First Steps to Maintain a Large Fleet of Building Energy Models

Daniel L. Villa, P.E. *Member ASHRAE*
Joshua R. New, PhD *Member ASHRAE*
Mark Adams *Member ASHRAE*
Aaron Garrett, PhD

Gerald R. Gallegos, P.E. *Member ASHRAE*

ABSTRACT

Continuous maintenance of large fleets of building energy models (BEMs) is a new challenge that may soon approach economic parity for large institutions as data and modeling connectivity become streamlined. In the past, BEMs have mostly been used for analyses during design with little or no reuse of the model. The Energy Independence Security Act (EISA) of 2007's requirement to complete energy and water evaluations for federal facilities is changing this. EISA requirements can be met through performance of ASHRAE energy audits that allow BEMs to be used for identifying energy savings opportunities and energy use breakdowns. Several years ago, Sandia National Laboratories (SNL) developed a fleet of 121 BEMs for site-wide energy assessments. Applications for this fleet has now been expanded to EISA compliance. The authors propose maintaining the BEM fleet on a 4-year cycle. In this process models undergo a quality check (QC) and recalibration whenever a building energy audit is performed for its corresponding building. For recalibration, auto-calibration technology is being used. This paper outlines the first year of efforts to construct a streamlined process and results for the first 5 models. The first BEM underwent both manual and auto-calibration for a direct comparison. Manual calibration incurred more cost and required more time from staff members. Auto-calibration required significantly less effort, slightly less cost, and resulted in slightly better accuracy. ASHRAE Guideline 14-2014 was met by 3 of the 5 buildings after auto-calibration. Even so, significant improvements to Normalized Mean Bias Error (NMBE) and Coefficient of Variation for Root Mean Square Error (CV(RSME)) were achieved for all five buildings. Data and model quality issues are suspected as causes for the non-compliance rather than inadequacy of the auto-calibration procedure. Many improvements to the processes used to prepare data and models have been identified including issues that require major changes to SNL's energy tracking infrastructure.

INTRODUCTION

Potential markets for Building Energy Models (BEM) are expanding and rapidly changing (Hong et. al., 2018). This includes applications for large fleets of BEM on existing buildings (Villa et. al., 2017). Meanwhile, existing energy efficiency practice's profitability may be slowing (Stuart et. al., 2018) and may benefit from innovative uses of BEMs. Many efforts are underway to automatically generate entire campuses and cities through urban-scale building energy modeling techniques (Nagpal and Rienhart, 2018; ORNL 2018; NREL, 2018; Chen et. al., 2017; Reinhart and Davila, 2016) yet such efforts are mostly accurate on a larger scale than individual building energy assessments. In comparison to urban scale planners, institutional planners often need asssessments for large fleets of buildings with accuracy for each building. Institutions also often have access to data that has the potential to keep BEM accurate. Institution-scale research efforts therefore require focus on the maintenance of large fleets of BEM that leverage automation as much as possible to reduce modeling efforts.

SNL created 121 DOE 2.2 BEM from 2012 to 2017 for site-wide assessment of Energy Conservation Measures

Daniel Villa and **Gerald Gallegos** are technical staff members at Sandia National Laboratories Albuquerque, NM. **Joshua New** is founder and CEO of Tunation, LLC. Knoxville, TN. **Mark Adams** and **Aaron Garrett** are employees of Tunation, LLC.

(ECM). The results of these analyses were used to estimate the feasibility of institution-wide energy efficiency goals (Villa et. al., 2017). These site-wide assessments are only needed every couple of years. One new application starting in 2018 has been using the fleet as a resource for compliance to the Energy Independence and Security Act of 2007 (EISA) through energy-use breakdown estimates (Fisher, 2014). Legal compliance to EISA section 432 (EISA, 2007) requires applicable federal facilities to undergo energy audits every four years. EISA compliance can be met by the American Society of Heating, Refrigeration and Air-conditioning (ASHRAE) commercial building energy audit levels one to three (ASHRAE, 2011). To do this, the models must produce estimates of the energy and water savings opportunities for each building. Unfortunately, an audit of SNL's fleet in 2016 found only 52% of the models met ASHRAE Guideline 14 calibration requirements (Villa et. al., 2017). Even though calibration status is not an exact indicator, this result probably shows that energy savings opportunities are inaccurate for the fleet. This led to planning for maintaining SNL's fleet of models through a systematic process. The need for methods to reduce workloads for 121 models suggested automated calibration techniques might be helpful (Chaudhary et. al., 2016). This in turn led to collaboration for auto-calibrating the fleet of models. Manual intervention was also planned through quality checks of the model using energy audit reports (Villa et. al., 2018). This paper presents the design of a quality check and automatic calibration procedure on our BEM fleet that is applied after building energy audits. It then walks through the results for the first five buildings involving both quality checks and auto-calibration.

