Analysis of a simplified calibration procedure for 18 design-phase building energy models

Holly W. Samuelson[a], Arash Ghorayshi[b] & Christoph F. Reinhart[c]

[a] Graduate School of Design, Harvard University, 48 Quincy Street, Cambridge, MA 02138, USA

[b] MMM Group Ltd., formerly Enermodal Engineering, 582 Lancaster Street West, Kitchener, ON Canada N2K 1M3;

[c] Department of Architecture, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
Published online: 07 Jan 2015.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

Analysis of a simplified calibration procedure for 18 design-phase building energy models

Holly W. Samuelson[a*], Arash Ghorayshi[b] and Christoph F. Reinhart[c]

[a]Graduate School of Design, Harvard University, 48 Quincy Street, Cambridge, MA 02138, USA; [b]MMM Group Ltd., formerly Enermodal Engineering, 582 Lancaster Street West, Kitchener, ON Canada N2K 1M3; [c]Department of Architecture, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

This paper evaluates the accuracy of 18 design-phase building energy models, built according to LEED Canada protocol, and investigates the effectiveness of model calibration steps to improve simulation predictions with respect to measured energy data. These calibration steps, applied in professional practice, included inputting actual weather data, adding unregulated loads, revising plug loads (often with submetered data), and other simple updates. In sum, the design-phase energy models underpredicted the total measured energy consumption by 36%. Following the calibration steps, this error was reduced to a net 7% underprediction. For the monthly energy use intensity (EUI), the coefficient of variation of the root mean square error improved from 45% to 24%. Revising plug loads made the largest impact in these cases. This step increased the EUI by 15% median (32% mean) in the models. This impact far exceeded that of calibrating the weather data, even in a sensitivity test using extreme weather years.

Keywords: energy simulation; plug loads; process loads; unregulated loads; weather; design-phase models; compliance models

1. Introduction

Commercial buildings are complicated thermodynamic objects. Occupant behaviour and local weather conditions may strongly impact their energy use, making it difficult to detect operational inefficiencies such as malfunctioning equipment and inefficient control strategies. Unfortunately, such inefficiencies are ubiquitous and waste substantial energy and money. A study of 60 buildings by Lawrence Berkeley National Laboratory showed that, before intervention, at least half had problems with their system controls, a minimum of 40% had faulty heating ventilating and air conditioning (HVAC) equipment, and 25% or more had problems with specific "green" HVAC features (Piette, Nordman, and Greenberg 1994). One group of researchers estimated that finding and fixing these types of problems can save $18 billion or more per year in commercial buildings in the USA alone (Mills et al. 2004).

Fortunately, researchers have shown that state-of-the-art computer-based energy simulation can aid in the search for operational problems in processes such as existing-building commissioning (Adam et al. 2004; Andre et al. 2003) and ongoing-commissioning (O'Neill et al. 2011; Song et al. 2003). Carling, Isakson, and Blomberg (2004) demonstrated the value of whole-building energy simulation to support new building energy commissioning. At the same time, they noted that their procedure was too

expensive for routine application as preparing the detailed model took several weeks. This situation helps to explain why, despite advances in research, practitioners rarely use energy simulation beyond a building's design phase (Claridge 2011). The post-occupancy task of measurement and verification (M&V) is an occasional exception. While an established M&V methodology exists which uses a calibrated energy model to isolate the impact of installed energy conservation measures, even its proponents caution that the methodology "is best applied where . . . the budget for M&V is generous". (Efficiency Valuation Organization 2006). A key obstacle to the use of simulations during building operation is hence the cost of generating an energy model to start.

This obstacle may become less relevant for many projects as building energy codes and green building standards, such as the UK's Building Regulations Part L (Stationery Office London 2010), California's Title 24 (California Energy Commission 2007), and LEED (US Green Building Council 2009) promote and/or require the wider spread use of building energy modelling during building design. Could these design-phase energy models then be used to improve building operations?

The answer is, "probably not directly", since researchers have shown design-phase models to often be poor predictors of actual energy use in buildings (Ahmad

*Corresponding author. Email: hsamuelson@gsd.harvard.edu

and Culp 2006). In fact, ASHRAE 189.1 specifically alerts the reader that design-phase energy models may substantially diverge from actual building energy use "due to variations such as occupancy, building operation ... energy use not covered by this procedure ... " and so on. (ANSI, ASHRAE, USGBC, & IES 2010). However, despite these well-known facts regarding model accuracy, in a recent survey of 116 energy modellers 75% believed that (once calibrated) their design-phase models could provide value during building commissioning and operations (Samuelson, Lantz, and Reinhart 2012). Given this optimism within the energy simulation community, the question becomes, what shortcomings of design-phase models can be improved with calibration steps?

Model calibration can be time-consuming and has been described as "an art form that inevitably relies on user knowledge ... and an abundance of trial and error" (Reddy 2006). Various researchers have advanced the calibration process (Clarke, Strachan, and Pernot 1993; Haberl and Bou-Saada 1998; Pan, Huang, and Gang 2007; Reddy, Maor, and Panjapornpon 2007) and developed time-saving techniques, especially for use with hourly measured data. The greater the granularity of measurements, in terms of time steps and end uses, the more accurate the calibration procedure can be. For instance, a widely researched calibration method called signature analysis involves analysing the fit between simulated and measured hourly data to deduce the necessary model adjustments (Liu and Liu 2011; Liu et al. 2003; Liu et al. 2006; Wei, Liu, and Claridge 1998). Practitioners with less granular information, such as monthly metered data only, face additional challenges. Other problems that modellers face during calibration are that most simulation packages rely on significant approximations of HVAC systems (Bannister 2009). Finally, design-phase models pose unique challenges, since practitioners built these models when many model inputs were yet unknown and with the intent of comparing design options or demonstrating code compliance rather than predicting absolute energy consumption.

In this study, the authors describe and examine the effectiveness of a simple, heuristic calibration procedure for design-phase energy models that has been used over several years in the second author's sustainable consultancy practice in projects for which the firm was hired as the energy modeller during design as well as the commissioning agent during the first year of building operation. In their latter capacity, the firm uses the calibration procedure to help ensure that a building functions as intended. The calibration procedure is evidence based in the spirit of Raftery, Keane, and O'Donnell (2011), in that the modellers only update input parameters based on more reliable evidence, such as direct measurements or better informed estimates regarding the actual operation of a building. In the following, the ability of the different calibration steps to increase simulation accuracy is evaluated for 18 buildings and recommendations are made

for others interested in adopting a comparable process in their practice.

2. Methodology

2.1. The data set

The authors analysed a set of 18 commercial building projects located within 100 km (60 miles) of Toronto, Ontario. For each project, modellers from the earlier mentioned energy consulting firm had built the energy models to support the new building design process and to document energy conservation measures for LEED Canada New Construction or Core-and-Shell version 1.0 certification (Canada Green Building Council 2007). The modellers then "partially calibrated" the models, typically over the first year of occupancy, to inform the measurement and verification (M&V) and extended commissioning process. Here, the authors use the term "partial-calibration" to describe the act of bringing the energy model inputs closer to as-operated conditions, while recognizing that the modellers could have calibrated the models further with additional effort.

The authors started the investigation with all 34 of the firm's model/calibration projects with at least one year of monthly measured utility data available at the time. The 10 different modellers who had built these models had not documented their projects with the idea that future researchers would re-create their work. Therefore, nine projects needed to be eliminated due to inadequate documentation. The authors also removed another seven models that were built in software packages other than EE4, a DOE2.1-based simulation program used to demonstrate compliance with LEED Canada (Canada Green Building Council 2007), in order to streamline the time-intensive analysis process across the remaining 18 cases. The authors manually re-ran the models for each calibration step in order to validate the original calibration process.

All 18 case studies were new buildings, completed between 2008 and 2012 and then monitored for at least 12 months. The buildings ranged in size from 416 m^2 (4480 ft^2) to 51,000 m^2 (550,000 ft^2) and one to 16 stories. Table 1 provides more project information.

