Applied Energy 203 (2017) 348–363

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Assessing the value of information in residential building simulation: Comparing simulated and actual building loads at the circuit level

Brock Glasgo [a,*], Chris Hendrickson [b], Inês Lima Azevedo [a]

[a] Engineering and Public Policy, Carnegie Mellon University, 129 Baker Hall, 5000 Forbes Ave., Pittsburgh, PA 15213, United States
[b] Civil and Environmental Engineering, Carnegie Mellon University, 3050 Hamburg Hall, 5000 Forbes Ave., Pittsburgh, PA 15213, United States

HIGHLIGHTS

- 106 homes are simulated in EnergyPlus using energy audit and survey records.
- Simulation results are compared to monitored data at the device level.
- Modeling reveals large discrepancies between simulated and actual energy use.
- Sensitivity analysis is used to identify factors most important for accurate models.
- The role of EnergyPlus in residential energy code design and analysis is discussed.

ARTICLE INFO

Article history:
Received 20 January 2017
Received in revised form 17 May 2017
Accepted 26 May 2017
Available online 20 June 2017

Keywords:
EnergyPlus
Residential
Simulation
Energy code

ABSTRACT

Building energy simulation tools are now being used in a number of new roles such as building operation optimization, performance verification for efficiency programs, and – recently – building energy code analysis, design, and compliance verification in the residential sector. But increasing numbers of studies show major differences between the results of these simulations and the actual measured performance of the buildings they are intended to model. The accuracy and calibration of building simulations have been studied extensively in the commercial sector, but these new applications have created a need to better understand the performance of home energy simulations.

In this paper, we assess the ability of the DOE's EnergyPlus software to simulate the energy consumption of 106 homes using audit records, homeowner survey records, and occupancy estimates taken from monitored data. We compare the results of these simulations to device-level monitored data from the actual homes to provide a first measure of the accuracy of the EnergyPlus condensing unit, central air supply fan, and other energy consumption model estimates in a large number of homes. We then conduct sensitivity analysis to observe which physical and behavioral characteristics of the homes and homeowners most influence the accuracy of the modeling.

Results show that EnergyPlus models do not accurately or consistently estimate occupied whole-home energy consumption. While some models accurately predict annual energy consumption to within 1% of measured data, none of the modeled homes meet ASHRAE criteria for a calibrated model when looking at hourly interval data. The majority of this error is due to appliance and lighting energy overestimates, followed by AC condensing unit use. These inaccuracies are due to factors such as occupant behaviors and differences in appliance and lighting stocks which are not well-captured in traditional energy audit reports. We identify a number of factors which must be specified for an accurate model, and others where using a default value will produce a similar result.

The use of building simulation tools reflects a shift from a component-focused approach to a systems approach to residential code analysis and compliance verification that will serve to better identify and deploy efficiency measures in homes. By better understanding the limitations of home energy simulations and adopting strategies to mitigate the effects of model errors, simulation models can serve as valuable decision making tools in the residential sector.

* Corresponding author.
E-mail addresses: bglasgo@andrew.cmu.edu (B. Glasgo), cth@andrew.cmu.edu (C. Hendrickson), iazevedo@cmu.edu (I.L. Azevedo).

http://dx.doi.org/10.1016/j.apenergy.2017.05.164
0306-2619/

1. Introduction

Increased attention to building energy performance, improved software packages, and decreasing computing power requirements have led to the use of building energy simulation tools in a large and growing number of applications [1–3]. These tools are now being used in their traditional role as decision support for building and retrofit design in the commercial sector, but also in new roles such as building operation optimization [1], performance verification for energy efficiency programs like LEED [4], and – recently – building energy code analysis, design, and compliance verification in the residential sector [3,5]. However, increasing numbers of studies show major differences between the results of these simulations and the actual measured performance of the buildings they are intended to model [6–13]. These discrepancies, combined with the new application of building simulation tools to policy and investment decision-making in the residential sector, have created a need to better understand the accuracy of their results and develop methods for calibrating them to ensure reliable outputs. Doing so will allow policymakers to apply these tools in a way that will ensure that residential building energy codes continue to deliver the energy savings for which they are intended.

The Department of Energy's EnergyPlus is the most prominent simulation package being used in these new residential applications. As part of their work for the Building Energy Codes Program, Pacific Northwest National Laboratory (PNNL) established a method for analyzing potential changes to residential building codes based on EnergyPlus [3]. The method first involved the construction of prototype EnergyPlus models of simple single- and multi-family residences that meet existing region-specific building codes. The cost-effectiveness of potential changes to these codes is evaluated by incorporating a proposed change in the model – reducing allowable building leakage rate, for instance – simulating the building's energy performance using local weather data, and observing the resulting change in energy consumption. The simulated energy cost savings are then compared to the first cost to estimate the lifecycle cost of implementing the change. These results then serve to inform the DOE's position on whether to approve a code change proposed by the International Code Council (ICC), but are also used to inform state and local jurisdictions about the expected effects of adopting a new code when they are considering a change.

