

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

Eric Garrison

Joshua New, PhD
Member ASHRAE

Mark Adams
Member ASHRAE

ABSTRACT

ASHRAE has recently named a TC4.7 subcommittee on Multi-Scale Energy Modeling and there have been many technical developments focused on urban-scale energy modeling with the goal of creating building energy models for all buildings in a city to facilitate simulation-informed energy optimization for the built environment. These efforts often rely on non-scalable, local sources of information (e.g. county's tax assessor's database) to create more accurate models of those buildings. In partnership with the local utility, a year's worth of real-world 15-minute electrical use data from 178,377 premise buildings has been assessed. This smart meter data is statistically analyzed and compared to corresponding simulation data for 97 different building prototypes in order to better assess simulation accuracy for various utility-defined use cases. To our knowledge, this is the first time real world and simulated building energy use data have been compared at this scale.

INTRODUCTION

There are approximately 124 million buildings in the United States which consume approximately \$395 billion per year in energy bills. This constitutes 45% of national primary energy consumption and 73% of all electrical consumption. The U.S. Department of Energy's (DOE) Building Technologies Office has the goal of reducing energy use intensity of U.S. buildings 45% by 2030 compared to a 2010 baseline. While the U.S. Energy Information Administration provides some top-down statistics to guide a measure of success against those goals, utilities – which hold the high-value building-specific data – are frequently conservative in sharing that data due to legal obligations to keep the energy use data of their customers secure. Finding a secure way to recognize value for themselves and their rate payers would be very attractive to most utilities. The authors have partnered with a utility to create a digital twin of a utility's service area and compare their built environment to 15-minute electrical consumption from revenue grade meters.

In partnership with the Electric Power Board of Chattanooga (EPB), we have analyzed a full year's worth of raw 15-minute electrical energy use data gathered by utility-grade smart meters for 178,377 premise identifiers (premise IDs) in the Chattanooga, TN area. The premise ID is an integer value representing a building in EPB's business systems and must be mapped to the EPB building ID corresponding to the building energy model created by the

Eric Garrison is a graduate student in the Electrical Engineering and Computer Science Department of The University of Tennessee, Knoxville, TN. Joshua New and Mark Adams are R&D Staff employed by Oak Ridge National Laboratory, Oak Ridge, TN.

authors. These values are different in order to obfuscate the energy use values from a given building energy model. However, premise IDs do not map 1-to-1 with building models—there are 178,377 distinct premise IDs but only 135,481 building models detected from current data sources. The energy use values associated with a premise ID can represent a building’s energy use values for an entire year or only a subset of the year. This means multiple premise IDs will often map to a single building ID. While EPB has global positioning system (GPS) coordinates for all electric meters that can be used to map their billing data to GPS-registered building energy models, there remains an issue in correcting GPS coordinates or properly registering 24% of the electrical meters using this crude modeling approach, so premise IDs cannot all be combined or linked to their respective building ID.

As it comes from real-world sensors, the energy use data for each premise ID needs to be checked for possible errors. From the raw 15-minute energy use data, the minimum, maximum, average, and standard deviation of energy use were recorded for each premise ID. Then, threshold analysis [1] was applied to each premise ID’s energy use values to estimate its average 15-minute energy use over the year. The estimated average was used to calculate the relative error (RE), absolute error (AE), and root mean square error (RMSE) between this threshold average and the reported values. From here, outliers in the data could be detected.

After analyzing the statistical data for each premise ID, a subset of premise ID energy use vectors were normalized by their building square footage to get their 15-minute energy use per square foot for the year – also known as the energy use intensity (EUI_{15min}). Only 27% (48,268) of the premise IDs had a floor area registered with the utility. These normalized premise ID energy use vectors were compared 97 different simulated buildings (residential and commercial building types and vintages) to determine which energy use profile most closely matched the energy use per square foot of a given premise ID over the entire year. Once a comparable prototype building was determined, the coefficient of variation of the root mean square error (CVRMSE) and normalized mean bias error (NMBE) were calculated.

