

Baseline Analyses Using DBP (2006) & AMP (2008) Program Data

Steven Braithwait
Christensen Associates Energy Consulting

Conference Call

May 26, 2009

Project Objectives (2006)

- ❑ Assess the *accuracy* and *bias* of different versions of the 3-in-10 day baseline methods
- ❑ Assess whether different types of *baseline adjustments* can reduce the anticipated downward bias of unadjusted baselines
 - Event-day usage
 - Notification-day usage

Project Objectives (2008)

- Compare performance of:
 - Aggregator-level and “Sum-of-Customer” baselines
 - Baselines constructed from different numbers of non-event days (*e.g.*, 3-, 5-, or 10-in-10 day baselines)
- Assess the effect of *baseline adjustments* on the tendency of unadjusted baselines to *understate* the “true” baseline (*i.e.*, downward bias)
- Test whether “gaming” was avoided for customers/aggregators who selected the adjusted baseline option in 2008

Baseline Performance Measures

- Accuracy:
 - Measured as relative *inaccuracy* using Relative Root Mean Square Error – a fraction between 0 and 1 (*e.g.*, 10 percent relative error)
 - When assessing individual customer results (*e.g.*, DBP), use *median* of distribution of relative errors

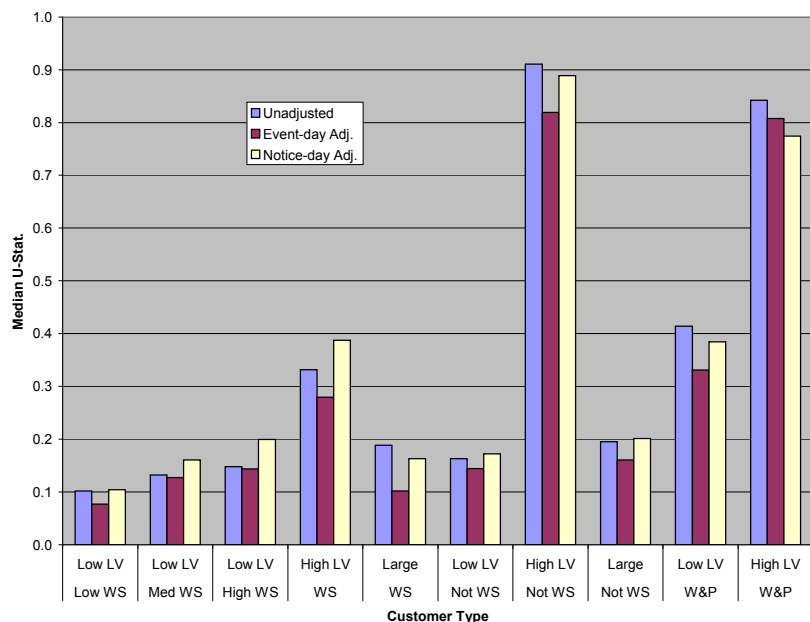
- Bias:
 - *Median* of distribution of % errors across events (& customers, where relevant)
 - By convention, $Error = True\ BL - Estimated\ BL$; so *positive* errors indicate *downward* bias
 - *Distributions* of % errors around the median also examined