To understand how these cases compare to other LEED-certified projects, Figure 1 shows the measured versus predicted energy use intensity (EUI) of the 18 case studies alongside a data set from Turner and Frankel (2008), which consists of LEED (US) Version 2 (US Green Building Council 2001) certified buildings completed between 2001 and 2007.[1]

The Canadian cases have a slightly higher median EUI of 244 kWh/m^2 compared to 198 kWh/m^2 for the Turner and Frankel cases. However, model accuracy was similar in both groups; the design-phase simulation predictions deviated by 31% median from the measured EUI for the Canadian buildings compared to 36% in the older study.

Table 1. Characteristics of building cases.

Bldg. #	Area (m²)	Building type	Stories	LEED Canada 1.0 version	Sub-metered end uses	Notable systems and conservation measures[a]
1	416	Ecology centre	2	New construction (NC)		Radiant floors, occupancy sensors
2	1037	Ecology centre	2	NC	Mech. systems, plug loads, lighting	Photovoltaic, ground source heat pump (GSHP), radiant floors, daylight sensors
3	1185	Ecology centre	1	NC	Plug loads, lighting	GSHP, radiant floors, occupancy sensors, DCV, all-electric building
4	5575	Office	3	NC	Mech. systems, lighting, emergency system	
5	7934	Office	4	Core & Shell (CS)	Mech. systems, lighting, elevators, emergency system, fire pumps, other power	Water-loop heat pump, DCV, occupancy and daylight sensors
6	4148	Police office with laboratory	2	NC	Mech. systems, plug loads & cooling, lighting	Occupancy sensors
7	10,879	Police office	2	NC		
8	13,000	Warehouse-type retail, non-food	1	NC	Lighting, battery charging	DCV
9	10,590	Warehouse-type retail, non-food	1	NC	Mech. systems, lighting, special plug loads	GSHP, DCV, occupancy sensors
10	51,000	Office	8	CS	Mech. systems, plug loads, lighting, tenant loads, elevators, emergency system, retail tenant loads	DCV, occupancy sensors, district heating/cooling plant
11	12,600	Higher education (labs., classrooms, office, assembly)	5	NC	Emergency system	DCV, variable-flow fume hoods, district heating/cooling plant
12	16,568	Social multi-unit residential building (MURB)	15	NC (MURB)	Mech. systems, common area plug loads, tenant loads, emergency system, snow-melt, parking and exterior lighting	
13	45,700	MURB (two buildings)	14 & 16	NC (MURB)	Common area: mech. systems, plug loads, elevators, emergency system, swimming pool equip., exterior lighting	Occupancy sensors, district heating/cooling plant
14	1730	Mail sorting warehouse	1	NC	Mech. systems, plug loads, lighting, motorized doors & battery charging	Occupancy and daylight sensors, DCV
15	1147	Mail sorting warehouse	1	NC	Mech. systems, plug loads, lighting, motorized doors & battery charging	Occupancy and daylight sensors, DCV
16	4130	Higher Education (classrooms)	2	NC	Common area: mech. systems, plug loads, lighting, elevators, emergency system, cafe tenant loads, exterior shades	occupancy sensors, district heating/cooling plant
17	18,013	Luxury MURB (two buildings)	10 & 8	NC (MURB)	Common area: mech. systems, plug loads, elevators, exterior and parking lighting	District heating/cooling plant
18	27,000	Office/Call Centre	10	CS	Mech. systems, plug loads, lighting, emergency system	Occupancy sensors

[a] All buildings have energy recovery ventilation. All buildings have a natural gas-fired boiler, unless otherwise noted. Occupancy sensors, where listed, are part of the lighting control.
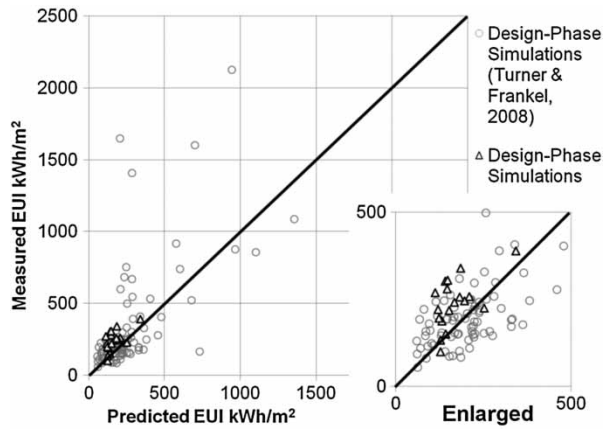
Figure 1.   Current cases shown in context.

2.2.   *Modelling protocol and software*

In the current cases, the energy modellers originally followed the Performance Compliance for Buildings protocol to demonstrate compliance with the Model National Energy Code for Buildings (MNECB) for Commercial Building Incentive Program compliance (Candian Commission on Building and Fire Codes 1997) and LEED Canada 1.0. Per the MNECB/LEED compliance process, the models are peer-reviewed, which may help improve the pre-calibration model quality, and modellers follow a detailed modelling procedure that creates more consistency between models than would otherwise be the case. The modellers used the EE4 version 1.7 software (Natural Resources Canada 2005), based on the DOE2.1E simulation engine[2] (J. J. Hirsch & Regents of the University of California 1999) which references DOE-2 hourly .bin format weather files.

Importantly, the Performance Compliance protocol encourages modellers to use default values, according to building type, for the following loads: number of occupants, receptacle power, service water heating, outdoor air requirements, space temperature schedule, and occupancy/lighting/operation schedules. Modellers can, however, add special loads/heat gains as an additional input. EE4 contains a library of input selections related to the loads above. These values can be overwritten, but the onus then falls on the modeller to document non-standard conditions.

The EE4v1.7 software itself poses limitations for the prediction of absolute energy performance. For example, it lacks the capacity to model the following features: shading/reflection from vegetation or neighbouring buildings, exterior lights, elevator usage, and steam humidifiers (Natural Resources Canada Office of Energy Efficiency & CANMET Energy Technology Centre 2008) as well as special equipment whose energy use is not regulated by the code/incentive program. With the exception of vegetation/neighbouring buildings, the modellers added

loads from missing features as part of the calibration process.

2.3.   *The calibration method*

The modellers performed a series of steps to improve the accuracy of the models with reasonable effort, ensuring that each revised input came from a more reliable source than the original input, as noted above. In some cases, the modellers made adjustments within the energy model, that is, the EE4 software interface. In other cases, the modellers post-processed the simulation results in a spreadsheet, for example, replacing simulated with sub-metered plug loads. This external approach overcame two hurdles. First, it allowed the modellers to add loads from building features, such as exterior lights, that the software interface could not support. Second, it allowed the modellers to overcome a discrepancy between the building's coarsely zoned sub-metering systems and the finely zoned model geometry.[3] The benefits and shortcomings of this external calibration approach will be discussed in Section 4.1.

2.4.   *Analysis procedure*

2.4.1.   *Overview*

The original modellers performed the calibrations in an ad hoc fashion, for example, inputting three months of actual weather data while simultaneously changing the occupancy schedules. This process prohibited the authors from ascertaining which calibration tasks substantially affected the simulations. Therefore, the authors re-ran all simulations and systematically isolated each step. The authors made minor revisions to the modellers' original approaches, if they suspected an error or if expanded sub-metered data became available.

The starting point for each case was the "design-phase model", that is, the energy model submitted to demonstrate compliance with the LEED energy credits. For analysis, the authors divided the calibration steps into categories, described in Sections 2.4.2 and 2.4.3. The authors changed one category of inputs at a time and, after each step, identified the net change from the previous model's results.