EnergyPlus is also in the process of being incorporated into a tool being developed by the Residential Energy Services Network (RESNET) to standardize residential energy benchmarking for energy code compliance. RESNET is a not-for-profit membership corporation that develops standards used in home energy efficiency ratings [14]. RESNET's Home Energy Rating System (HERS) is an industry standard calculation specification that allows certified energy raters to assign efficiency scores to homes that can be used to demonstrate their energy code compliance in most states and jurisdictions [5,15]. Efficiency scores are currently calculated using any one of a number of software programs that have been approved and accredited by RESNET [16]. In March of 2016, however, RESNET and the DOE announced that this suite of software packages is going to be replaced by a single-source tool based on EnergyPlus [5]. While the tool has not yet been released or described in detail, its announcement alone highlights the need to better understand the ability of EnergyPlus to accurately model residential buildings.

Each release of EnergyPlus is thoroughly validated using three types of methods [17]. Analytical verification compares EnergyPlus results to mathematically determined results for individual building components and systems. Comparative testing compares EnergyPlus simulation results to the results of other simulation packages. These two validation methods are described in a variety of technical standards including ANSI/ASHRAE Standard 140, *Method of Test for the Evaluation of Building Energy Analysis Computer Programs*. Finally, empirical validation compares simulation results with measured energy consumption from actual buildings.

Previous empirical studies of the accuracy of building energy simulations have focused almost exclusively on the commercial sector and have often found large discrepancies between modeled and actual performance. These studies typically involve the construction of a model of a building in which extensive data gathering has been conducted. Using measured and observed details of the building and its operation, a detailed model is constructed and its simulated performance is compared to measured data such as electric or gas utility data [10,17], environmental sensor data [11], or submetered system-level data [8,12,13]. Models are often then modified to observe the effects of varying certain parameters to observe their effect on simulated energy consumption [17]. Based on these results, conclusions are drawn about which parameters are most important to specify and the suitability of the chosen model and application, and recommendations are made to improve modeling efforts in the future. The results of some of these comparison studies have called into question the basic ability of simulation tools to predict energy use in buildings given all of the uncertainties involved in building an accurate model [17].

In addition to empirical validation efforts, there are a growing number of papers dedicated solely to the methods by which these models can be calibrated. Coakley et al. summarized these methods in a literature review of around 70 papers addressing issues of calibration in building simulation modeling [18]. The authors propose four classes of calibration methods: (i) calibration based on manual, iterative and pragmatic intervention, (ii) calibration based on a suite of informative graphical comparative displays, (iii) calibration based on special tests and analytical procedures, and (iv) analytical and mathematical methods of calibration. The paper generally finds no consensus method for building simulation calibration, nor does it find a widely accepted set of criteria for validating these models. However, given the large body of literature found by the authors, they conclude that the work already available could inform the development of standardized methods for model calibration. Recently, attention has turned to the development of automated model calibration methods that rely less on modeler expertise and more on mathematical and analytical approaches [19]. These methods generally use optimization tools to minimize an error term between simulated and actual data by tuning specified model parameters [8,19].

Both empirical validation studies and calibration studies are typically limited by data availability to a small number of buildings. The conclusions that can be reached from such studies are therefore limited as well. To address this issue and increase sample sizes, research is now turning to batch simulations in which large numbers of buildings are modeled in parallel. Rhodes et al. used one such method to simulate 54 homes in the Pecan Street study using energy audit and survey records as model inputs [20]. A baseline model of each home was built using actual building characteristics and simulated using Typical Meteorological Year (TMY) data. Three alternate scenarios were then simulated which (1) used actual weather data, (2) updated default thermostat settings with actual thermostat settings, and (3) simplified each home's geometry into a rectangular footprint. Each set of simulation results were compared to measured whole-home annual electricity consumption. Results indicate that including actual thermostat settings improves model accuracy, actual weather data unexpectedly worsened accuracy, and simplifying home geometries had little effect on outcomes. Errors for individual homes ranged from underestimating actual annual consumption by 60% to overestimating by over 100%. However, when results are aggregated to measure the model's ability to predict the combined electricity consumption of all the homes, errors are reported as less than 3%.