Baseline Analysis Results

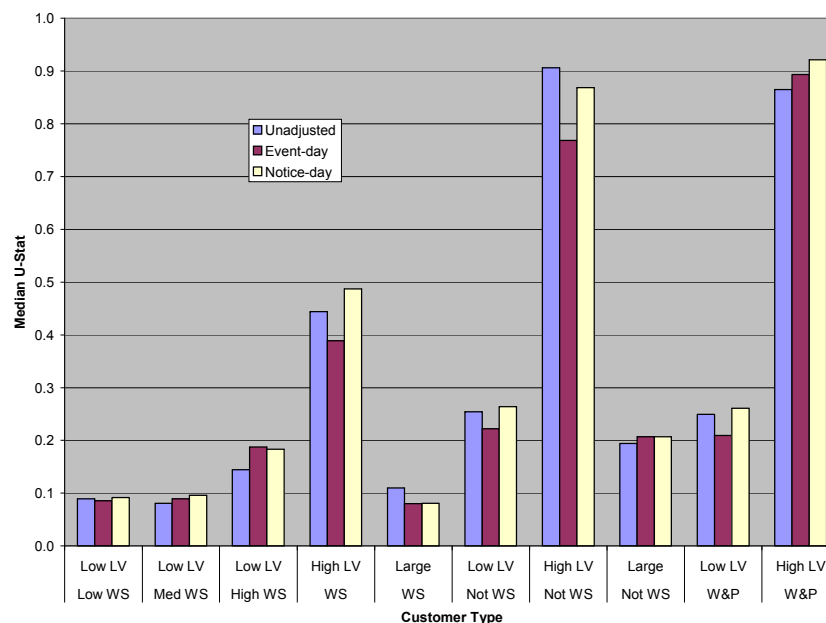
- ❑ Performance of 3-in-10 Baselines for Individual Customer (2006 DBP)
 - Accuracy and bias, by customer type
- ❑ Performance of Alternative Baselines for Aggregations of Customers (2008 AMP)
 - Accuracy and bias of aggregate vs. sum-of-customer, by aggregator

DBP 2006: Unadjusted and Adjusted 3-in-10 – Accuracy, by Weather Sensitivity & Load Variability

PG&E



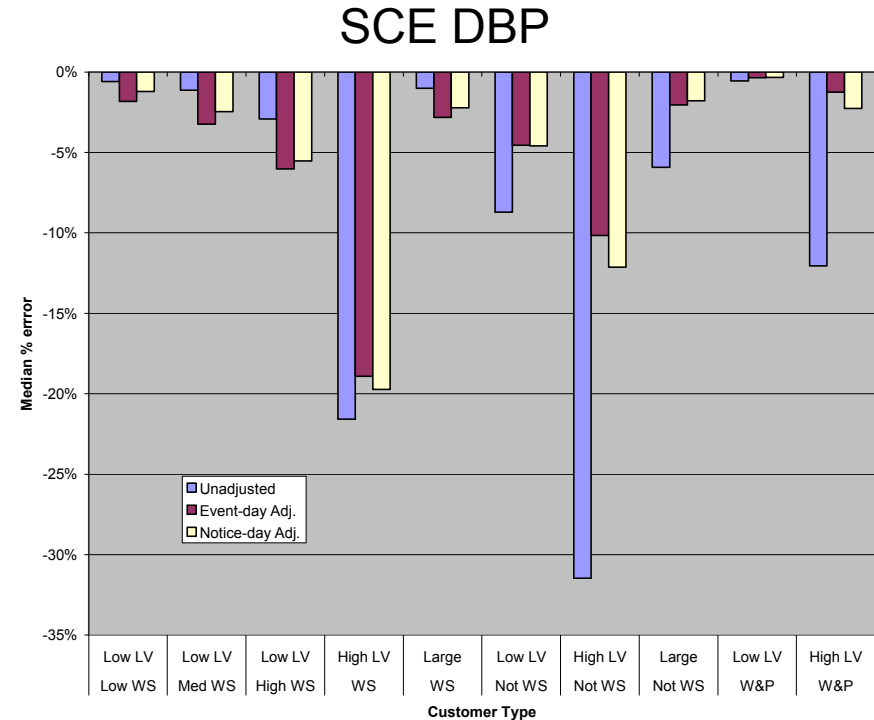
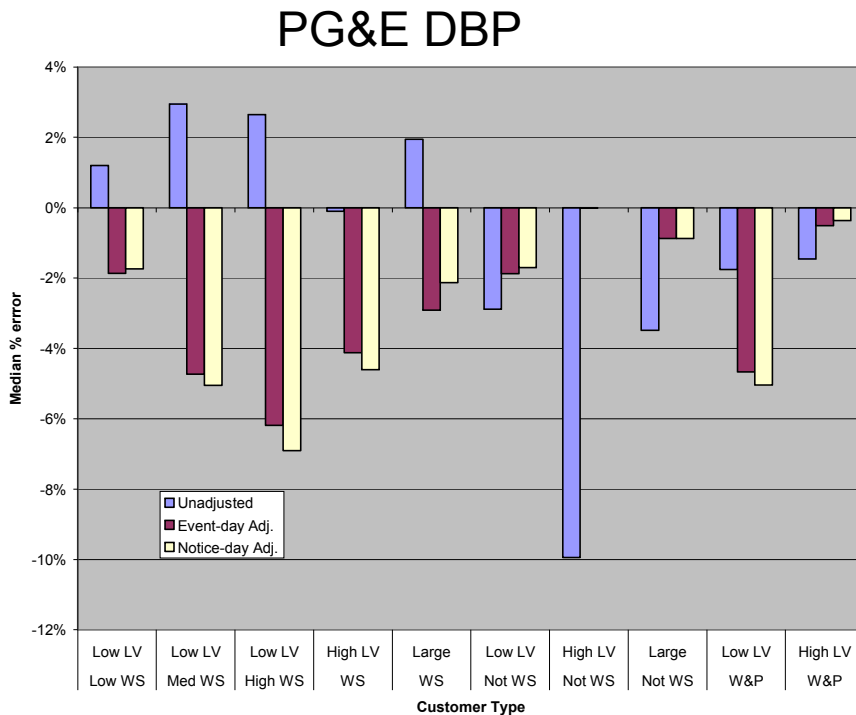
SCE