2.4.2.   *Weather*

The authors compared simulation results using typical versus historical weather data corresponding to the building metering periods. The modellers/authors used typical weather data included in the EE4 library, originating from the Canadian Weather for Energy Calculations (CWEC) database, an amalgamation of 1960–1991 weather data. They obtained the historical weather data from the National Climate Data and Information Archive and formatted it for use in DOE2.1E using the DOEWTH.exe converter

(J. J. Hirsch & Regents of the University of California 1999). Since these historical data did not include solar radiation information, a common problem, the modellers used solar radiation data from the CWEC typical weather file, a known modelling inconsistency. The weather converter uses monthly average clearness, which the modellers estimated from the hourly CWEC horizontal irradiance and monthly average extraterrestrial radiation for the latitude, per Duffie and Beckman (1991).

The authors also performed a sensitivity analysis to investigate the effect of calibrating weather inputs in extreme weather years. Crawley (2008a, 2008b) investigated the impact of historical weather data from 1961 to 1999 on the annual simulation results of multiple test buildings in various cities including Toronto. He found that 1998 and 1972 resulted in the lowest and highest simulated energy consumption, respectively, in his Toronto cases. Therefore, the authors used 1998 and 1972 Toronto weather data (Environment Canada 2001) in their sensitivity analyses.

Due to the known modelling inconsistency with the solar radiation data in the original calibration procedure, the authors performed a second sensitivity analysis to understand the potential impact of solar inputs on resulting simulated energy consumption. The 1998 and 1972 weather files (Environment Canada 2001) did include solar radiation data (albeit modelled data, since measured historic radiation data are rare.) The authors performed simulations first using the unadulterated 1998 and 1972 weather data and then exchanging the solar radiation inputs for values taken from the CWEC typical weather file.

2.4.3. *Plug loads, unregulated loads, and other calibration categories*

The authors used the historical weather data when testing the remaining calibration steps. They broke these steps into the following categories: HVAC equipment, HVAC operation schedule, occupant density/schedule (also called occupancy), lighting, infiltration, plug loads, and unregulated loads. The lighting category included any changes to the installed lighting power density, controls, or schedules. The plug-loads category included revisions to energy use from office equipment, home appliances, and other equipment related to the occupants' business or activity, as listed in Section 3.2. In all cases, the modellers included plug-load estimates in the design-phase models. In many cases, they used the MNECB default values for plug-load densities and operating schedules, as described in Section 2.2, which they revised during calibration. In contrast, the "unregulated loads" category included loads that the modellers excluded from the design-phase model, because they are unregulated by the code/incentive protocol and excluded from the EE4 software interface. This category included exterior lights and other building equipment listed in Section 3.3.
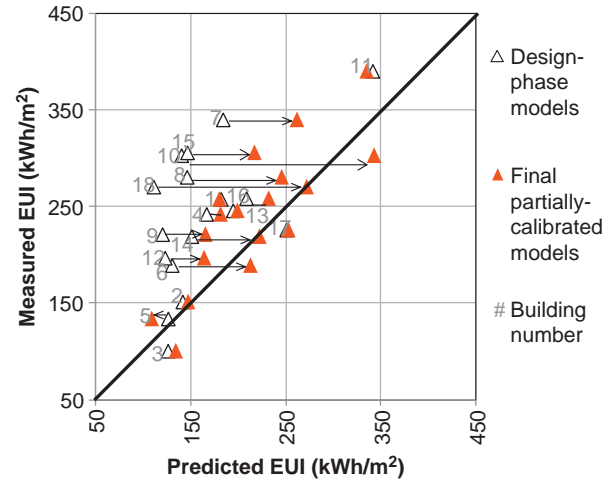

Figure 2. Goodness of fit: design-phase models vs. partially calibrated models.

3. Results

3.1. *Overview*

The "goodness of fit" between measured and predicted results improved substantially with the partial calibration, as illustrated in Figure 2. In sum, the design-phase energy models underpredicted the measured energy consumption for all 18 buildings by 36%. Following calibration, this error decreased to a net 7% underprediction. The annual mean bias error improved from −28% to −9%. The maximum difference between measured and simulated annual energy use for any building reduced from 59% to 34%. Using the formulas shown in the equation below, the accuracy also improved in terms of monthly normalized mean bias error (nMBE) and coefficient of variation of the root mean square error CV(RMSE), from 41% to 18% and 45% to 24%, respectively.

Aggregate statistics like these can help the modellers evaluate the overall effectiveness of their modelling and calibration practices. These numbers suggest that the partial calibration process helped to rectify the discrepancy, but some noteworthy building energy use remained unexplained by the models. Trends such as the pervasive design-phase underpredictions may indicate that the modellers need to adjust their future design-phase assumptions. The following analysis helps to deduce which design-phase inputs may be the most problematic.

Single-building statistics:

$$\text{monthly nMBE} = \frac{\sum_{i=1}^{12} (\text{pred}_i - \text{meas}_i)}{\overline{\text{meas}} \times (12 - p)},$$

$$\text{monthly CV(RMSE)} = \frac{1}{\overline{\text{meas}}}$$

$$\times \sqrt{\frac{\sum_{i=1}^{12} (\text{pred}_i - \text{meas}_i)^2}{(12 - p - 1)}}.$$

Combined statistics:

$$\text{weighted mean of all buildings} = \frac{\sum_{i=1}^{18} |Z_i \times A_i|}{\sum_{i=1}^{18} (A_i)},$$

where pred = predicted monthly EUI, meas = measured monthly EUI, $\overline{meas}$ = mean of measured monthly EUI, p = number of predictor variables (1), Z = either: nMBE, or CV (RMSE) for each building, and A = annual measured energy consumption for each building

Figure 3 shows examples of the step-by-step calibration process for Building 8, a warehouse-type, or "big box" store. The authors included this particular case because it most clearly delineates the process from one step to another. In contrast, Figure 4 shows the same data for Building 3, an ecology centre, which had the highest error statistics and least improvement with calibration. These results illustrate that the process did not always improve the model's goodness of fit.

Figures 5–7 show the results of each calibration step for each building compared to measured electricity and natural gas consumption. The authors list the changes included in the most notable calibration steps. Table 2 lists each calibration step and indicates in how many cases the modellers implemented that step. The authors also calculated the median impact of the step, where implemented, and evaluated the relative effort involved in the calibration step (excluding data gathering.)

The modellers implemented three calibration steps in most of the cases, namely, updating plug loads, adding unregulated loads, and replacing CWEC data with historical weather data, since revised values for these inputs were readily obtainable. Therefore, for the group as a whole, those three relatively easy steps accounted for the majority of the model improvement. The modellers implemented other calibration steps only on a limited portion of the cases, that is, the specific cases where more accurate inputs were available and the modellers believed that their impact would warrant the effort. For example, the modellers had no additional information regarding infiltration rates. Therefore, they kept the original infiltration rates except in two special cases where they had reason to suspect significant problems with the default assumptions, such as the case described in Section 3.5.