In this paper, we advance this approach by modeling and simulating 106 homes from the same Pecan Street study used by Rhodes et al. [20]. Simulating a large number of actual, occupied homes in parallel ensures that a wide range of physical home characteristics and occupant behaviors are included in our modeling. Doing so allows us to draw conclusions that are applicable to a broader population of homes than if our results were based on any single home, its occupants, and their energy consumption. We use PNNL's residential prototypes as a starting point, then modify these prototypes with information from Pecan Street's energy audit records, homeowner survey results, and occupancy profiles estimated from device-level energy consumption data to more closely resemble the actual monitored homes. We compare the results of these simulations to device-level monitored data from the actual homes to provide the first ever empirical validation of the EnergyPlus condensing unit, central air supply fan, and other energy consumption model estimates in a large number of homes. By comparing simulated and actual energy consumption at this detailed level, we are able to more closely identify the source of EnergyPlus model errors and provide detailed recommendations on how to address them and mitigate their effect on decision-making in the residential sector.

We then conduct a sensitivity analysis to observe which physical and behavioral characteristics of the homes and homeowners most influence the accuracy of the modeling. These results provide context for the use of EnergyPlus as a decision-making tool in the residential sector. Specifically, they show the approximate level of accuracy that researchers can expect when simulating homes displaying a range of physical characteristics, appliance stocks, and occupant behaviors. Sensitivity analysis results also provide a measure of which characteristics most influence model accuracy and need to be included for accurate system-level modeling.

The rest of this paper is organized as follows. Section 2 describes the data, modeling methods, and assumptions used in the analysis. Section 3 presents results of the modeling and sensitivity analysis. Sections 4 and 5 provide a discussion of these results, conclusions reached, and policy implications.

2. Material and methods

2.1. Residential building prototypes

As a starting point for our modeling, we use PNNL's single-family detached home EnergyPlus prototypes built according to the IECC 2006 residential building energy code. Prototype models are available that were built to simulate homes compliant with IECC 2006, 2009, and 2012 [21]. The oldest available prototypes were chosen to more closely match the older Pecan Street building stock. Prototypes are also differentiated by location to account for variations in building energy codes by climate zone. Homes in the Pecan Street sample are located in and around Austin, Texas, so we select prototypes designed for San Antonio, which is located in the same climate zone [22]. A summary of some of the key characteristics of these prototypes can be found in Table 1.

2.2. Appliance-level energy use data

Appliance- and home-level energy consumption data were obtained from the Pecan Street Research Institute's Dataport for the year 2015 [23]. Pecan Street Inc. is a 501(c)(3) not-for-profit corporation and research institute headquartered at The University of Texas at Austin. Volunteers from in and around Austin elect to join the study and work with researchers at Pecan Street to decide which circuits and devices in their homes to monitor. The resulting dataset includes records for approximately 722 homes, with data available for up to 28 circuits per home at 15-min intervals. We apply validation criteria which require at least one full year of whole-home use data with less than one week of missing values.

To ensure a fair comparison between EnergyPlus simulations and monitored data, energy consumption from electric vehicles, garages, pool lights, pool pumps, and sprinklers are subtracted from monitored whole-home consumption when these devices were monitored. These devices are not modeled in EnergyPlus simulations and would otherwise be a source of error.

Average whole-home electricity consumption in the Pecan Street sample is around 33% less than comparable homes in the EIA's 2009 Residential Energy Consumption Survey (RECS) and around 13% less than the average Austin Energy customer

Table 1
Summary statistics for PNNL prototype models and simulated Pecan Street homes.

	PNNL prototype	Pecan street		
	Value	Min	Mean	Max
Occupant characteristics				
Residents (qty)	3.0	1.0	2.6	6.0
Household income ($/yr)	–	$30k	$130k	$230k
Building envelope characteristics				
Number of floors	2	1	1.5	2
Foundation type (slab/pier)	Varies	–	94/12	–
Home age (yrs)	~10	6	28	96
Area (ft^2)	2400	800	2100	4000
Ceiling height (ft)	8.5	7.5	9	14
Building infiltration (ELA, cm2)	960	310	960	2400
Window U-factor (W/m^2 K)	4.3	2.2	4.1	7
Attic insulation R-value (m^2 K/W)	4.3	1.1	5.3	9.2
Appliance characteristics				
Condensing unit efficiency (EER)[a]	13.5	5.0	10.9	17.0
Condensing unit age (yrs)	0	4	10.1	27
Heating setpoints (daily avg., °F)	72	63	69	75
Cooling setpoints (daily avg., °F)	75	68	77	82
Heat pump/gas furnace (qty/qty)	Varies	–	7/99	–
Programmable thermostat (y/n)	Yes	–	98/8	–
Water heater fuel (gas/electric)	Varies	–	100/6	–

[a] Nameplate efficiency.

[24,25]. Thus, while the sampled homes are more efficient than the average Texas home, they are likely to provide a reasonable estimate of household electric consumption around Austin.