Similar patterns at PG&E and SCE:

- Most accurate – Low load-variability
- Accuracy somewhat lower as weather sensitivity increases
- Event-day adj. usually improves accuracy more than notice-day

DBP 2006: Unadjusted and Adjusted 3-in-10 – Bias, by Weather Sensitivity & Load Variability

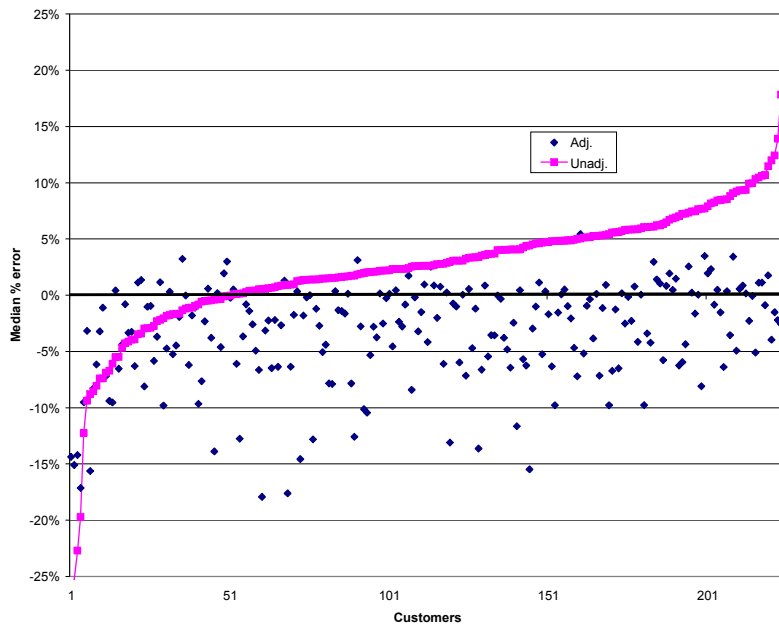


Some major differences between PG&E and SCE:

- Unadj. BL biased downward for WS (PG&E); Biased upward (SCE)
- Upward bias (non-WS) worst for High load variability (Both)
- Adjusted BL shifts errors toward upward bias (Both)
- Greatest improvement from adj. BL for Non-weather sensitive (Both)