3.2. Updating plug loads

The modellers produced the biggest overall calibration impact, by far, by revising the plug loads. As shown in Figure 8, they implemented this step in 14 cases which increased the predicted EUI by 15% median (32% mean). Revised plug loads included the following: miscellaneous occupant equipment loads (in 13 cases), computer servers (in two cases), laboratory equipment, broadcast equipment, retail displays, battery charging stations, communal laundry, and café equipment (in one case each). They used
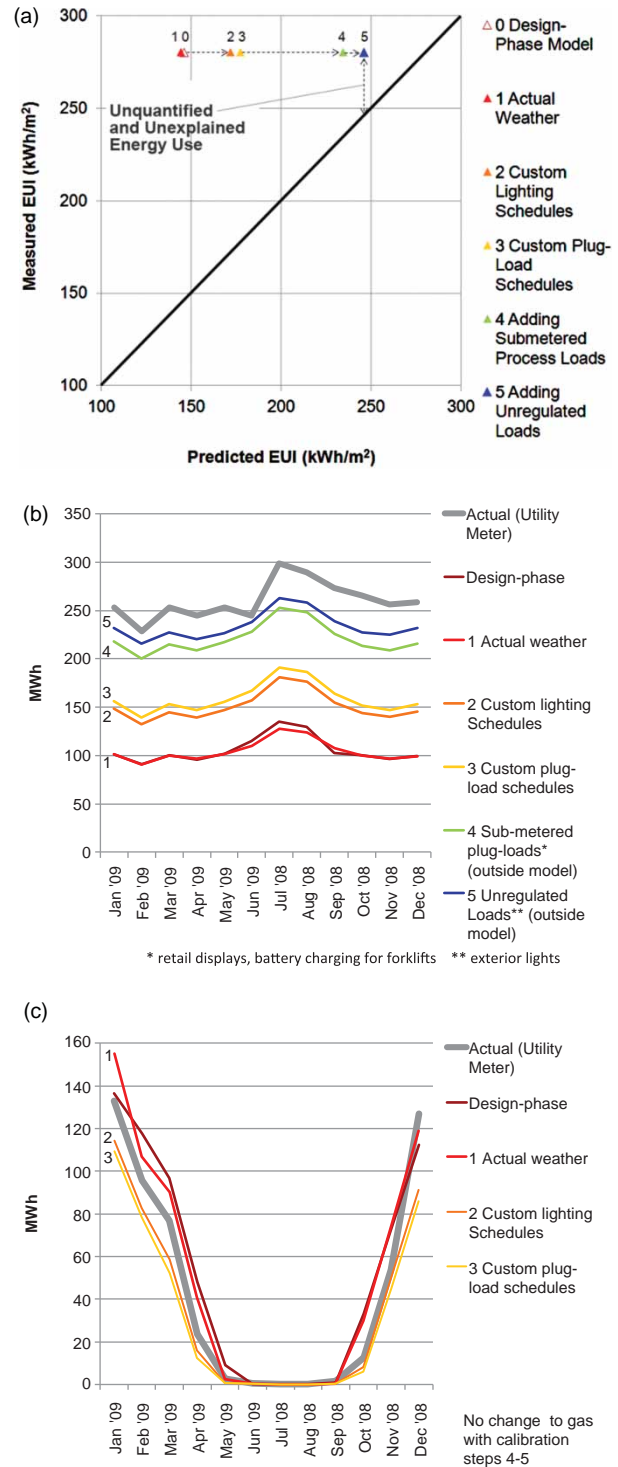

Figure 3.   (a)–(c): The calibration steps for Building 8, a warehouse-type store. (a) EUI. (b) Monthly electricity. (c) Monthly natural gas.

monthly sub-metered data unless otherwise noted. As can be seen in Figure 8, the plug loads decreased slightly (2% and 7%) in only two (warehouse-type) buildings and increased considerably in four notable cases. In Buildings 10 and 18, both large core-and-shell office buildings, the
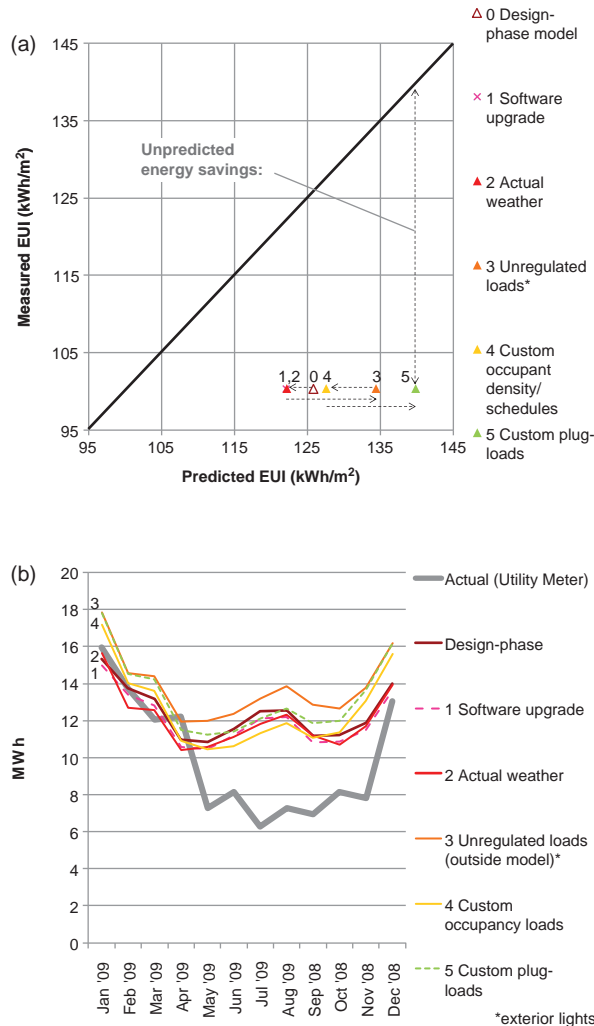
Figure 4. (a)–(b): Calibration steps for building 3, an ecology centre. (a) EUI. (b) Monthly electricity (all electric building).

modeller assumed typical office functions using MNECB default loads and schedules. After construction, Buildings 10 and 18 housed more energy-dense functions including a call centre and a broadcast station, respectively. Both cases resulted in much higher monthly sub-metered loads than predicted plug loads. In Building 6, police offices, the monthly sub-metered forensic laboratory loads proved to be much higher than modelled. In Building 8, a warehouse store, the modeller ultimately added missing loads from retail displays (estimated from short-term measurement) and battery charging stations for equipment such as forklifts (sub-metered). Revising the plug loads in these four buildings increased the EUI by 145%, 142%, 48%, and 32% respectively.

3.3. Adding unregulated loads

The modellers added unregulated loads in all but three cases. Considering the group as a whole, this calibration step made the second largest impact on the predicted EUI.

The modellers added loads from exterior or parking lighting (calculated via installed lighting power and predicted hours of operation) in 15 cases, resulting in a median EUI increase of 4% (5% mean). The modellers similarly added missing elevator loads (estimated based on measurements from a benchmark building) in eight cases, resulting in a median EUI increase in 1% (1% mean). They also added estimated or metered loads from security equipment, emergency equipment, pool heating, motorized doors, and a snow-melt system (1 case each). In total, the modellers added unregulated loads in 15 cases, which increased the EUI by a median of 5% (7% mean) as shown in Figure 9. The largest impact (32% increase in EUI) from this step occurred in Building 12, a multi-unit residential building, when the modeller added monthly metered loads from a large parking facility.

3.4. Updating weather inputs

The modellers replaced the CWEC weather data with historical data in all 18 cases, which decreased the EUI by a median of 2% (2% mean). The years 2008–2012 were all warmer than average in Toronto (Environment Canada 2001), and in this heating-dominate climate, in all but two cases, the updated weather data resulted in lower simulated energy use. Contrary to expectations, calibrating the weather data generally *increased* the discrepancy between measured and simulated energy use, both annually and monthly, because the less accurate weather data partially counteracted other inaccuracies in the model at that point in the calibration.

Figure 10 shows the results of these weather calibrations as well as the sensitivity analysis using Toronto extreme weather years, as described in Section 2.4.2. Compared to the CWEC data, the extreme cool year had −18% cooling degree days (CDD) and +23% heating degree days (HDD), whereas the extreme warm year had +20% CDD and −25% HDD. Despite this wide variability, the cool and warm extreme years resulted in only a 4% increase and 8% decrease, respectively, in median simulated EUI (4% and 8% mean) across the 18 cases, compared to the typical weather year. These results give some indication of the effect that weather calibration *could* make in more extreme years. One should note that this low effect of weather on the 18 buildings is a testament to their high-quality thermal envelopes. The number would very likely be larger for more standard Canadian buildings. In the analysis of solar radiation data in the extreme years, as described in Section 2.4.2, the change in solar radiation data produced a 1% median (1% mean) change in EUI over the 18 cases.

3.5. Other calibration steps

This section includes the highest impact cases for each of the remaining calibration categories. In Building 14,
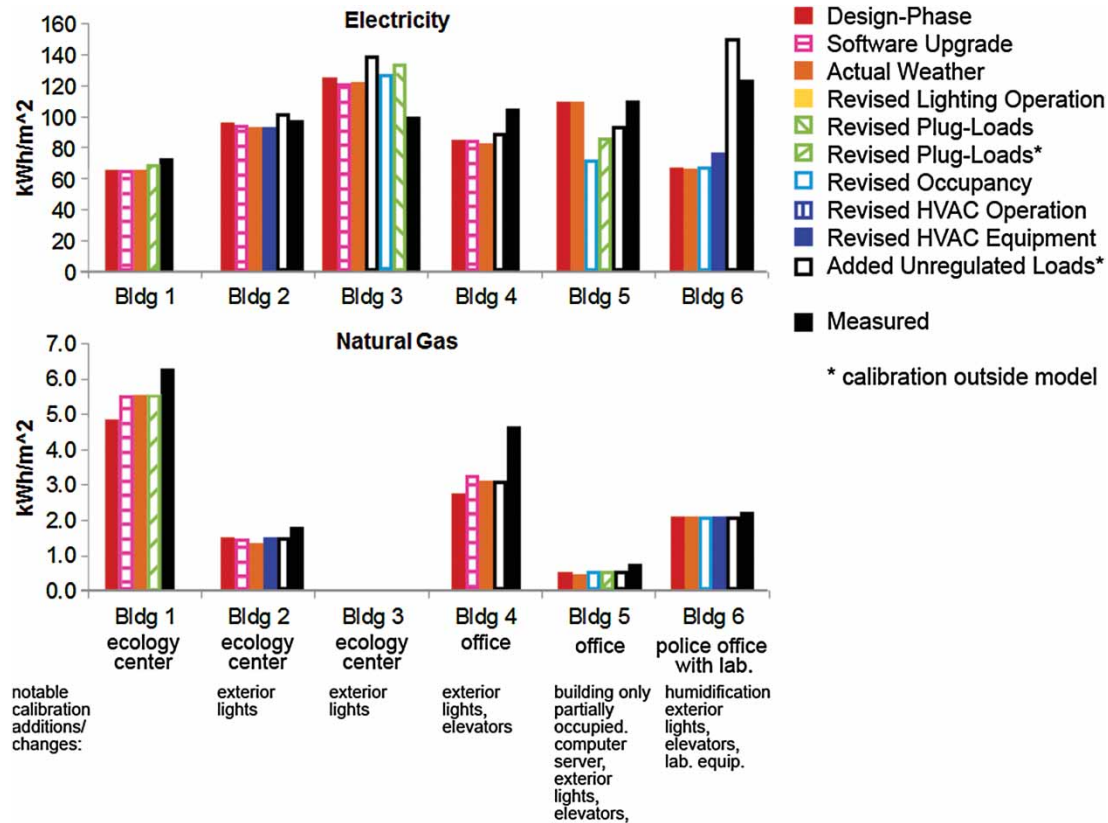
Figure 5.   Buildings 1–6 measured versus simulated annual energy at each calibration step.

a mail sorting warehouse, revising the infiltration rate to consider the often-open loading dock doors made a substantial impact, increasing the simulated EUI by 75%. The modeller calculated the new infiltration rate by hand based on opening sizes. Therefore, the value is still estimated, but based on more information than the default assumption. In Building 18, an office/call centre described above, revising the HVAC operation schedules, based on an interview with the building manager, increased the EUI by 20%. The installed HVAC equipment differed from the design-phase assumptions in Building 7, a police office. This change increased the predicted EUI by 19%. In Building 8, a warehouse store, the modellers revised the lighting schedules, based on short-term measurements, resulting in a 19% increase in EUI. Conversely, in Building 3, an ecology centre, their revised occupancy schedules, based on an interview with the building owner, decreased the EUI by 9%.

4.   Discussion and outlook

4.1.   Partial model calibration

The previous section revealed that the straightforward and low-effort partial calibration process described above reduced the annual and monthly mean bias errors between measured and simulated energy use in 18 building by more than 50% compared to using the original design-phase

models. This finding is both important and encouraging and the authors recommend that design teams and owners adopt a version of this procedure to improve their compliance-type models.[4] It is important to note that the 18 case study projects did not receive any additional research funding to support the calibration process described. This demonstrates that partial model calibration can, in fact, be conducted in a for-profit context.

In these cases, like many compliance-type models, design-phase inputs did not match the realities of the as-occupied/as-operated building, and unregulated loads were not included in the original models. Here, revising the plug loads generally created the largest impact followed by adding unregulated loads. Both steps required relatively little effort. In some cases, the modellers implemented a spreadsheet shortcut for updating these loads. In those cases, they minimized the calibration effort, but excluded the interrelated effects of the revised load on other systems in the model.[5] Even so, this spreadsheet approach could be a cost-effective initial calibration step that the modeller could replace with a more exacting approach later if desired.

4.2.   Remaining inaccuracies

As discussed, the partial calibration did not correct all model inaccuracies. Parameters which were not monitored here should not be considered insignificant. For example,
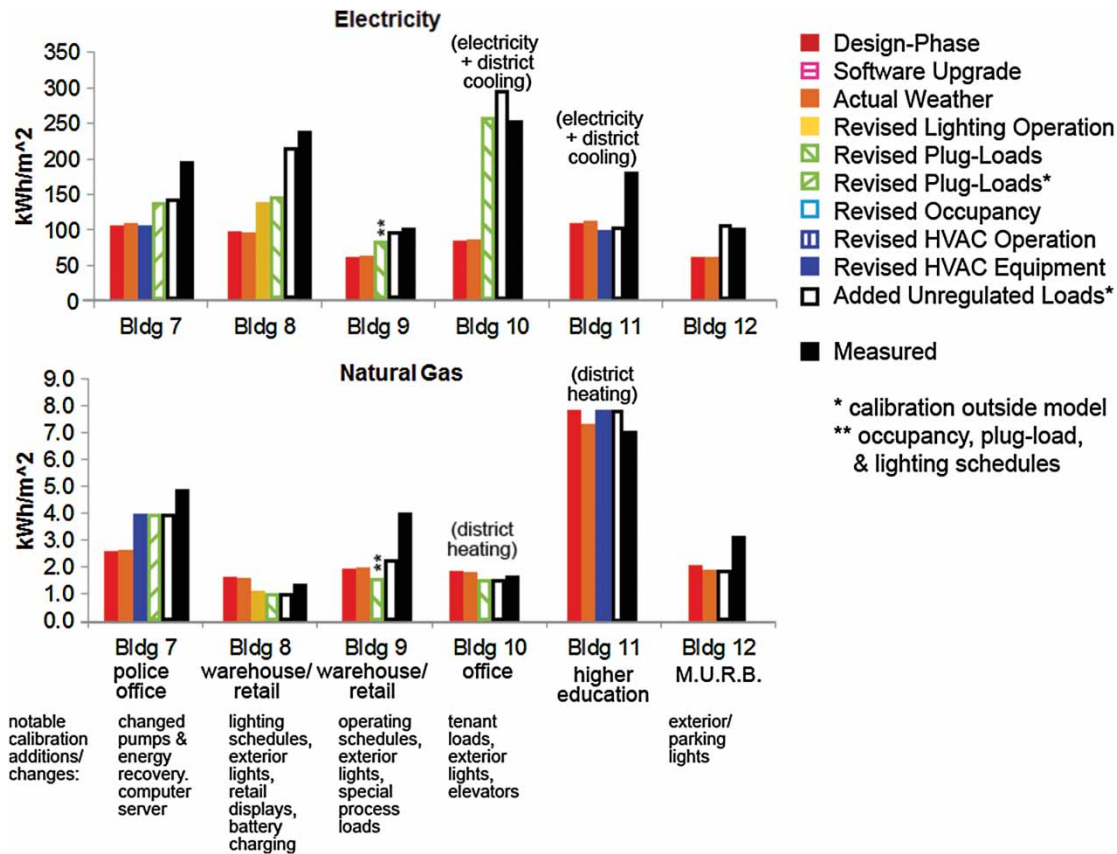
Figure 6.   Buildings 7–12 measured versus simulated annual energy at each calibration step.

researchers have shown that the typical approach to estimating window U-factors used here, based on standard test dimensions rather than actual window sizes, may be inaccurate (Hirsch et al. 2011). Software limitations also remained unaddressed. Notably, modellers in EE4 cannot model demand controlled ventilation (DCV) and instead implement a work-around, described in the EE4 manual, to reduce the fan operating hours and thus the amount of outside air being heated or cooled. As shown in Table 2, eight of the studied buildings include DCV. The DOE2 engine also poses challenges to accurately representing ventilation energy recovery (ERV) by restricting one's ability to model upstream preheat coils and reducing the ERV effectiveness to a single number which may not be accurate under all control situations. All 18 buildings included ERV. Furthermore, researchers have also shown that DOE2 may overestimate boiler efficiency at low loads (Kenna and Bannister 2009). Twelve cases had boilers. The authors recommend using a more flexible software package for future modelling and calibration efforts.

4.3.   Accuracy of compliance models

A sobering takeaway from these case studies is that the difference between simulated and measured energy use in certified green buildings is actually quite large, with a median of 31% for the whole group and reaching as

high as 59% in the worst case. This study shows that routinely voiced expectations that future energy use can be predicted to within 10–20% during design should be upwardly adjusted. Notably, the bulk of these discrepancies were not caused by the simulation algorithms, the availability of good weather data, or the modellers' ability to reliably model building envelope properties. The main discrepancies were rather a result of the compliance modelling protocols as well as how buildings are used and operated, which, to some extent, could be anticipated in the design phase. For example, in Building 1, a small ecology centre and the building most impacted by weather variation, switching from the 1998 (extreme warm year) to 1972 (extreme cool year) weather increased the annual EUI by 23%. This is large effect, but it nevertheless pales in comparison to the impact of revising the plug loads in Buildings 10 and 18, which increased the EUI by over 140% in each case.

Modelling protocols in North America require the modeller to compare his/her as-designed building to a code-compliant baseline model, identical to the former except for the energy conservation measures. This approach penalizes designers for high loads that fall outside of their design control. This partially explains why modelling protocols traditionally excluded certain loads or de-emphasized custom plug-load inputs, and why some
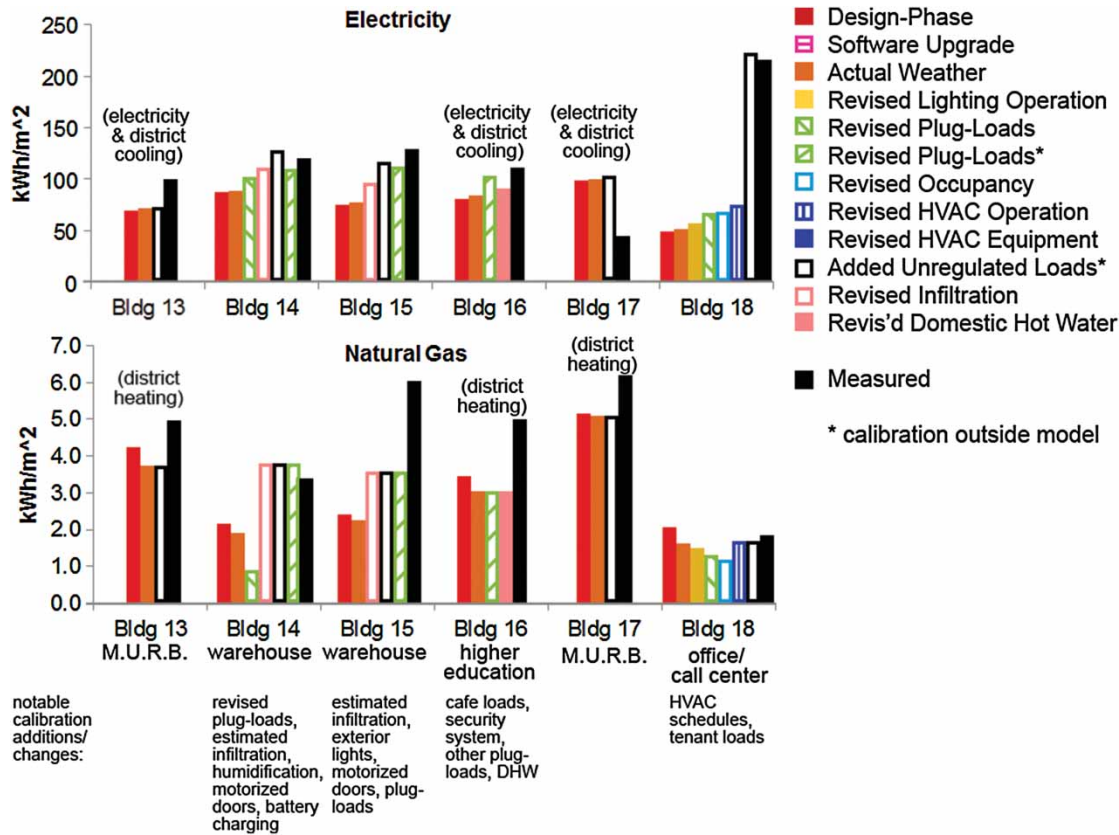
Figure 7.   Buildings 13–18 measured versus simulated annual energy at each calibration step.

Table 2.   The calibration steps implemented, their impact, and the effort involved.

Calibration Step	Number of Cases Including this Step	Median [mean] Change in EUI, where implemented[a]	Effort Involved
Replacing typical with actual weather	18	− 2% [− 2%]	Easy to moderate
Revising plug loads	14	15% [32%]	Via spreadsheet = easy[b] In model = moderate[b]
Adding unregulated loads	15	5% [7%]	Via spreadsheet. easy
Revising occupant density/schedules	4	− 6% [− 4%]	Moderate[b]
Revising lighting density/schedules	3	11% [11%]	Moderate[b]
HVAC Updates (not including schedules)	3	7% [9%]	Depends
Infiltration	2	58% [58%]	Easy (to manipulate in text file)
HVAC schedules	1	20% [20%]	Moderate
Revising domestic hot water	1	− 6% [− 6%]	Via spreadsheet. easy

[a]Mean Change in EUI from Previous Step $= \left(\left(\sum_{i=1}^{n} (a_i - b_i)/b_i \right) /n \right)$, where
a_i = predicted EUI after this calibration step, b_i = predicted EUI after the prior calibration step, and n = number of projects including this calibration step.
[b]if information is available. The effort level varies. More zones in model = more effort.