2.3. Energy audit and homeowner survey records

Energy audit and homeowner survey records are available for many homes in the Pecan Street sample. As part of their ongoing research, the Pecan Street Research Institute has implemented several interventions in volunteer residents' homes and apartments. These include providing residents access to an online portal to observe their energy use, simulating time-of-use pricing schemes, and providing new appliances to homeowners, among others. A description of these programs and the number of participants in each can be found in Appendix A.

Energy audits were conducted and recorded in the Pecan Street dataset between January 2011 and September 2014 as part of two separate programs. The majority of audits were conducted by an outside contractor before monitoring installations began in January 2012. The remaining audits, conducted in late 2013 and early 2014, were conducted by Pecan Street personnel. Annual surveys are administered to participants in the Pecan Street study. Records from these surveys provide demographic and other information about the homes and their occupants.

In selecting homes to be included in final simulations we include only those which meet the monitored data validation criteria described above, and which also have an energy audit record and at least one annual survey record. A total of only 106 homes remain that meet these requirements. Of these, 102 have complete condensing unit data and 75 have complete central air supply fan data.

A summary of key data collected in the audits, surveys, and monitoring installations is shown below in Table 1.

2.4. Occupancy estimation

Home occupancy is an important determinant of the timing and quantity of energy consumption in homes. Survey results provide an incomplete accounting of occupied hours for each home, so we estimate occupancy based on device-level monitored data.

To do this, monitored circuits and devices are separated into accompanied and activated loads. Accompanied loads are loads which indicate that the home is likely occupied if they are consuming energy. Activated loads are loads which can be consuming energy even if the home is unoccupied. For these loads, occupancy can only be estimated by looking for events where the device or circuit sees a significant change in load, indicating that someone has activated or deactivated the circuit. See Table 2 for a list of accompanied and activated loads. Note that loads which do not vary significantly based on occupancy, such as refrigerators or air conditioners, are not included in either load class and are not used to estimate occupancy with this method.

Using these allocations, annual energy consumption profiles for each appliance or circuit in the 106 homes are used to estimate occupancy for every 15-min interval for the year 2015. If any accompanied load in a home is consuming over 50 W in a given interval, the home is flagged as being active in that period with *P(Active) = 1.0*. Similarly, if any activated load sees an increase or decrease in demand of over 50 W, the home is flagged as being active in that period with *P(Active) = 1.0*. If no activity is identified, the home is flagged as inactive with *P(Active) = 0*. Averaging these activity profiles for every day in 2015, we generate a probabilistic daily activity profile for each home. For example, if a home is flagged as active every other day at 10 AM, the probabilistic daily activity profile will have *P(Active) = 0.5* at 10 AM for that home.

Because this method relies on device activity, it fails to identify hours where homes are likely occupied, but the residents are inactive or sleeping. To correct for this, we assume that most homes follow a typical occupancy pattern of waking up between 3 AM and 11 AM and arriving home between 6 PM and midnight. To estimate inactive – but likely occupied – hours, peak activity is identified for each of these periods. Prior to peak activity in the morning, and following peak activity in the evening, it is assumed the home is always occupied with *P(Occupied) = 1.0*. Between each home's wakeup and arrival hours, it is assumed that activity actually reflects occupancy, with *P(Occupied) = P(Active)*.

A sample home showing the difference between the calculated activity profile and estimated occupancy profile is shown in Fig. 1. This method is used to estimate an occupancy profile for each home in the final sample.

2.5. Modeling operations

EnergyPlus version 8.4.0 was used to run this study's simulations. Fully specified EnergyPlus models of the 106 Pecan Street homes are generated by modifying the PNNL prototypes' input data files (IDFs) with the home and occupant characteristics described above. These IDFs are text files which describe the physical, operational, and behavioral characteristics of the modeled homes and their occupants. To modify the prototype IDFs, a Matlab program was used to open each file, locate the fields to be edited, and replace the default values or descriptions with the actual home's characteristics. The characteristics for which we have data from either Pecan Street's energy audit or survey records are listed and described below.

- Number of floors.
- Primary space heating fuel – PNNL prototypes are available with heating supplied by a heat pump, electric resistance heating, an oil furnace, or a gas furnace. Pecan Street audit records are used to assign the correct heat source for each home.
- Foundation type – Foundations in the Pecan Street homes are all either slab or pier-and-beam. Without additional information to specify slab thicknesses or other details, we rely on the baseline slab and pier-and-beam foundations defined by PNNL in their prototypes. The ground heat transfer method used is the preprocessor method, which is also the default used in the PNNL prototypes. More details on these factors and the prototypes can be found in [21].
- Building square footage – