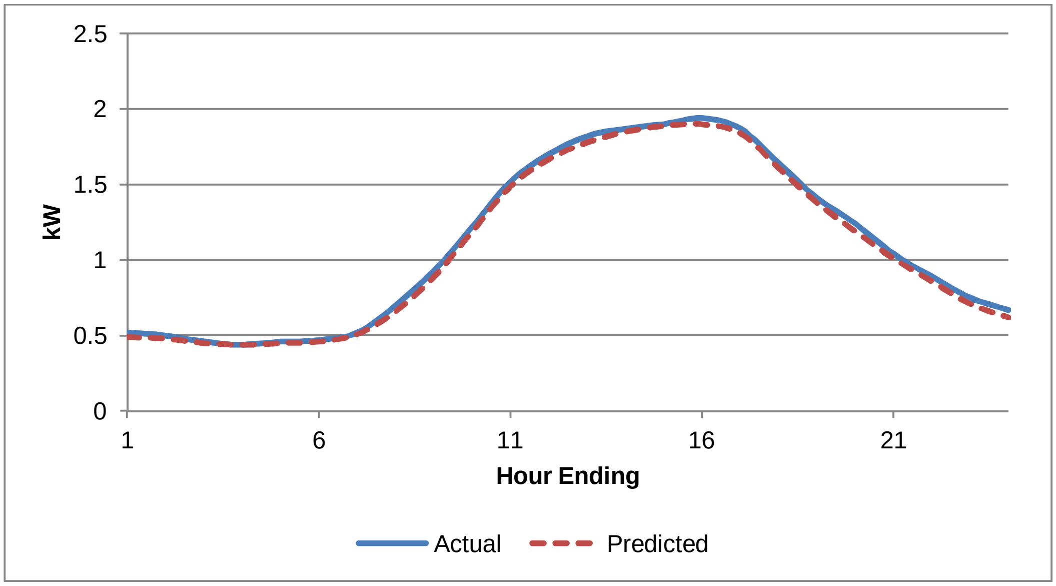
---

21 Additional information about the predictive power of the regression model can be found in the 2010 *peaksaver®* evaluation.

22 For SMB customers, the five 2010 days are July 16, July 27, August 9, August 10 and August 19.

can accurately estimate impacts on the 2011 event day. This introduces uncertainty into the impact estimates, although it is not certain whether this skews impacts up or down.

Figure 4-1: Average SMB CAC Unit Actual and Predicted Load for Out-of-sample Days



4.2 Results and Discussion

Table 4-1 shows the average per CAC unit event impact estimate and average temperature during event hours for the SMB *peaksaver*® population during the July 21, 2011 event. Unlike the residential *ex post* impacts, this table does not include any smart meter-based estimates.

The average impact across province-wide events is 0.66 kW per CAC unit, which is 30% of CAC load. This impact is only slightly higher than the impact measured in FSC's evaluation of the 2010 events. While temperatures during the 2011 event were considerably hotter than those measured during the 2010 events, the 2011 event took place fairly late in the day; many businesses were likely already closed and were not running their CAC units with the same intensity they would be running during the day. There were also differences in the geographical location of customers between 2010 and 2011; in 2011, there were relatively fewer customers in high-impact LDCs and more customers in low-impact LDCs, thus decreasing the overall results.

In addition to the possibility that event impacts are smaller because CAC units would not have been running at full capacity during the event period, event impacts in Table 4-1 are influenced by control device communication failure, which was only directly observable for a fairly small number of customers and events in the SMB *peaksaver*® load-research sample. However, based on what was observable through field work, communication failure appears to be a major issue in Toronto Hydro's territory and possibly important in Hydro Ottawa's territory. The effects of communication failure are automatically incorporated into the *ex post* impact estimates because devices that fail to respond to an event simply show no impact in the CAC data collected.