The contributions of this paper are two-fold: 1) a quality assurance analysis was performed on the raw 15-minute electrical energy data of 178,377 premise IDs and was used to improve utility data, and 2) a building type classification algorithm based on 15-minute energy consumption profiles was assessed for 48,268 premise IDs using 97 different simulated buildings.

PREVIOUS WORK

EnergyPlus [2] and OpenStudio [3] are the building physics simulation engine and middleware software development kit in which DOE has invested \$93 million since 1995. As part of this work, new features, better runtime performance, and creation of new prototype models of building types from assessments of real buildings has been constructed [4][5] to facilitate accurate and timely modeling of the U.S. building stock. In an effort to retain individual building information but at larger, city-scale geographical areas, ASHRAE has hosted 8 seminars focused on Urban-Scale Energy Modeling [6][13]. These efforts by government, universities, and industry has focused on leveraging increasingly diverse data sources and algorithmic capabilities to automatically construct individual building energy models in a scalable, compute-driven manner [14].

EXPERIMENTS

Data

The original electrical energy consumption data came from 15-minute measured electrical use for each building in the EPB service area over an entire calendar year. The raw data given from EPB is approximately 50 gigabytes, has several million tuples consisting of a 15-minute time period, electrical energy use values in kW consumed during each 15-minute time period (the sum of four consecutive values provides the kWh), and an associated premise ID. The data was reorganized by premise ID and ordered chronologically to facilitate comparison to the energy use data of simulated buildings.

As expected with this volume of raw data, there are several issues involved with data formatting. First, not all 15-minute intervals have energy use values for many of the premise IDs. Second, some data is incorrectly formatted, either through an incorrectly formatted date/time string or a non-numeric value for energy use. Third, some customers change rate structures or owners which may result in a mid-year change of premise ID. Finally, some of the 15-minute intervals for the same premise ID have duplicates in the raw data. The result of these issues means that most premise IDs are missing at least some of their energy use values and some premise IDs may be missing almost all of their energy use values. Ideally, each premise ID would have a complete set of 15-minute energy values for the entire year (35,040 in total). Most premise IDs have only a small amount of missing data, with over 93% of the premise ID energy use vectors missing 2% or less of their energy use values. A breakdown of missing data is shown in Table 1.

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

Each premise ID had its energy use values stored in a 35,040-element vector for the annual data. For example, January 1, 00:00-00:15 is index 0, January 1, 00:15-00:30 is index 1, and so on. Missing or badly formatted data is represented by a token ‘NaN’ rather than a zero value in order to differentiate between actual zero values and missing or badly formatted data.

Prototype Vectors

All commercial prototype buildings and a group of prototype residential buildings were simulated for every vintage using climate zone ASHRAE-169-2006-4A building codes and simulated using Actual Meteorological Year (AMY) weather data for the calendar year corresponding to the measured data. Each prototype building’s 15-minute energy use was simulated for 17 different building types including Full Service Restaurants, Hospitals, Office buildings (small, medium, and large), Primary Schools, and more. Each building type uses several different vintage types averaging 6.75 vintages per building type. A total of 97 different prototype vectors were used. Based on this simulated data, the energy use profile is used to determine building type from actual, building-specific energy use intensity.

Statistical Analysis

Real-world data can be inaccurate or contain errors, through recording errors with equipment, transcription errors, or from several other sources. To get an idea of how likely the real-world data is to contain such errors, statistical information about the data can be computed and analyzed. Although it is not possible to know whether or not the real-world data is accurate (because there is no ground truth against which it can be compared), it is important to know how likely the data represents its real-world counterpart before significant time is spent comparing it to models.

Based on previous work, threshold analysis [1] is a more accurate way to determine an average of a series of data with missing values than several other methods. Threshold analysis works by finding the average of sliding averages of data within a certain threshold over the data set. First, the naïve average, μ , and standard deviation, σ , of the full data set is calculated. From these, a user selected value ϵ is chosen to create a “threshold” for valid values, where

$$threshold = \mu \pm \epsilon \sigma \tag{1}$$

For a particular window size n (determined by the user), the average of the first x_1, x_2, \dots, x_n values is calculated. Only values inside the threshold are used to calculate this average—values outside the threshold are ignored. Next, the window shifts by one value, and the average of the next x_2, x_3, \dots, x_{n+1} values (within the threshold) is calculated. This process repeats until the averages of all windows have been calculated. Finally, the average of the window averages is the threshold average of the series.

The threshold analysis ($c = 3, n = 6$) was performed as described in [1] with a modification to properly accommodate questionable data; since values are missing at unpredictable locations, the sliding window size cannot be guaranteed to always contain all valid values. A minor change was made to average only valid values. With a window size of six (i.e., 1.5 hours if all data is present), if two values in the window are missing or invalid, the mean of only the four valid values are taken as the mean of the window. Windows with no valid values are not included in the threshold average. This process repeats every time the window shifts. The data was split with 70% of valid values used for training and 30% for testing.

Once the threshold average had been calculated for a given premise ID, an average energy use vector was created with each value being the threshold average. The AE (equation 2), RE (equation 3), and RMSE (equation 4) were calculated between each premise ID energy vector and its average energy use vector. Again, only valid data points were used to calculate the error measures. In addition, the following metrics were recorded from each premise ID energy use vector: minimum and maximum energy use values, number of badly or incorrectly formatted data, number of missing data points, number of duplicate data points, number of correctly formatted data points, naïve average and standard deviation, and the threshold average. As discussed in the results section, these measures helped give confidence that most of the real-world data did not contain errors.

Euclidean Distance

In addition to analyzing each building's actual electrical energy use using statistical information, the primary goal was to determine how similar each premise ID energy use vector was to each of the simulated prototype vectors so that building type could be assigned based on the electrical signal (i.e., the one that matters most for utilities). A naïve Euclidean distance metric was employed to compare the similarity of each premise ID's actual EUI_{15min} to that of each prototype building. Any 15-minute intervals without an energy use value for the premise ID were excluded from the distance calculation (each prototype vector had energy use values for all 35,040 time intervals).

Of the 178,377 initial premise IDs, only 48,268 could be mapped to a building area, which is required to find EUI_{15min} and measure the Euclidean distance. The prototype building with the smallest distance was considered to be its match and used to assign the premise ID's building type based on the closest prototype/vintage EUI. For each of the 48,268 premise IDs, CVRMSE (equation 5) and NMBE (equation 6) were calculated between it and its matched prototype building.

Utilities are very cost-sensitive to the time-varying cost of energy as energy is more expensive during peak energy use times (esp. peak demand during critical generation times.) A common measure of this is the ratio of each building's total energy use to the maximum continuous energy use, known as load factor (equation 7). Once the building type was assigned as described above, the load factor for the building was computed. For utilities, this could be used to inform operational decisions for the electrical distribution network, such as moving transformers to handle larger loads.

$$AE = \sum_{i=1}^n |y_i - \bar{y}_i| \quad (2)$$

$$RE = \frac{\sum_{i=1}^n \left| \frac{\bar{y}_i - y_i}{y_i} \right|}{n} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n}} \quad (4)$$

$$CVRMSE = 100 \times \left[\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n - p} \right]^{1/2} / \bar{y} \quad (5)$$

$$NMBE = 100 \times \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{(n - p) \times \bar{y}} \quad (6)$$

where n is the number of valid values in the premise ID energy use vector, y_i are the valid premise ID energy use values, \hat{y}_i are the corresponding prototype energy use values, \bar{y} is the mean of the valid values in the premise ID energy use vector, and p is 1.

$$Load\ Factor = 100 \times \frac{kWh}{(Peak\ kW \times \# \text{ hours in billing period})} \quad (7)$$

RESULTS

All work for implementing the experiments described above was written in Python scripts for preprocessing, statistical analysis, and Euclidean distance calculations. All code was run on a secure internal, 24-core Linux server.

Statistical Analysis

For the 178,377 premise ID buildings, the previously-described statistical data was calculated and threshold analysis run to determine error measures, which were then used to look for outliers. Initial analysis revealed several outlier premise IDs that warranted further investigation.

A total of 3,658 premise IDs had minimum and maximum energy use values of 0, indicating that these premise IDs had zero vectors for energy use and were likely not representative of an actual building. Additionally, several premise ID energy use vectors were very close to a zero vector with only a few non-zero 15-minute energy use values, and these values spiked into the tens of thousands. With all energy values measured in kWh, these high values in this pattern seem unrealistic. Average statistical data for the premise ID energy use vectors was calculated for several cases: using the data from all premise IDs, using only premise IDs with non-zero vectors, and using premise IDs with non-zero vectors without spiking. Premise IDs with spiking were defined as having both a maximum kWh value of 10,000 and a maximum value more than 1,000 times the threshold average. This definition does not remove all premise IDs with high outlier values, but testing different thresholds shows that these values remove the most obvious outliers. The statistical data was averaged for each of these cases and values are shown in Table 2.

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

It is interesting to note that removing the premise IDs with zero vectors did not significantly change the average statistical information of the premise IDs, which indicates that many of the premise IDs with non-zero data have an average energy use close to a zero vector. Only 124 premise IDs (0.06%) were filtered with the spiking criteria but caused a dramatic change in almost every measure except the threshold average and minimum values (which should be expected). These dramatic decreases are best explained by how significantly the maximum energy use of these spiking premise IDs lies outside of standard energy use. This gives support to the intuition that the spiking premise IDs are not likely accurate representations of real energy usage.

The averages of formatting errors were also tracked, shown in Table 3. The most significant changes came from removing the zero vectors, which reduced the average number of outliers and greatly reduced the average number of missing values.

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

From this analysis, only 2% of premise IDs seemed to have unrealistic energy patterns. Enough of the data was useful that it could be included in the next part of the analysis.

Euclidean Distance

Comparing the real-world data to simulated data in a meaningful way requires normalizing the energy use of each building. Of the initial 178,377 premise ID buildings, a total of 48,268 had square footage data available, which was required to normalize the premise ID energy use vectors and compare with the prototype energy use vectors. The number of premise IDs matched with each building type and the averages of several metrics are displayed in Table 4.

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

Of the 17 possible building prototypes, only PrimarySchool is not represented. The MediumOffice and SecondarySchool building prototypes only have one match. The number of primary schools, secondary schools, and medium offices (the most common commercial building type) in reality is much larger than what was found with this building type assignment. While any energy conservation measures applied to inappropriately labeled building types would be physically unrealistic, it is important to note that simply matching the EUI profile for load factors is beneficial for the utility use cases driving this study.

Nearly 80% of the 48,268 premise IDs matched with the IECC Residential building type. The IECC building types have significantly lower EUI than any of the other building prototypes. As noted in the statistical analysis, some spiking premise ID vectors have over 35,000 energy use values of zero and only one non-zero value. This contributes to large average CVRMSE and NMBE errors and can cause their individual CVRMSE to be in the millions. The low EUI_{15min} of these premise IDs match with the IECC building type. For the statistical analysis, thresholds were set to filter premise IDs with spiking energy use. Extreme CVRMSE values are the result of more factors than maximum energy use, so a different filter was needed to remove unrealistic premise IDs. Filtering out CVRMSE greater than 10,000 removes 584 premise IDs and reduces the CVRMSE and NMBE for IECC buildings to 259.338% and -144.423% respectively. Although this filter is not optimal, testing has shown that it is a good starting point and can be optimized in future work.

While ASHRAE Guideline 14 [18] specifies CVRMSE<30% and NMBE<10% for hourly data for a calibrated model, existing reports covering over 3,000 buildings show manually-created building energy models have monthly CVRMSE between 23% and 97% [19][20]. For either calibrated data or preliminary manual BEM creation, there is not yet an equivalent 15-minute data. If CVRMSE and NMBE quadruple with the number of data points (i.e., 15-minute instead of hourly), these comparisons are within range of those typically constructed manually. As such, the Euclidean distance method is adequate for this application.

The largest average distance is from the QuickServiceRestaurant (QSR) building type, which is nearly 90 times larger than the second largest average distance, the Outpatient building type. The QSR distance is heavily weighted by several premise IDs whose distance values were measured in the thousands. The best explanation for this is that the QSR has the highest total EUI of all building prototypes over the course of a year. Any premise IDs with higher EUI than any of the prototype buildings will match with the QSR. Of the 305 premise IDs matching the QSR, 242 match with the DOE-Ref-1980-2004 vintage, which is the prototype building type/vintage combination with the highest EUI over the year. This indicates that at least a subset of the 242 premise IDs matching the indicated vintage are not well represented by any of the 97 prototype buildings. A breakdown of metrics by vintage type for the QSR is in Table 5.

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

CONCLUSION

A quality assurance analysis was performed on raw 15-minute electrical energy data of 178,377 premise IDs in the Chattanooga, TN area. Statistical information was analyzed to determine the usefulness of the data and whether individual premise IDs were likely to represent actual buildings. The analysis determined that nearly 98% of premise IDs were a reasonably representative and could be used as a baseline to compare with simulated data.

In partnership with a utility, actual 15-minute electrical whole-building energy use from each of 178,377 premise IDs was compared to 97 prototype buildings (different combinations of building types and vintages) using Euclidean distance to assign each premise ID a vintage and building type based on the most similar 15-minute energy use profile. For 47,684 premise IDs which met the data availability and statistical requirements, crude building energy models had a 15-minute error added over the entire year which is likely within range of error rates from previous studies for manually-created building energy models.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability, under contract number DE-AC05-00OR22725. This work was also funded by field work proposal CEBT105 under DOE Building Technology Office Activity Numbers BT0201000. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC05-00OR22725.

Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the US Department of Energy under contract DE-AC05-00OR22725. This manuscript has been authored by UT-Battelle, LLC, under Contract Number DEAC05-00OR22725 with the US Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

REFERENCES

- [1] 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.
- [2] U.S. Dept. of Energy (2011). DOE releases new version of EnergyPlus modeling software. Available: <https://energyplus.net/downloads> and <https://github.com/NREL/EnergyPlus>
- [3] Roth, Amir, Brackney, Larry, Parker, Andrew, and Beitel, Annette (2016). OpenStudio: A Platform for Ex Ante Incentive Programs. Available: <https://www.osti.gov/scitech/biblio/1324396>, <https://www.openstudio.net/downloads> and <https://github.com/NREL/OpenStudio>
- [4] Malhotra, Mini, New, Joshua R., and Im, Piljae (2018). "Prototype Courthouse Building Energy Model: Building and System Characteristics." ORNL internal report ORNL/TM-2017/2, February 2018, 79 pages. [PDF]
- [5] Goel, S. et al. (2014). "Enhancements to ASHRAE Standard 90.1 Prototype Building Models." Pacific Northwest National Laboratory Internal Report PNNL-23269, April 2014. [PDF] [website]
- [6] Cerezo, Carlos, Heo, Yeonsook, and Joshua New (*Seminar Chair*) (2016). "Seminar 39 - Data Sources toward Urban-Scale Energy Modeling, Part 1." In *Proceedings of the ASHRAE Annual Conference*, St. Louis, MO, June 28, 2016. [MIT] [UK]
- [7] Brown, Jason, Allen, Melissa, Scheer, David, and Joshua New (*Seminar Chair*) (2016). "Seminar 56 - Data Sources toward Urban-Scale Energy Modeling, Part 2." In *Proceedings of the ASHRAE Annual Conference*, St. Louis, MO, June 29, 2016. [GATech] [ORNL] [Autodesk]
- [8] Hong, Tianzhen, Muehleisen, Ralph, Long, Nicholas, and Joshua New (*Seminar Chair*) (2017). "Seminar 43 - Urban-Scale Energy Modeling, Part 3." In *Proceedings of the ASHRAE Winter Conference*, Las Vegas, NV, January 31, 2017. [LBNL] [ANL] [NREL]