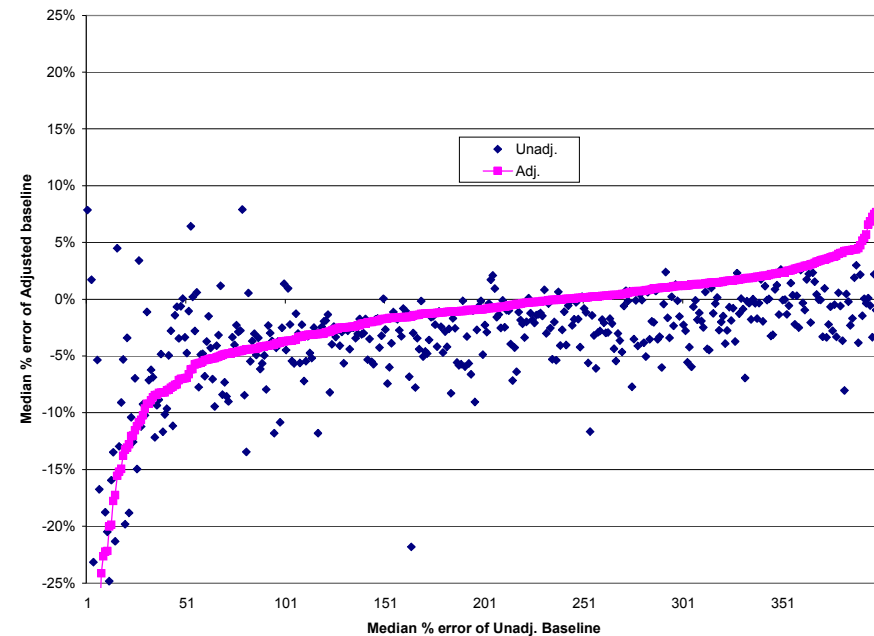
Distribution of % Errors – PG&E and SCE, WS Low-Variability Customers

PG&E



- Unadj. BL biased downward
(More *positive* values)
- Adj. BL shifts errors to mostly
negative (-7% to 3%)

SCE



- Unadj. BL biased upward
(More *negative* values)
- Adj. BL reduces some negative values,
but moves most in negative direction

Explanation of Differences in Bias Results for PG&E and SCE

- Composition of WS group
 - PG&E – Dominated by office buildings
 - Regular loads, strong WS
 - SCE – Dominated by retail stores, shopping centers and supermarkets
 - Less regular loads (sometimes higher on pre-event days than on event days)

Conclusions -- DBP

- ❑ Baseline performance depends greatly on the nature of customers and their loads – in particular *weather sensitivity* (WS) and *load variability* (LV)
 - Greater accuracy for WS
 - Much greater accuracy for low LV than high LV (suggests testing to exclude high LV customers from bidding programs)
- ❑ *Unadjusted 3-in-10* BL showed expected *downward bias* for WS customers for PG&E, but not for SCE
 - Main reason appeared to be major difference in composition of WS DBP customers (offices at PG&E; and retail stores and supermarkets at SCE)
- ❑ *Morning adjustments* generally improved the accuracy of the unadjusted 3-in-10 BL, and shifted the distribution of % errors toward upward bias
 - Adjusted baseline actually improved accuracy more for NWS than for WS customers
- ❑ BL performance varied by *event type* – better performance for isolated events than for second or more in series of sequential events
- ❑ Examining *distributions* of % errors provides insights beyond median values

2008 AMP: Unadjusted & Adjusted Baselines – Accuracy

		<i>Aggregator</i>			<i>Sum of Customers</i>		
Agg.	Level	Unadjusted			Unadjusted		
		3-in-10	5-in-10	10-in-10	3-in-10	5-in-10	10-in-10
1	Total	0.057	0.069	0.092	0.054	0.057	0.091
2	Total	0.065	0.074	0.102	0.055	0.065	0.102
3	Total	0.049	0.056	0.080	0.068	0.052	0.080
4	Total	0.061	0.053	0.049	0.120	0.093	0.049
All	TOTAL	0.056	0.062	0.083	0.075	0.062	0.083

		<i>Aggregator</i>					<i>Sum of Customers</i>				
Agg.	Level	Symmetric Adjustment			Upward-only		Symmetric Adjustment			Upward-only	
		3-in-10	5-in-10	10-in-10	5-in-10	10-in-10	3-in-10	5-in-10	10-in-10	5-in-10	10-in-10
1	Total	0.022	0.023	0.022	0.022	0.022	0.034	0.025	0.027	0.044	0.024
2	Total	0.025	0.028	0.027	0.034	0.030	0.033	0.030	0.026	0.039	0.029
3	Total	0.022	0.021	0.020	0.025	0.020	0.043	0.037	0.034	0.071	0.033
4	Total	0.044	0.039	0.037	0.053	0.037	0.087	0.071	0.041	0.118	0.063
All	TOTAL	0.029	0.028	0.027	0.034	0.028	0.051	0.043	0.036	0.074	0.039

- *Aggregator* BL more accurate than Sum-of-customers
- *Adjusted* BLs more accurate than Unadjusted
- Unadjusted BL less accurate the more days included
- Adjusted BL accuracy similar across # of days
- Upward-only adjustment less accurate than symmetric

2008 AMP: Unadjusted & Adjusted Baselines – Bias

		Aggregator			Sum of Customers		
Agg.	Level	Unadjusted			Unadjusted		
		3-in-10	5-in-10	10-in-10	3-in-10	5-in-10	10-in-10
1	Total	4.42%	5.59%	8.45%	-0.37%	2.57%	8.28%
2	Total	1.39%	3.23%	7.76%	-2.75%	0.75%	7.68%
3	Total	3.51%	4.82%	8.60%	0.89%	3.09%	8.55%
4	Total	0.01%	1.07%	4.14%	-4.70%	-2.71%	4.14%
All	TOTAL	2.47%	3.75%	7.24%	-0.90%	1.55%	7.15%

		Aggregator					Sum of Customers				
Agg.	Level	Symmetric Adjustment			Upward-only Adjustment		Symmetric Adjustment			Upward-only Adjustment	
		3-in-10	5-in-10	10-in-10	5-in-10	10-in-10	3-in-10	5-in-10	10-in-10	5-in-10	10-in-10
1	Total	-0.03%	0.72%	0.97%	0.72%	0.97%	-2.12%	-0.76%	1.51%	-2.81%	0.64%
2	Total	-1.59%	-1.13%	-0.12%	-2.41%	-1.17%	-3.63%	-2.33%	0.56%	-4.49%	-0.51%
3	Total	-0.98%	-0.52%	0.22%	-0.92%	-0.05%	-1.72%	-1.29%	1.37%	-2.75%	0.33%
4	Total	-0.70%	-0.59%	-0.05%	-2.29%	-0.80%	-3.03%	-2.79%	-0.48%	-5.31%	-2.14%
All	TOTAL	-0.71%	-0.36%	0.26%	-1.29%	-0.38%	-2.25%	-1.52%	0.70%	-3.76%	-0.40%

- Aggregator – Unadjusted BL shows *downward* bias (median 2.5% for 3-in-10)
- Downward bias increases w/ number of days included (across columns)
- Adjusted BL shifts distribution to small *upward* bias for 3 and 5-in-10
- Adjusted 10-in-10 appears to have smallest bias for both Agg. & Sum of Cust.

Tests for Gaming Under Adjusted Baseline Option

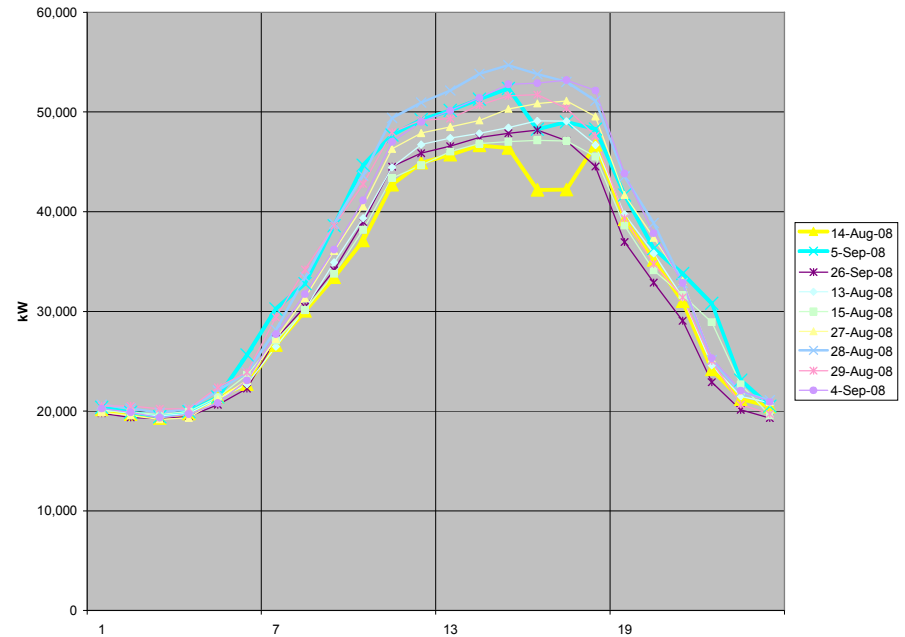
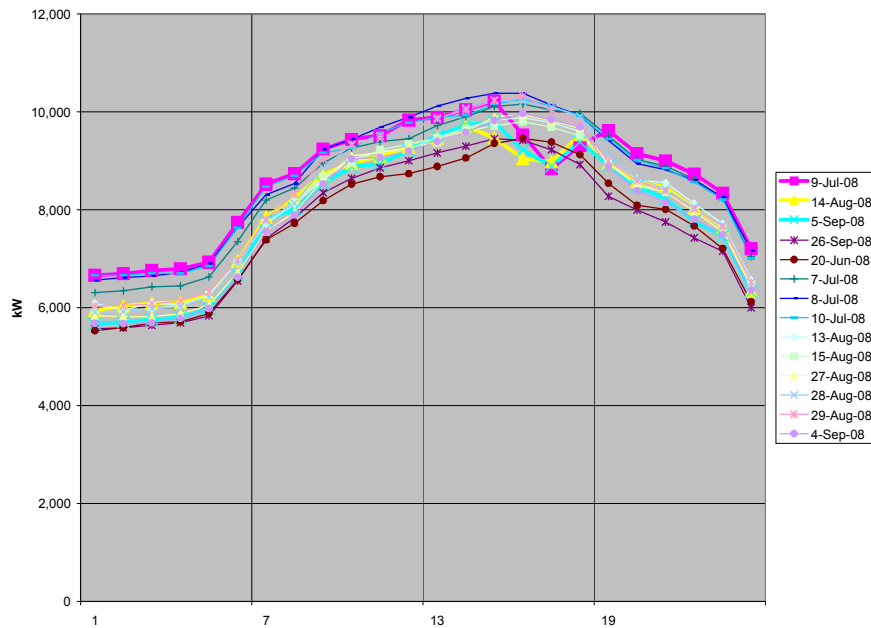
Ratios of Morning Usage on Event & Non-event Days,
by Industry Type and Choice of Adjusted BL

Customer type	Count		Ave. AM kWh - Event/ Non-event		Standard Deviation		Coeff. of Variation	
	No	Adj. BL	No	Adj. BL	No	Adj. BL	No	Adj. BL
1. Ind	193	56	0.98	0.98	0.39	0.38	0.39	0.39
2. Comm'l	94	109	0.99	0.99	0.05	0.18	0.05	0.18
3. Schools	9	6	1.01	1.00	0.18	0.11	0.18	0.11
Grand Total	296	171	0.99	0.98	0.31	0.26	0.32	0.26