METHODS

The preparation of models for auto-calibration is challenging. To reduce the need for modelers to insert new expressions in every model, software was developed to automate the process. The DOE2.2 BEMs already had many Building Design Language (BDL) expressions and global parameters for ECMs but these were not designed to touch every part of the model needed for calibration purposes. Generalized BDL expressions were therefore derived that produce no change to the original BDL expressions if calibration parameters are default values. Two types of parameters were inserted by the software: multipliers and base-load offset parameters (Villa et. al., 2018). Multipliers have a default numerical value of one and output the product of the multiplier and any existing BDL expression. Base-load offset parameters are applied to schedules and have a default value of zero. At negative one, the baseload offset stretches the schedule to zero baseload. At one, the baseload offset compresses the schedule to baseload equal to peak load as seen in Figure 1. The functional relationship to transform any function this way is given in equations (1) and (2).

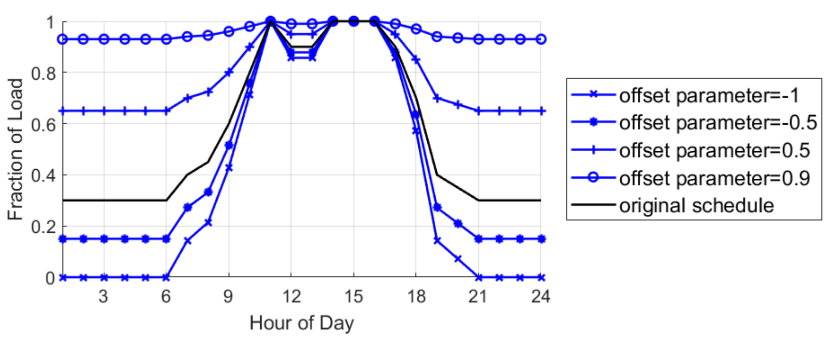


Figure 1. Baseload offset parameter effects on an hourly load schedule for one day as defined by equation (1).

$$f_{new}(t) = \begin{cases} (1-p_o)f(t) + f_{max}p_o & 1 \geq p_o \geq 0 \\ -p_o f_{new_{-1}}(t) + (p_o+1)f(t) & -1 \leq p_o < 0 \end{cases} \tag{1}$$

$$f_{new_{-1}}(t) = \frac{f_{max}}{f_{max}-f_{min}}f(t) - \frac{f_{max}f_{min}}{f_{max}-f_{min}} \tag{2}$$

Here, p_o is the base-load offset parameter used by the calibration algorithm, $f_{max} = \max(f(t))$, $f_{min} = \min(f(t))$.

Sixteen multipliers and five base-load fraction parameters were designed to alter building envelope, HVAC cooling efficiency, equipment schedules, exhaust schedules, fan schedules, heating efficiency, cooling efficiency, heat rejection efficiency, infiltration, lighting schedules, occupancy schedules, outdoor air flow, plug loads schedule, pumping efficiency, and non-electrical source load schedules (Villa et. al., 2018). The underlying software applies the parameters across a broad range of systems and BDL command types. Over fifty-five BDL keywords per building are referenced by a total of 255 tunable calibration parameters (13 for Building 1, 56 for Building 2, 49 for Building 3, 70 for Building 4, and 67 for Building 5). The changes to BEM from insertion of these parameters are extensive with two examples shown in Figure 2. The lower example illustrates that existing expressions are preserved. There were approximately 49 previously defined parameters that were leveraged for site-wide energy assessments when available (number of parameters varied by building). Even though it was intended for our software to not alter model output with default parameter values, some changes in predicted energy were observed. Investigation concerning why this is the case is still underway but BDL's handling of default and missing values is probably the issue. These differences were deemed sufficiently small to neglect. The resulting software can prepare any DOE2.2 model for auto-calibration.