- [9] Allen, Melissa, Bobker, Michael, Khan, Haider, Crawley, Drury, and Joshua New (*Seminar Chair*) (2017). "Seminar 55 - Urban-Scale Energy Modeling, Part 4." In *Proceedings of the ASHRAE Winter Conference*, Las Vegas, NV, January 31, 2017. [[ORNL](#)] [[CUNY](#)] [[ICF](#)] [[Bentley](#)]
- [10] Muehleisen, Ralph, Crawley, Drury, and Joshua New (*Seminar Chair*) (2017). "Seminar 55 - Urban-Scale Energy Modeling, Part 6." In *Proceedings of the ASHRAE Annual Conference*, Long Beach, CA, June 28, 2017. [[ANL](#)] [[Bentley](#)]
- [11] New, Joshua R., Chen, Yixing, Choi, Joon-Ho, and Bass Abushakra (*Seminar Chair*) (2017). "Seminar 28 - Urban-Scale Energy Modeling, Part 5" presenting "Automatic Building Energy Model Creation (AutoBEM)." In *Proceedings of the ASHRAE Annual Conference*, Long Beach, CA, June 26, 2017. [[ORNL](#)] [[LBNL](#)] [[USC](#)]
- [12] Leung, Luke, Phillips, Duncan, and New, Joshua R. (*Seminar Chair*) (2018). "Seminar 27 - Urban-Scale Energy Modeling, Part 7." In *Proceedings of the ASHRAE Winter Conference*, Chicago, IL, January 22, 2018. [[RWDI](#)] [[SOM](#)]
- [13] Nagpal, Shreshth, Hong, Tianzhen, Cox, Matt, and New, Joshua R. (*Seminar Chair*) (2018). "Seminar 16 - Urban-Scale Energy Modeling, Part 8." In *Proceedings of the ASHRAE Conference*, Houston, TX, June 23, 2018. [[MIT](#)] [[LBNL](#)] [[GLG](#)]
- [14] Reinhart, C. and Davila, C. (2016). Urban building energy modeling – A review of a nascent field. *Building and Environment*, 97, 196-202. [[DOI](#)]
- [15] New, Joshua R., Omitaomu, Olufemi, Yuan, Jiangye, Yang, Hsiuhan (Lexie), Carvalhaes, Thomaz, Sylvester, Linda, and Adams, Mark (2017). "AutoBEM: Automatic Detection and Creation of Individual Building Energy Models for Each Building in an Area of Interest." In *Proceedings of the 2nd International Energy and Environment Summit*, Dubai, UAE, November 18-20, 2017. [[PDF](#)] [[PPT](#)]
- [16] New, Joshua R., Adams, Mark, Im, Piljae, Yang, Hsiuhan, Hambrick, Joshua, Copeland, William, Bruce, Lilian, Ingraham, James A. (2018). "Automatic Building Energy Model Creation (AutoBEM) for Urban-Scale Energy Modeling and Assessment of Value Propositions for Electric Utilities." In *Proceedings of the International Conference on Energy Engineering and Smart Grids (ESG)*, Fitzwilliam College, University of Cambridge, Cambridge city, United Kingdom, June 25-26, 2018. [[PDF](#)] [[PPT](#)]
- [17] Ingraham, James A. and New, Joshua R. (2018). "Virtual EPB." Presented to Building Technologies Office following the *BTO Peer Review*, 87 slides. Arlington, VA, May 3, 2018. [[PPT](#)]
- [18] ASHRAE (2014). Measurement of Energy and Demand Savings. ASHRAE Guideline 14-2014. [[PDF](#)]
- [19] Earth Advantage Institute and Conservation Services Group, Energy performance score 2008 pilot: Findings and recommendations report.
- [20] D. Roberts, N. Merket, B. Polly, M. Heahey, S. Casey, J. Robertson, Assessment of the U.S. department of energy scoring tool, Tech. rep., National Renewable Energy Laboratory (2012).