OAK RIDGE National Laboratory

2019 ASHRAE WINTER CONFERENCE
JANUARY 12–16 // AHR EXPO JANUARY 14–16

Conference Paper Session 1 (Intermediate) – Energy Modeling Simulations

Eric Garrison

University of Tennessee

egarris4@vols.utk.edu

Track: HVAC&R Fundamentals and Applications

Learning Objectives

- Explain how threshold analysis can be used to perform a quality assurance analysis of measured data with missing values
- Apply a distance metric to compare EUI for simulated and real-world buildings
- Explain the building modeling advantages
- Identify the shortcoming of each modeling packages

Acknowledgements

Accuracy of a Crude Approach to Urban Multi-Scale Building Energy Models Compared to 15-min Electricity Use

Co-authors

- Dr. Joshua New
- Mark Adams

Outline/Agenda

- Goals
- Data
- QA and Statistical Analysis
- Distance Metric
- Conclusions

Goals

Through partnership with a utility, the real-world 15-minute electrical data for a service area was analyzed and compared to simulations.

Goals:

1) QA of real-world data.

2) Building type classification algorithm using each building's EUI_{15min}.

Data

- 135,481 buildings in the service area.
- 178,377 premise IDs with 15 minute energy-use data for a full calendar year (00:00 – 00:15 Jan. 1 through 23:45 – 24:00 Dec. 31).
- Each premise ID should have 35,040 energy-use values (they did not).

Table 1. Missing Energy Use Data in Premise IDs

Number of Premise IDs	Missing Data (Nearest %)	Percent of Data (%)
146,318	1	82.03
20,617	2	11.56
11,387	3-99	6.38
55	100	0.03

Only "valid values" were used for each premise ID.

QA and Statistical Analysis

Statistical Information to Track (per premise ID)

- Minimum Energy Use
- Maximum Energy Use
- Naïve/Raw Average
- Naïve/Raw Standard Deviation
- Root Mean Square Error (RMSE)
- Relative Error (RE)
- Absolute Error (AE)
- Threshold Average

Is the data statistically consistent?

QA and Statistical Analysis

Threshold Analysis (Castello et al., 2013)

1) Calculate the naïve average, μ, and standard deviation, σ, of the data to determine the threshold:

$$threshold = \mu \pm c\sigma$$

2) Select a window size, *n*.
3) For each window of size *n*, compute the average of the window for values within the threshold.
4) Find the average of the window averages. This is the threshold average.

This paper uses c = 3 and n = 6.

QA and Statistical Analysis

- Some premise IDs had only 0 EUI_{15min} values (0-vectors).
- Other premise IDs had either 0 or extremely high EUI_{15min} values (spiking).

Table 2. Averages of Statistical Information for Premise IDs

Premise IDs Included	Min.	Max.	Threshold Avg.	Raw Avg.	Raw Std. Dev.	RMSE	RE	AE	Total Removed
All	0.084	3,510.513	1.644	15.088	147.439	67.641	1.153	105,748.077	0
Removing 0-vectors	0.085	3,510.493	1.674	15.351	150.320	68.951	1.176	107,961.451	3,658
Removing 0-vectors and spiking	0.085	194.291	1.573	4.822	21.014	6.253	1.027	19,441.787	3,782

Table 3. Average Data Formatting Statistics for Premise ID Data

Premise IDs Included	Bad Formats	Outliers	Missing	Duplicates	Valid Values
All	8.345	580.588	829.295	37.039	34,202.360
Removing 0-vectors	8.373	592.740	756.888	36.855	34,274.739
Removing 0-vectors and spiking	8.373	593.122	753.170	36.849	34,278.455

Only 2% of premise IDs have unusual energy patterns.

Distance Metric

- Finding EUI_{15min} requires building area.
- Of 178,377 premise IDs, only 48,268 had area available.
- These were compared to EUI_{15min} of ASHRAE prototype buildings using Euclidean Distance (using only valid values).
- 17 ASHRAE prototypes with an average of 6.75 vintages per prototype.
- Of the 17 prototypes, only the PrimarySchool prototype was not represented.

Distance Metric

Table 4. Average Values for Each Building Prototype

Building Type	Distance	Valid Data Points	Load Factor	CVRMSE (%)	NMBE (%)	Total Matches
FullServiceRestaurant	0.322	34,409.412	0.348	96.541	3.535	51
HighriseApartment	0.052	34,585.214	0.214	87.439	-6.297	3,324
Hospital	0.232	34,331.950	0.260	108.868	7.232	558
IECC Residential	0.025	34,508.290	0.164	**531,380.321**	**-477,671.045**	38,380
LargeHotel	0.316	34,229.087	0.303	208.323	8.997	403
LargeOffice	0.102	32,217.706	0.278	87.600	3.980	34
MediumOffice	7.250	34,640.000	0.000	5,262.907	48.370	1
MidriseApartment	0.059	34,299.668	0.219	117.646	-21.999	1,192
Outpatient	16.658	32,595.766	0.377	539.962	13.409	64
QuickServiceRestaurant	**1,433.371**	34,085.328	0.402	604.051	48.724	305
RetailStandalone	0.101	26,219.800	0.294	111.854	3.926	5
RetailStripmall	0.108	33,985.231	0.215	108.856	5.978	26
SecondarySchool	1.886	20,618.000	0.002	1,880.802	18.612	1
SmallHotel	0.085	34,709.770	0.197	97.889	3.693	3,751
SmallOffice	0.200	18,784.000	0.190	302.404	14.007	4
Warehouse	0.096	26,383.550	0.191	2,664.157	-2,070.092	169

There are several noticeable outliers.

Distance Metric

Table 5. Average Values for QuickServiceRestaurant by Vintage.						
Building Vintage	Distance	Valid Data Points	Load Factor	CVRMSE (%)	NMBE (%)	Total Matches
90.1-2004	0.540	33,794.500	0.156	134.236	-0.153	2
90.1-2007	0.303	34,611.941	0.426	66.880	11.569	17
90.1-2010	0.253	34,826.000	0.393	65.032	9.258	1
90.1-2013	0.314	34,626.000	0.185	100.160	3.804	2
DOE-Ref-1980-2004	0.280	34,158.122	0.429	59.015	11.915	41
DOE-Ref-Pre-1980	**1,806.445**	34,030.876	0.399	744.402	58.509	242

- QuickServiceRestaurant has the highest EUI_{15min} of all prototypes with the DOE-Ref-Pre-1980 vintage being the highest of its prototype.

- IECC Residential has the lowest EUI_{15min}. Nearly 80% of premise IDs match to IECC Residential, including 0-vectors and most spiking premise IDs.
- Removing 0-vectors and spiking premise IDs brings CVRMSE to 259.338% and NMBE to 14.423%.
- There are no ASHRAE guidelines for 15-minute data, but these values seem like a reasonable start.

Conclusions

- Statistical analysis of real-world, 15-minute data determined nearly 98% of premise IDs were reasonably representative and could be used as a baseline to compare to simulated data.

- For the 47,684 premise IDs where EUI_{15min} could be computed, the error of crude building energy models over the year is likely within the error rates manually created buildings of previous studies.

Bibliography

Castello, Charles C., New, Joshua R., and Smith, Matt K., 2013. "Autonomous Correction of Sensor Data Applied to Building Technologies Using Filtering Methods." In *Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Austin, TX, December 3-5, 2013.

Questions?

Eric Garrison

egarris4@vols.utk.edu

Dr. Joshua New

newjr@ornl.gov

Mark Adams

adamsmb@ornl.gov