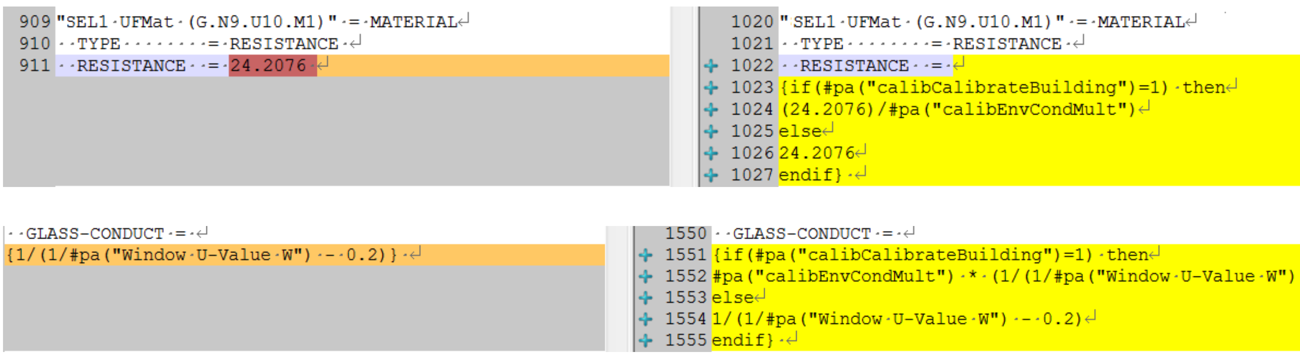


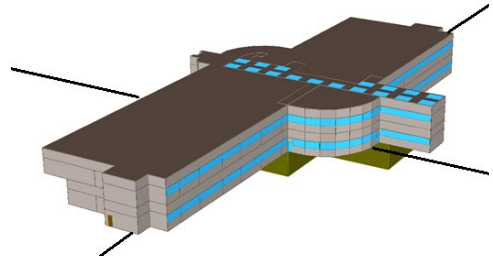
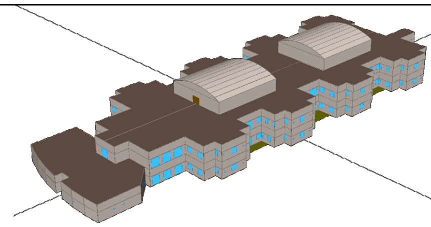
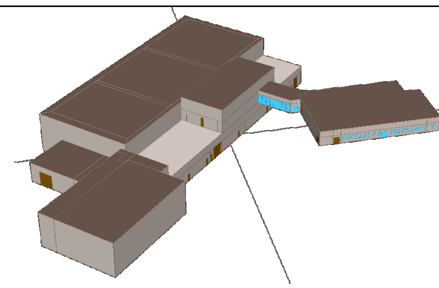
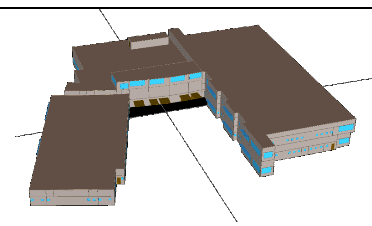
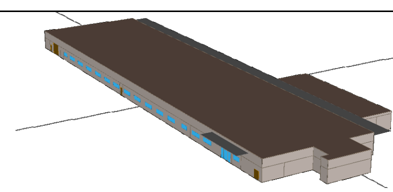
Figure 2 A BDL expression before and after insertion of calibration parameters shows how calibration is accommodated or else preserves the original parameter value.

Quality checks (QC) were confined to verification that building-specific details from existing energy audits were consistent. The reviewer verified that a BEM could run before and after changes were made. An average of 6 hours per building was spent generating a spreadsheet that lists where the model was consistent, areas that were corrected, and issues that were not correctable (and why) in the time frame available. The BEM model was also checked for errors in modeling methodology with investigation of all output warning messages from DOE2.2. Finally, a list of information in the energy audit that could not be usefully applied to the BEM was generated. The reviewer finished the quality check by assigning a grade to the BEM: A = ready for calibration, B = ready for calibration but with known issues, C = known issues are serious and the reviewer is uncertain whether to calibrate, D = calibrating is not recommended but known corrective action may correct issues, F = fundamental flaws exist or DOE2.2 is incapable of simulating the building accurately.