Table 4-1: Average SMB per CAC Unit Reference Loads, *Ex post* Impacts and Temperatures During Event Hours

Event Date	Event Hours	Average Reference Load (kW)	Average Event Impact (kW)	Percent Impact	Aggregate Energy Savings (MWh)	Average Event Temp. (℃)
7/21/11	**4 PM – 8 PM**	**2.18**	**0.67**	**31%**	**9.9**	**34**

As described in the 2010 evaluation, load impact estimates in Table 4-1 have a correction factor applied to account for missing, disconnected and broken control devices. For devices in Hydro Ottawa's territory, the correction factor is 0.99, while for devices in Toronto Hydro's territory, the correction factor is 0.84. Without the correction factor, the average event impact is 0.70 kW and represents an increase of about 5% over the original value of 0.67 kW. This value is an estimate of the potential load impact per device if all broken, disconnected and missing devices were accounted for and removed from the list of available devices. As was discussed in last year's evaluation, another issue that reduces load impacts is control device communication failure. The 0.70 kW value includes the effect of control device communication failure. There is not currently enough information to accurately estimate how large load impacts would be if there was no communication failure in the control device population.

4.3 Recommendations

For future evaluations, it is recommended to use the smart meters in place for almost all residential customers to provide data for use in impact estimation. Additionally, it is recommended to put in place an EM&V plan in advance of the 2012 season to allow for more accurate impact estimates. This can best be accomplished through the use of a control group. Also, events called on small subsets of customers specifically for the sake of EM&V can provide significantly more information about program operation while impacting customers minimally.

For the program itself, it is recommended to gather further data on device failures related both to failure of paging signals to reach devices and due to breakage or customer removal.

Additionally, evidence gathered in the 2010 evaluation and in other similar programs strongly suggests that load impacts per customer can be improved by targeting high use customers. This could be highly beneficial for the program and might be quite feasible given the planned expansion of the program in 2012.

Finally, the program currently uses a 50% simple cycling strategy, which may not provide large impacts under the moderate temperatures typically experienced during events. More sophisticated strategies, such as those that make use of a baseline algorithm to provide higher load impacts may work better. Making use of the large number of smart meters in the territory, a strategy such as this could be tested side-by-side on the same event days as the current standard strategy. Load impacts could then be compared to determine whether the more sophisticated strategy produced significantly larger impacts. Post-event surveys could be used to ensure that customers do not feel significantly more discomfort from either strategy.

Appendix A Load Impact Tables

Table A-1: Residential – July 21, 2011 *Ex Post* Impacts

Hour Ending	Reference Load (kW)	Estimated Load w/ DR (kW)	Load Impact (kW)	Temp (C)
1	0.23	0.23	0.00	26.9
2	0.23	0.23	0.00	26.3
3	0.22	0.22	0.00	26.2
4	0.21	0.21	0.00	26.1
5	0.27	0.27	0.00	25.8
6	0.48	0.48	0.00	26.4
7	0.50	0.50	0.00	27.6
8	0.72	0.72	0.00	29.1
9	0.94	0.94	0.00	31.2
10	1.29	1.29	0.00	31.9
11	1.47	1.47	0.00	33.1
12	1.54	1.54	0.00	34.9
13	1.66	1.66	0.00	35.6
14	1.68	1.68	0.00	35.9
15	1.71	1.71	0.00	36.4
16	1.71	1.71	0.00	36.3
17	1.85	1.18	0.67	35.9
18	1.96	1.25	0.71	35.1
19	1.89	1.16	0.73	33.8
20	1.86	1.12	0.74	31.1
21	1.69	2.20	-0.51	29.3
22	1.60	1.97	-0.37	28.1
23	1.47	1.75	-0.29	27.4
24	1.33	1.55	-0.22	26.4

Table A-2: Commercial – July 21, 2011 *Ex Post* Impacts

Hour Ending	Reference Load (kW)	Estimated Load w/ DR (kW)	Load Impact (kW)	Temp (C)
1	0.34	0.34	0.00	27.0
2	0.38	0.38	0.00	26.4
3	0.42	0.42	0.00	26.3
4	0.59	0.59	0.00	26.2
5	0.57	0.57	0.00	25.9
6	0.60	0.60	0.00	26.4
7	0.79	0.79	0.00	27.6
8	1.10	1.10	0.00	29.1
9	1.45	1.45	0.00	31.3
10	1.86	1.86	0.00	31.9
11	2.12	2.12	0.00	33.1
12	2.30	2.30	0.00	35.0
13	2.36	2.36	0.00	35.7
14	2.44	2.44	0.00	36.0
15	2.48	2.48	0.00	36.5
16	2.52	2.52	0.00	36.3
17	2.45	1.74	0.71	36.0
18	2.27	1.55	0.72	35.2
19	2.10	1.49	0.61	33.9
20	1.91	1.28	0.63	31.2
21	1.74	1.89	-0.15	29.3
22	1.59	1.68	-0.09	28.2
23	1.46	1.52	-0.05	27.5
24	1.36	1.39	-0.04	26.5

Appendix B Mapping Between LDCs and Settlement Zones

For general background information, Table B-1 shows the settlement zones served by each LDC. Note that some LDCs serve more than one zone. This information comes from enrolment lists supplied by the LDCs. Customers were mapped to a Settlement Zone based on the postal code in their records.

Table B-1: Settlement Zones Served by Each LDC

LDC	Settlement Zone	Number of Customers
Bluewater Power	Niagara	4
Bluewater Power	Southwest	4
Bluewater Power	West	1,084
Brantford Power	Long Point	1,133
Brantford Power	South Central	3
Brantford Power	Toronto	1
Brantford Power	West	1
Burlington Hydro	South Central	4,914
Chatham-Kent Hydro	South Central	1
Chatham-Kent Hydro	West	211
Clinton Power	Southwest	13
Cooperative Hydro Embrun	Ottawa	109
Enersource	Toronto	11,905
Erie Thames Powerlines	South Central	37
Erie Thames Powerlines	Southwest	32
Erie Thames Powerlines	West	587
Greater Sudbury Hydro	Northeast	133
Halton Hydro	South Central	948
Horizon	East	4
Horizon	Essa	1
Horizon	Niagara	3,669
Horizon	Ottawa	1
Horizon	South Central	8,076
Horizon	Toronto	6
Hydro 2000	Ottawa	17
Innisfil Hydro	Essa	247
Innisfil Hydro	South Central	1
Innisfil Hydro	Toronto	1
Middlesex Power	Southwest	1
Middlesex Power	West	197

LDC	Settlement Zone	Number of Customers
Oakville Hydro	South Central	3,539
Oakville Hydro	Toronto	3
Orangeville Hydro	Georgian Bay	190
Ottawa Hydro	Niagara	1
Ottawa Hydro	Ottawa	23,846
Powerstream	Essa	4,189
Powerstream	Toronto	12,695
Renfrew Hydro	East	139
Toronto Hydro	Toronto	65,783
Veridian	East	960
Veridian	Essa	115
Veridian	South Central	2
Veridian	Toronto	6,071
West Perth Power	Southwest	4

As discussed in the SMB section, settlement zones were assigned to either Ottawa or Toronto. This table contains those assignments.

Table B-2: Assignments Between Un-represented Settlement Zones and Toronto Hydro or Hydro Ottawa Modeling Results

Settlement Zone	Assigned Model Group
Northwest	Ottawa
Northeast	Ottawa
Essa	Ottawa
East	Ottawa
Georgian Bay	Toronto
Bruce	Toronto
Southwest	Toronto
South Central	Toronto
West	Toronto
Long Point	Toronto
Niagara	Toronto

Appendix C Claiming Electricity Savings from the Installation of a Programmable Communicating Thermostat

FSC was asked by the OPA and at least one LDC to investigate the possibility of claiming electricity savings from the installation of Programmable Communicating Thermostats (PCTs) in conjunction with their use as load control devices. Conceptually, there are at least three ways in which the installation of a PCT in conjunction with a load control program might lead to reductions in electricity use:

- The PCT might replace a manual thermostat (that is, a thermostat that can only accommodate a single set point) and be programmed to conserve energy;
- The PCT might replace a programmable thermostat that was not programmed with energy savings in mind and through the installation process of the PCT, customers might program it in a more conservation minded manner;
- A reduction in energy use as a result of adjusting the thermostat during a load control event.

The third type of savings is covered under the existing load impact evaluations and, thus, is not relevant to this review.

FSC contacted numerous experts from utilities and consulting firms in an attempt to identity any empirical studies that estimated conservation effects stemming from the replacement of existing manual or programmable thermostats with a PCT. Specifically, we contacted representatives from PG&E, SCE, SDG&E, Lawrence Berkeley National Laboratory, KEMA, The Brattle Group, Itron, Navigant Consulting and Christensen Associates Energy Consulting. We heard back from nearly everyone we reached out to and no one claimed to be aware of any studies or relevant information beyond what is summarized below.

A contact at a California utility indicated that California utilities used to be able to claim energy savings due to switching from manual thermostats to PCTs, but that the California Energy Commission (CEC) stopped this practice based on studies showing that customers are likely to change a pre-programmed set point to a more comfortable level if it is especially hot or cold outside. In other words, while a customer may have programmed energy-saving set points, those savings will not be realized when the customer overrides those set points.

Only one relevant study was identified. The study, authored by KEMA and Southern California Edison (SCE) found that, in general, customers do not use PCTs to save energy. The authors cited both engineering simulation and statistical analyses of self-reported thermostat set points and found that customers with PCTs and manual thermostats have similar set point behavior for cooling (Dyson, Samiullah, Rasmussen, & Cavalli, 2005). These analyses also found that consumers with PCTs use slightly more heating energy than their counterparts with manual thermostats. In addition, the same paper cited a survey of purchasers of PCTs finding that few consumers used the factory settings designed to maximize energy savings. Accepting these findings, California's three large investor-owned utilities decided not to offer rebates for programmable thermostats in their 2006-2008 residential programs.

In addition, one of FSC's contacts at another consultancy stated that her organization had investigated energy efficiency savings from PCTs from a few angles in various territories, and that their conclusion was similar. She said that "[w]hile lowering heating set points or raising cooling set points certainly saves energy, PCTs themselves usually result in minimal energy savings." Some people use them for convenience to do what they were already doing, some never change set points regardless of thermostat type, and some use [them] in an anti-savings manner, to make the home comfortable before they get up or arrive [home], rather than only cranking up the heat or cooling when they get up or get home." She further stated that "PCTs are ubiquitous in many markets"; in other words, even if some savings could be attributed to switching from a manual thermostat to a PCT, it is unlikely that many customers still had manual thermostats that could be upgraded. Her views were confirmed as being current and accurate by professionals from other utilities and consulting firms.

In conclusion, the three major California utilities do not currently claim energy efficiency savings from switching customers to PCTs, and research shows that any such savings are negligible. OPA discontinued incentivizing programmable thermostats as part of OPA's HVAC program because cooling-related (CAC) savings were found to be limited.

Generally speaking, a study seeking to estimate the savings from PCTs should randomly assign a large number of customers without PCTs to a treatment and a control group. Customers in the treatment group would get a PCT, and their energy usage would be monitored for a full year. If the treatment and control groups are large enough and randomly assigned, the difference in usage can be attributed to the effect of installing a PCT using a difference-in-difference approach. However, such a study would be difficult to carry out because most customers already have a PCT. This could mean that customers *without* a PCT may be systematically different from customers who already have a PCT, meaning that the study's results may not be externally valid and may not be applicable to the general population.

Source: Dyson, C., Samiullah, S., Rasmussen, T., & Cavalli, J. (2005). Can Programmable Thermostats Be Part of a Cost-Effective Residential Program Portfolio? Energy Program Evaluation Conference, 2005, 243-254.