modellers may hesitate to overestimate these loads. The 5% median increase caused by adding unregulated loads found in this study further underscores the need to include all loads in the design-phase model, if possible. Policy-makers have begun requiring modellers to include custom plug-load and other miscellaneous load predictions in the model (ASHRAE & IESNA 2010) in an effort to take a more comprehensive view of energy performance and to bring compliance-type model predictions closer to reality. More research is needed to determine whether modellers
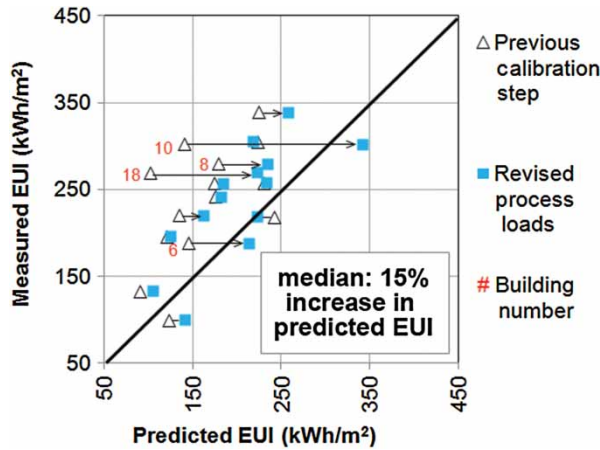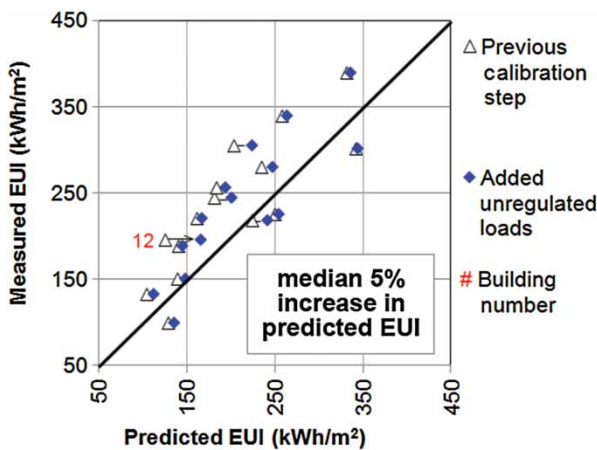
Figure 8. Effect of revising plug loads.

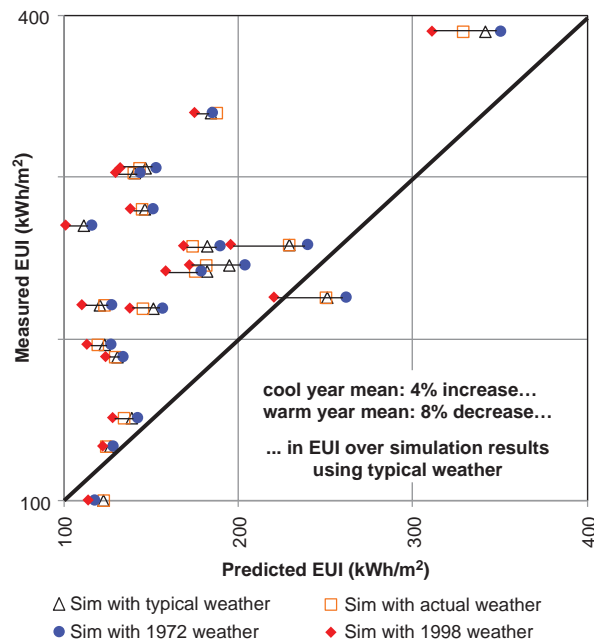Figure 9. Effect of adding unregulated loads.

Figure 10. Weather sensitivity analysis.

are actually capable of predicting these loads accurately in the design phase.

4.4. Building sub-metering

Where installed, building sub-metering provided valuable data for calibration. Nevertheless, in this research, sub-metering seldom supplied data in the desired level of detail for input into the model. Translating the coarse-grained measured data into more granular end uses and building zones required effort and judgement. Software developers can help address this difficulty. In addition, the practice of building sub-metering may grow more prevalent and granular with the increasing affordability of equipment (Claridge 2011) and the influence of codes and standards (ANSI et al. 2010; US Green Building Council 2009), which in turn will decrease the calibration effort level even further.

4.5. Calibration triage

In these cases, the remaining predicted versus measured energy discrepancies helped to guide the search for operational problems in the buildings. For example, according to the commissioning agent for Building 8, shown in Figure 3, the simulation helped to uncover a problem with the building automation system, a fault in the energy recovery ventilation equipment, and unintentional off-hours lighting. Higher quality models and calibrations would help project teams uncover more operational problems like these, but at an increased cost. The ASHRAE 14 standard (ASHRAE 2002) lists goodness-of-fit suggestions for calibrated models. In these 18 cases, none of the final, partially calibrated simulations met all of the ASHRAE 14 suggestions. On the one hand, higher quality models and calibrations may have uncovered more building problems and provided value for other applications such as ongoing commissioning. On the other hand, building owners may hesitate to invest in an extensive model calibration process to start the investigation, since the process might not lead to remunerable savings if the building already performs generally as it should. A coarse first-pass calibration approach can provide a cost-effective means to identify projects that warrant further calibration effort, just as triage helps to prioritize care in a hospital emergency room.

The simulated/measured discrepancy can shed light on the potential magnitude of the problem, as illustrated in Figure 11. This graph highlights the unexplained energy expenditures of Building 13, the worst case, in which the building consumed an extra $188,000 (USD) per year in electricity and natural gas compared to the partially calibrated model. For building owners, who must weigh the cost of the investigation and remediation against the potential operational savings, this type of estimate can help decide whether it is worthwhile to proceed with a more detailed investigation or not.
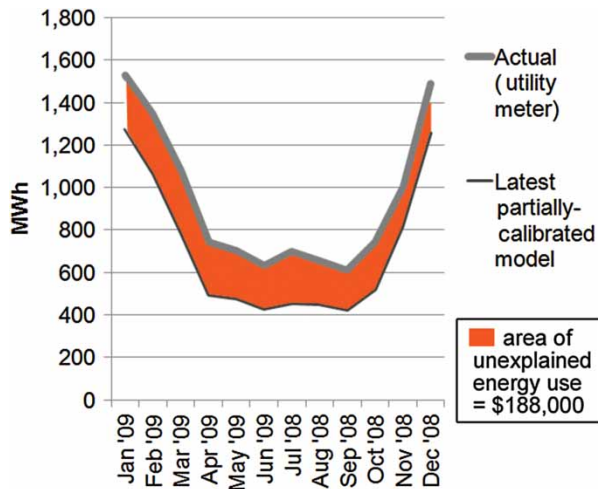
Figure 11. Building 13 theoretical potential for improvement.

5. Conclusion

In this research, the authors analysed the partial calibration process of design-phase energy models, performed within the context of for-profit projects. Practitioners originally built the 18 models to demonstrate compliance with the LEED Canada version 1.0 rating system in buildings constructed between 2008 and 2012. The partial calibration effort focused on adding value to the design-phase models within limited budgets and schedules. The modellers performed the original calibration steps in an ad hoc fashion, and the authors re-created the process systematically in order to discover the calibration tasks that provided the highest impact for the least effort. In aggregate, the partial calibration process improved the annual and monthly mean bias errors of the design-phase models by more than 50%. The bulk of this improvement came from revising the plug loads and adding unregulated loads. In general, the impact of each of these calibration steps far exceeded that of the change from CWEC to historic weather data, and each of these calibration steps required relatively little effort to implement.

Acknowledgements

Notes

1. This includes the 92 buildings for which measured and predicted EUI data were available.
2. The authors have made available the custom Python scripts created for this research to batch process DOE (.sim) results files and extract key monthly and annual results. These scripts may be helpful to EE4, eQuest, and other .sim file users. The public can download them here: https://gist.github.com/hollywas/4757770 and here: https://gist.github.com/alexstorer/4219834.
3. The modellers followed the EE4 modelling protocol (Natural Resources Canada Office of Energy Efficiency & CANMET Energy Technology Centre 2008) building zoning principles, thereby creating separate model zones per HVAC system, space operation/function, heating/cooling loads, and solar orientation.
4. Compliance-type model: energy model used to demonstrate compliance with a building code or standard.
5. For example, in Building 8, the added retail displays not only consumed electricity directly, which the modeller tallied in the spreadsheet, but also added heat to the space, which would be ignored with this method.