Seventeen buildings were chosen to undergo auto-calibration in 2018 of which five were finished in time for this paper. The five buildings are listed in Table 1 which gives the general attributes, calibration status, and QC grade with the exception of building 1 whose calibration was completed before QC checking was implemented. The buildings are highly complex, have multiple uses that vary sporadically, and many instances of 24-7 operations.

The calibration methodology used in this study is the Autotune technology (New et. al., 2012). The core Autotune capabilities were developed at Oak Ridge National Laboratory over the course of 3 years by leveraging high performance computers, and development of a software system for running a suite of machine learning algorithms (MLSuite) on these resources in parallel (Edwards, 2013a) composed of proprietary, open source, and computational complexity

Table 1 Building Attributes and Calibration Status 2014

Building	Description	Model Picture from eQUEST®
Building 1 Albuquerque New Mexico Built 1987	• 3 level 72,200ft^2 (6,710m^2) Light Lab with pre-cast concrete panels with exterior metal panels • 2 24-7 exhaust systems. • Mixed single duct and dual duct • 5.42 GWh (18,490 MBTU) electricity consumption 2017 • Calibrated om 2014 to NMBE -2.92% CV(RSME) 5.41%. No QC Grade	
Building 2 Albuquerque New Mexico Built 1995	• 3 Level 98,200ft^2 (9,120m^2) Light Lab with concrete masonry unit construction • 2 24-7 exhaust systems • Compressed air services • 2.26 GWh (7,710 MBTU) electricity consumption 2017 • 2,210 MCF (62,580m^3) natural gas used 2017 • Out of compliance for calibration in 2014. NMBE 5.50% CV(RSME) 6.40% QC Grade A	
Building 3 Albuquerque New Mexico Built 1984	• 2 Level 76,100ft^2 (7,070m^2) Highbay area with office space attached by a skybridge. Mostly precast with Double tee structural walls • Compressed air services • 1.57 GWh (5,360 MBTU) electricity consumption 2017 • 4,047 MCF (114,600m^3) natural gas used in 2017 • Out of compliance for calibration in 2014. NMBE 12.05% CV(RSME) 13.00% QC Grade A	
Building 4 Livermore California Built 2003	• 2 Level 71,500ft^2 (6,643m^2) steel frame office building • Offices and large conference room center. • Small café and dining area • 3.91 GWh (13,300 MBTU) electrical consumption 2017 • Not calibrated in 2014. QC Grade B	
Building 5 Livermore California Built 1958	• 1 Level 32,600ft^2 (3030m^2) concrete office building • 4-ply built up cool roof • 1.82 GWh (6,210 MBTU) electricity consumption 2017 • 6,050 MCF (171,300m^3) natural gas used in 2017 • Not calibrated in 2014. QC Grade C	

improvements (Edwards, 2013b) to artificial intelligence (AI) algorithms Over 300,000 AI algorithmic instances were utilized that included the following classes of algorithms: linearly and non-linear regression, feed forward and recurrent neural networks, C- and K-means clustering with local models, support vector machines, Gaussian mixture models, self-organizing maps, regression trees, time modeling, and genetic algorithms. The best-performing algorithm on a benchmark dataset of 20,000 buildings was a custom modification to the NSGA-II algorithm, which resulted in an average hourly CV(RMSE) below 4% (Garrett et. al. 2013, Garrett and New 2015, Chaudhary et. al. 2016). More importantly, this algorithm used a new ANSI/RESNET standard (ANSI 1201-2016) to quantify the recovery of the actual input parameters of the building between 15% and 32%, with better performance achieved with more channels of higher-resolution energy and non-energy performance data (New et. al., 2018). The core Autotune algorithm was further extended to DOE2.2 simulations required for this study. This included modification of BDL variables and scalable parsing of DOE2.2 output reports to enable improved calibration.

While not often reported in the literature, there are several practical considerations that many calibration studies face including BEM modifications, sensed data cleanup, selection of tunable parameters, properties of those parameters (e.g. min, max, distribution, grouping, mathematical constraints), and re-calibration via modifications of these when improved comparison data is provided from a calibration. In this study, the original building energy models were specified from March 27 – December 31, 2012 and January 1, 2012 – March 6, 2013. For this study, we modified the time period to January 1 – December 31, 2017 (the year is not important, but the time series reported needs to align with the sensed data). There were several sensed data cleanups, which is an ongoing challenge for many organizations with respect to centralizing and providing quality assurance/control of metered data (that may cover multiple buildings or overlap measurement boundaries with other analysis). While it is standard practice to calibrate on the last 12 months of data, to minimize the odds of a non-routine adjustment, we find comparison of those 12 months to previous years of data to be a useful exercise. Building 2 experienced a faulty natural gas meter in June that was detected as anomalous compared to the previous three years, so the usage was doubled. Building 3's electrical use was increased 67% due to an anomalous change in operations that is unlikely to repeat in the future. Building 4's energy use for August – December was averaged from previous years due to similar operational concerns. The authors find that most organizations and individuals end up re-evaluating the tunable parameters and properties after calibration once the modeled and measured data are compared together. Most of the reasoning and selection for the parameter properties used in this study are discussed previously (Villa et. al., 2018). We also find providing clear guidance on what is expected in a calibration report is useful. For this study, the final report for each building included an interactive report with the following information:

- Documentation of the quality check process with comparison to walkthrough audit data
- Prominent display of the final NMBE and CV(RSME) values achieved
- The final building energy model
- The building energy model before calibration
- Full details in result files for the final calibrated model
- Graph providing monthly building energy performance data of electricity/gas (if provided) versus the final calibrated building energy model performance
- Spreadsheet providing all the parameter values determined by the calibration algorithm
- The Actual Meteorological Year (AMY) weather file corresponding to the time-period in which data was collected
- The measured calibration data used
- Meta-information regarding the computer used, dates of run, individual who performed the analysis, contact information, and additional important notes
- Graph and data showing yearly end-use break-down by Heating, HVAC cooling equipment, HVAC Fans, interior lights, pumps, plug loads, interior lighting, and other loads
- Short notes concerning whether the end-uses are close to typical end-uses for the building type being evaluated

Archiving these results provides a permanent record of the analysis that allows direct application to the EISA energy audits and creation of a permanent link that allows replication or further investigation as needed.

RESULTS

For Building 1, the authors began with the original BEM model and a manually-calibrated model. Automatic calibration was applied to the manually-calibrated model to determine if it could improve beyond the manual calibration and to achieve the most accurate model for assessing energy savings opportunities. However, this calibration-on-top-of-calibration can lead to further differentiation of model inputs from walkthrough audit data. In addition to addressing this concern, the authors sought to determine if automatic calibration of the original model could have saved the cost and expense of the manual calibration, especially for matching measured energy use. Building 1 was calibrated using hourly data, but accuracy metrics are also shown for daily and monthly in Table 2 to inform how well the method generalizes across different temporal resolutions.

Formal interviews by the authors with individuals holding the job title of "energy engineer" that typically perform calibration services for Energy Service Companies (ESCOs) have identified an average of approximately 24 hours to manually calibrate a building of this size at an average fully-burdened rate of $130/hour. Using this assumed hourly rate, we show improved calibration performance and a savings of approximately $1,000 (29%) for this building and 12-20 hours in turn-around time. This value grows non-linearly with the increasing amount of sensed data, complexity of operations, number of calibration parameters, and the size of the building portfolio being calibrated.

Table 2. Building 1's monthly, daily, and hourly accuracy of the original model, auto-calibration, manual-calibration (from original), and auto-calibration (from manually-calibrated) shows generally better accuracy than manual calibration, minor cost savings, and increased scalability to cost-effectively address larger portfolios of buildings.

		Original Model	Autotune (from original)	Manual Calibration	Autotune (from manual)
Monthly utility data	CV(RMSE)	8.37%	5.22%	5.23%	5.20%
	NMBE	2.66%	0.16%	-1.42%	0.66%
Daily utility data	CV(RMSE)	12.41%	9.11%	8.26%	7.53%
	NMBE	2.66%	0.16%	-1.42%	-0.26%
Hourly utility data	CV(RMSE)	19.50%	10.9%	11.54%	9.70%
	NMBE	2.66%	0.16%	-1.42%	-0.26%
Cost			$2.5k (15 hours, compute)	27 person-hours ($3.5k at $130/hr)	$2.5k (7 hours, compute)

In performing the calibration for all 5 buildings, there were several real-world considerations that were noted as part of the analysis. First, most of the tunable parameters selected only affected electrical energy use. While some natural gas parameters were added, this was done after-the-fact and had little impact on natural gas energy use. When calibrating to items beyond whole-building electrical (e.g. natural gas, water), calibration benefits from significant forethought in the uncertainty of the major variables that affect those points of measure. Here, our automated approach to our parameter design makes it much easier to adjust the entire fleet as we refine our processes. With regards to natural gas, we noted several recommendations including the following:

- Buildings 2 and 3 natural gas base loads are needed for summer and shoulder seasons
- Building 2 electricity values were unusual for February, July, and September compared to previous years,
- Building 3 likewise deviated in February and September compared to previous years
- Building 4 had identical values in July, August, and September compared to the previous year and could be a potential copy/paste error in the energy audit report
- Building 5 had 30 times greater energy consumption for natural gas than the calibrated model with a 77% weather-independent base load and 23% weather-dependent (assumed to be space heating)

- Building 5 results were within 1% on an annual basis but off by as much as 30% on a monthly basis due to 6–27% lower than measured models results for January-August but 27–33% higher model results for September – December with measured energy use falling sharply between July and September. Since no building or operational change was noted in the supplied audit, it was conjectured that a model calibrated on only the last half of the year may be more valuable going forward.

	Original				Tuned			
	Electricity		Gas		Electricity		Gas	
Building	CV(RMSE)	NMBE	CV(RMSE)	NMBE	CV(RMSE)	NMBE	CV(RMSE)	NMBE
1	8.4%	2.7%	--	--	5.2%	0.7%	--	--
2	125.9%	125.4%	81.5%	54.3%	17.3%	13.5%	37.8%	13.8%
3	13.3%	-7.9%	47.8%	-42.5%	9.5%	-0.26%	13.9%	-1.4%
4	5.6%	2.3%	--	--	4.8%	0.2%	--	--
5	29.7%	-28.7%	97.3%	-96.2%	6.5%	-1.8%	97.8%	-96.6%

ASHRAE Guideline 14 NOT G14

Figure 3. Monthly BEM accuracy pre- and post-calibration. While ASHRAE Guideline 14 is almost exclusively applied to whole-building electrical use, future versions are considering submetering, higher-resolution sampling, and the (generally) more difficult challenge of matching natural gas or water use.

To estimate how much change was required to achieve calibration, the percent change across the allowed minimum to maximum ranges was calculated for each parameter for Buildings 2 through 5. This resulted in 236 data points concerning how much the auto-calibration procedure varied parameters (Figure 4). The resulting distribution has an average of 1.56% which we assume is converging to zero and a sample standard deviation of 30.07%.