- No difference in ave. ratio between adj. & non-adj. BL choice
- More variability in ratio for Industrial vs. Commercial

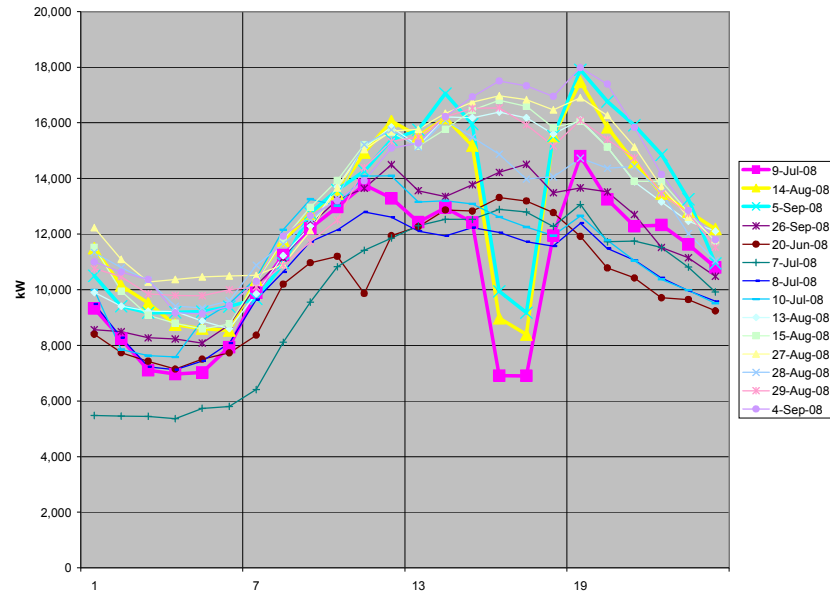
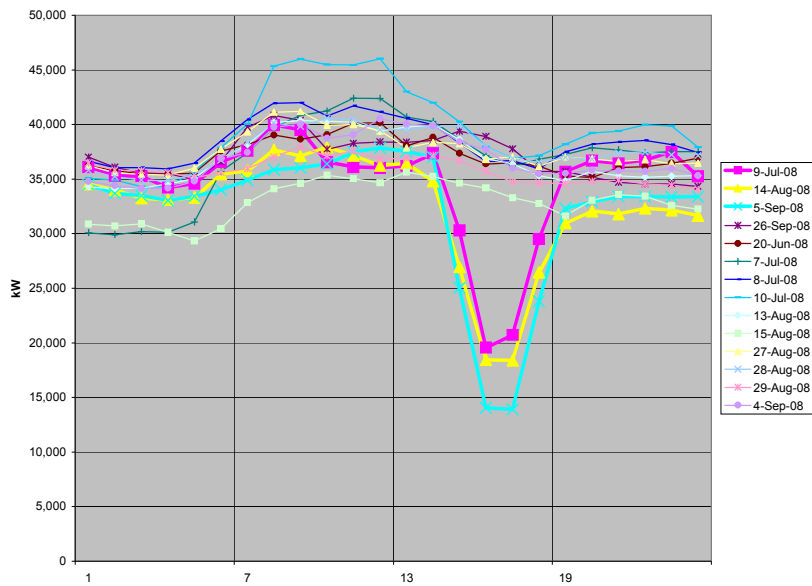
Illustrative Aggregator Loads (Commercial)

– Event Days and Event-type Days



Illustrative Aggregator Loads (Industrial)

– Event Days and Event-type Days



Conclusions -- Aggregator

- ❑ *Aggregator* method was more accurate than sum-of-customers method, though not by wide margin
- ❑ *Morning adjustments* improved the typical downward bias of unadjusted 3-in-10 BL
- ❑ *Adjusted 10-in-10* BL often produced greatest accuracy and least bias, by small margins
- ❑ *Event-day* results were comparable to event-like day findings
- ❑ No evidence found of systematic attempts to “game” the adjusted baseline option