Seeing the (Short-Term) Future: Assessment of Demand Response M&V Baseline Methods

ELIOT CROWE

LAWRENCE BERKELEY NATIONAL LABORATORY



Acknowledgements



Jessica Granderson (PI) Eliot Crowe

Sam Fernandes





Mrinalini Sharma

David Jump

Devan Johnson

Background

2015: Apply test procedure to evaluate proprietary and open source tools

2014: Develop test procedure to assess and compare predictive accuracy of auto-M&V tools

2016: Demonstrate software/methods using historical utility program data

2017: Pilots on 'live projects, transfer test procedure to industry, establish acceptance criteria and practitioner resources

2019- : DR algorithm prediction accuracy, NRE detection, load shift, grid services

Research Questions

How does the advanced M&V regression-based approach compare to the established averaging methods?

Does the duration and timing of DR event window have a significant impact on the prediction accuracy?

Are there notable differences in the distribution of prediction accuracy results across a large population of meters when employing different baseline prediction methods?



Collate a dataset of hourly load and ambient temperature data for commercial buildings

Method

Identify high load days

Use different algorithms to predict hourly load during the hypothetical DR event windows, calculate error metrics

Repeat steps above for all meters in the dataset

Compare the distributions and median error metrics for each algorithm

Algorithms Tested

Algorithm	Variant	Abbreviation
Day-Matching	Unadjusted	DMU
	Adjusted	DMPA
Weather-Matching	Unadjusted	WMU
	Adjusted	WMPA
Time-of-Week-and-Temperature (TOWT)	7-day baseline (no weighting)	UWTOWTU(7.0)
	7-day baseline (no weighting) (adjusted)	UWTOWTPA(7.0)
	70-day baseline (14-day weighting)	UWTOWTU(70.14)
	70-day baseline (10-day weighting)	UWTOWTU(70.10)





Results



Conclusions

How does the advanced M&V regression-based approach compare to the established averaging methods?	All algorithms under-predicted peak period consumption Unweighted TOWT had lowest median bias
Does the duration and timing of DR event window have a significant impact on the prediction accuracy?	No significant difference in bias observed for this data set
Are there notable differences in the distribution of prediction accuracy results across a large population of meters when employing different baseline prediction methods?	Wide distribution in bias values, with significant overlap between algorithms

Thank You



Jessica Granderson Staff Scientist, Deputy for Research Programs



Eliot Crowe Program Manager



Sam Fernandes Program Manager



Berkeley Lab Director's statement on Inclusion, Diversity, Equity, and Accountability (excerpt)

Fostering a diverse workforce - diverse in experiences, perspectives, and backgrounds - and a culture of inclusion are key to attracting and engaging the brightest minds and advancing our record of scientific excellence and groundbreaking innovations.

https://diversity.lbl.gov/directorsstatement/