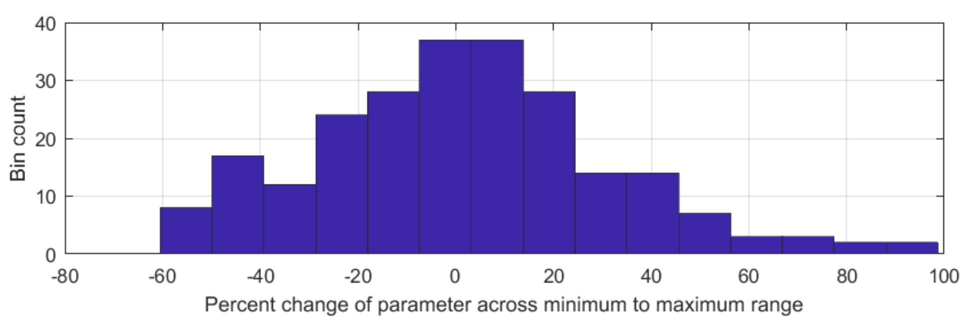


Figure 4. Auto-calibration parameter percent change

Such a high standard deviation is well beyond the desirable range of adjustment based on the authors' judgment. We hypothesize that it is indicative that the models are not sufficiently configured to represent the real building well. It is much more desirable for auto-calibration to tune a model that represents a real building accurately but for which unknown efficiencies and variations in schedules require some adjustments. There are many BEM attributes, such as thermal zone and HVAC configurations, which are prohibitively diffuclt to alter by auto-calibration. The authors had hoped for auto-calibration to achieve system identification of a core set of parameters such as efficiencies, changes to envelope, and changes to base and peak loads. Instead, we hypothesize that the models are being drastically changed by the optimization scheme to fit models that need to be configured more accurately before calibration. We think that this can be resolved by tighter connection to data sources in the building and we are moving forward to do this.

CONCLUSION

Maintenance of large fleets of BEMs for existing buildings requires a continuous cycle of quality checks and recalibration to increase model accuracy. In this work, we have taken first steps towards accomplishing this for a fleet of 121 BEM that are being used for EISA compliance on a four-year cycle. The auto-calibration design process has been completed and five BEMs have been successfully auto-calibrated with twelve more in process. For our case, the auto-calibration process has provided much needed deliverables for EISA 2007 compliance at economically competitive rates. Though reduced cost has been demonstrated, the reduced hours of labor required to calibrate models is of greater importance for our needs. Initiating the auto-calibration process has also elucidated the need for improvements to our data collection processes and methods for maintaining large fleets of BEMs. The quality checking procedure was found to be essential and even revealed misinformation in energy audit reports that was corrected, making it valuable beyond our application. Though the models for building 2 and building 5 did not achieve ASHRAE Guideline 14 compliance, all were drastically improved. For the failed cases, data and modeling accuracy issues are more likely at fault than any shortcomings of auto-calibration.

This work has revealed the need for refinement of many of the energy tracking processes at large institutions across our nation. For auto-calibration to be effective, BEM fleets need to be efficiently connected to building automation systems. Doing so involves challenges to understand sensor reliability, placement, and data interpretation. Yet, when connected thus, auto-calibration will better serve its designed purpose of tuning a model for identification of key parameters. Because of the lack of such connections, auto-calibration often may involve fitting models with configuration errors to data. The models therefore may need addition of missing systems, reconfiguration of existing systems, and corrections to schedule shapes. The drastic changes to parameters needed (Figure 4) for Buildings 2 through 5 suggest that this was the mode of operation for one or more buildings in this work so far. We plan to correct this through richer data connections. As these refinements are made, we expect to be able to open our BEM fleet to additional applications such as energy analytics that assist in identification of sudden unexpected changes in operations. To do this, auto-calibration will have to be done on a much more frequent basis than every four years. For the accuracy we desire out of our BEM's, we think that a QC/recalibration process must address every model parameter. To do so every parameter must be classified into four categories: 1) no changes needed, 2) discoverable by available data, 3) undiscoverable by available data, and 4) tuned by auto-calibration. The third category needs to be kept to a minimal set and can only be addressed by uncertainty analysis. The distribution of parameters within this classification scheme may serve as a basis to evaluate whether a model is likely to be accurate. Significant work is needed to be able to quickly classify BEM parameters into these categories. Regardless, our current efforts are a significant improvement over our previous practice and demonstrate methods that reduce resources required to maintain large fleets of BEMs.

ACKNOWLEDGEMENTS

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper has been approved by Sandia National Laboratories for unlimited release per SAND2018-13431 C but is still subject to any applicable ASHRAE release restrictions.

REFERENCES

ANSI/RESNET Standard 1201-2016 (2016) Standard Method of Test for the Evaluation of Ebuilding Energy Analysis Model Calibration Methods.

ASHRAE, 2011. "Procedures for Commercial Building Energy Audits 2nd Edition." Americal Society of Heating, Refrigeration and Air Conditioning Engineers, Inc.

Chaudhary, Gaurav. Joshua New, Jibonananda Sanyal, Piljae Im, Zheng O'Neill, Vishal Garg. 2016. "Evaluation of "Autotune" calibration against manual calibration of building energy models." *Applied Energy* 182: 115-134.

Chen, Yixing. Tianzhen Hong, Mary Ann Piette. 2017. "Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis." *Applied Energy* 205: 323-335.

Edwards, Richard E., Parker, Lynne E., and New, Joshua R. (2013). "MLSuite Software Transfer Document." ORNL internal report ORNL/TM-2013/131, May 8, 2013, 36 pages.

Edwards, Richard E., Zhang, Hao, New, Joshua R., and Parker, Lynne E. (2013). "Approximate l-fold Cross-Validation with Least Squares SVM and Kernel Ridge Regression." In *Proceedings of the IEEE 12th International Conference on Machine Learning and Applications (ICMLA13)*, Miami, FL, December 4-7, 2013.

EISA, 2007. Energy Independence and Security Act of 2007 § 432. Public law 110-140-Dec. 19, 2007 110th Congress.

Fisher, Maryl D. 2014. "EISA 432 Energy Audits Best Practices: Software Tools." Idaho National Laboratories INL/EXT-14-33070. November.

Garrett, Aaron, New, Joshua R., and Chandler, Theodore (2013). "Evolutionary Tuning of Building Models to Monthly Electrical Consumption." Technical paper DE-13-008. In *Proceedings of the ASHRAE Annual Conference* and *ASHRAE Transactions 2013*, volume 119, part 2, pp. 89-100, Denver, CO, June 22-26, 2013.

Garrett, Aaron and New, Joshua R. (2015). "Scalable Tuning of Building Energy Models to Hourly Data." In *Journal of Energy*, volume 84, pp. 493-502, April 2015.

Hong, Tianzhen, Jared Langevin, and Kaiyu Sun. "Building simulation: Ten challenges." *Building Simulation* 11:871-898.

ORNL 2018. 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.

New, Joshua R., Sanyal, Jibonananda, Bhandari, Mahabir S., and Shrestha, Som S. (2012). "Autotune EnergyPlus Building Energy Models." In *Proceedings of the 5th National SimBuild of IBPSA-USA*, International Building Performance Simulation Association (IBPSA), August 1-3, 2012.

New, Joshua R., Garrett, Aaron, Sanyal, Jibonananda, Slattery, Bob, Gehl, Anthony, and Miller, William A. (2019) "Big Data Mining for Assessing Calibration of Building Energy Models." To appear in *International Journal of Computer & Software Engineering (IJCSE)*, volume 3, issue 136, Sept. 8, 2018.

NREL 2018. UrbanOpt. Available at https://www.nrel.gov/buildings/urbanopt.html. National Renewable Energy Laboratory open-source software. Accessed October 25, 2018.

Reinhart, Christoph F. and Carlos Cerezo Davila. 2016. "Urban building energy modeling—A review of a nascent field." *Building and Environment* 97: 196-202.

Stuart, Elizabeth. Juan Pablo Carvallo, Peter H. Larsen, Charles A. Goldman, and Donald Gilligan. 2018. "Understanding recent market trends of the US ESCO industry." *Energy Efficiency* 11:1303-1324.

Villa, Daniel L., Jack H. Mizner, Howard D. Passell, Marlin S. Addison, Gerald R. Gallegos, William J. Peplinski, Douglas W. Vetter, Christopher A. Evans, Leonard A. Malczynski, Matthew A. Schaffer, and Matthew W. Higgins. 2017. "Institutional Transformation Version 2.5 Modeling and Planning." Sandia National Laboratories Technical Report SAND2017-1498. February.

Villa, Daniel L., Jack H. Mizner, and Gerald R. Gallegos. 2017 "Institutional Transformation: Using Detailed Building Energy Models for Site-Wide Energy Conservation Strategies." *Seminar 9 – Building Performance Analysis for Policymaking, Part 2 2017 ASHRAE Building Performance Analysis Conference.* September.

Villa, Daniel L., Joshua New, Gerald R. Gallegos, Mark Adams, William Peplinski, Aaron Garrett, and Thushara Gunda. 2018. "Reducing Building Energy Modeling Workloads Through Automated Calibration." Presented at the 2018 Building Performance Analysis Conference, Chicago, Illinois, September 26-28.