References

Adam, C., P. Andre, C. Aparecida Silva, J. Hannay, and J. Lebrun. 2004. "Commissioning-Oriented Building Loads Calculations. Application to the CA-MET Building in Namur (Belgium)." Paper presented at the international conference for enhanced building operations, Paris, France.
Ahmad, Mushtaq, and Charles Culp. 2006. "Uncalibrated Building Energy Simulation Modeling Results." *HVAC&R Research* 12 (4): 1141–1155.
Andre, P., C. Cuevas, P. Lacote, and J. Lebrun. 2003. "Re-commissioning of the CAMET HVAC system: A successful case study." Paper presented at the international conference for enhanced building operations, Berkeley, CA.
ANSI, ASHRAE, USGBC, & IES. 2010. *189.1 Standard for the Design of High-Performance Green Buildings*. Atlanta, GA: ASHRAE.
ASHRAE. 2002. *Guideline 14 Measurement of Energy and Demand Savings*. Atlanta, GA: ASHRAE.
ASHRAE, & IESNA. 2010. *90.1 Energy Standard for Buildings Except Low-Rise Residential*. Atlanta, GA: ASHRAE.
Bannister, P. 2009. "The Application of Simulation in the Prediction and Achievement of Absolute Building Energy Performance." Paper presented at the building simulation, Glasgow, Scotland.
California Energy Commission. 2007. Title 24, Part 6, of the California Code of Regulations: California's Energy Efficiency Standards for Residential and Nonresidential Buildings.
Canada Green Building Council. 2007. LEED Canada-NC (and CS) 1.0 Green Building Rating System.
Candian Commission on Building and Fire Codes. 1997. Model National Energy Code of Canada for Buildings. Ottawa.
Carling, P., P. Isakson, and P. Blomberg. 2004. Whole-Building Simulation in Commissioning – Experiences from the Katsan-Project (Vol. 2003): International Energy Agency ECBCS Annex 40.
Claridge, D. 2011. "Performance Simulation for Design & Operation." In *Building Simulation for Practical Operational Optimization*, edited by Hensen, J. and R. Lamberts, 365–401. London: Spon Press.
Clarke, J. A., P. A. Strachan, and C. Pernot. 1993. "An Approach to the Calibration of Building Energy Simulation Models." *ASHRAE Transactions* 99 (2): 917–927.
Crawley, D. 2008a. "Building Performance Simulation: A Tool for Policymaking." PhD, University of Strathclyde, Glasgow, Scotland. https://www.strath.ac.uk/media/departments

/mechanicalengineering/esru/research/phdmphilprojects/crawley_thesis.pdf

Crawley, D. 2008b. "Estimating the Impacts of Climate Change and Urbanization on Building Performance." *Building Performance Simulation* 1 (2): 91–115.

Duffie, J. A., and W. A. Beckman. 1991. *Solar Engineering of Thermal Processes*. 2nd ed. Hoboken, NJ: Wiley-Interscience.

Efficiency Valuation Organization. 2006. International Performance Measurement & Verification Protocol (IPMVP), Volume III *Part I*. Washington, DC: Efficiency Organization.

Environment Canada. 2001. *Canadian Weather Energy and Engineering Data Sets* (*CWEEDS Files*).

Haberl, J., and T. Bou-Saada. 1998. "Procedures for Calibrating Hourly Simulation Models to Measured Building Energy and Environmental Data." *ASME Journal of Solar Energy Engineering* 120 (August): 193–204.

Hirsch, A., S. Pless, R. Guglielmetti, and P. Torcellini. 2011. The Role of Modeling When Designing for Absolute Energy Use Intensity Requirements in a Design-Build Framework: National Renewable Energy Laboratory.

Hirsch, James J., and Regents of the University of California. 1999. DOE-2.1E.

Kenna, E., and P. Bannister. 2009. "Simple, Fully Featured Boiler Loop Modelling." Paper presented at the building simulation, Glasgow, Scotland.

Liu, M., D. E. Claridge, N. Bensouda, K. Heinemeier, S. Lee, and G. Wei. 2003. "Manual of Procedures for Calibrating Simulations of Building Systems." In California Energy Commission (Ed.), *High Performance Commercial Building Systems*, 1–53. http://energy.lbl.gov/BT/HPCBS/pubs/E5P23T2b.pdf

Liu, G., and M. Liu. 2011. "A Rapid Calibration Procedure and Case Study for Simplified Simulation Models of Commonly Used HVAC Systems." *Building & Environment* 46 (2): 409–420. doi:10.1016/j.buildenv.2010.08.002

Liu, M., G. Liu, D. Claridge, and J. Haberl. 2006. ASHRAE 1092-RP Report: Development of Procedures to Determine In-Situ Performance of Commonly Used HVAC Systems.

Mills, E., H. Friedman, T. Powell, N. Bourassa, D. Claridge, T. Haasl, and M. A. Piette. 2004. "The Cost-effectiveness of Commercial-buildings Commissioning: A Meta-analysis of Energy and Non-energy Impacts in Existing Building and New Construction in the United States". Lawrence Berkeley National Laboratory Report No. 56637. http://evanmills.lbl.gov/pubs/pdf/cx-costs-benefits.pdf

Natural Resources Canada. 2005. EE4 (Version 1.7 Build 2). Ottawa, Ontario: Natural Resources Canada. http://canmetenergy.nrcan.gc.ca/software-tools/ee4/754

Natural Resources Canada Office of Energy Efficiency, and CANMET Energy Technology Centre. 2008. EE4 Software Version 1.7 Modelling Guide.

O'Neill, Z., M. Shashanka, X. Pang, P. Bhattacharya, T. Bailey, and P. Haves. 2011. "Real Time Model-Based Energy Diagnostics in Buildings." Paper presented at the building simulation conference, Sydney, Australia.

Pan, Y., Z. Huang, and W. Gang. 2007. "Calibrated Building Energy Simulation and Its Application in a High-Rise Commercial Building in Shanghai." *Energy and Buildings* 39 (6): 651–657.

Piette, M. A., B. Nordman, and S. Greenberg. 1994, May. "Quantifying Energy Savings from Commissioning: Preliminary Results from the Pacific Northwest." Paper presented at the second national conference on building commissioning, St. Petersburg Beach, Florida.

Raftery, Paul, Marcus Keane, and James O'Donnell. 2011. "Calibrating Whole Building Energy Models: An Evidence-Based Methodology." *Energy and Buildings* 43 (9): 2356–2364. doi:10.1016/j.enbuild.2011.05.020

Reddy, T. 2006. "Literature Review on Calibration of Building Energy Simulation Programs: Uses, Problems, Procedures, Uncertainty, and Tools." *ASHRAE Transactions-Research* 112 (1): 226–240.

Reddy, T., I. Maor, and C. Panjapornpon. 2007. "Calibrating Detailed Building Energy Simulation Programs with Measured Data-Part I: General Methodology." *HVAC&R Research* 13 (2): 221–241.

Samuelson, H. W., A. Lantz, and C. F. Reinhart. 2012. "Non-technical Barriers to Energy Model Sharing and Reuse." *Building and Environment* 54: 71–76. doi:10.1016/j.buildenv.2012.02.001

Song, L., M. Liu, D. Claridge, and P. Haves. 2003, September 17–20. "Study of On-Line Simulation for Whole Building Level Energy Consumption Fault Detection and Optimization." Paper presented at the building integration solutions, architectural engineering, Austin, Texas, USA.

Stationery Office London. 2010. UK Building Regulations 2010, Conservation of Fuel and Power, Approved document L1A Conservation of Fuel and Power in Dwellings. London.

Turner, C., and M. Frankel. 2008. *Energy Performance of LEED for New Construction Buildings*. Washington, DC: US Green Building Council.

US Green Building Council. 2001. LEED-NC Version 2.0. Washington, DC, USA.

US Green Building Council. 2009. LEED-NC Version 3.0. Washington, DC, USA.

Wei, G., M. Liu, and D. E. Claridge. 1998. "Signatures of Heating and Cooling Energy Consumption for Typical AHUs." Paper presented at the eleventh symposium on improving building systems in hot and humid climates, Fort Worth, TX.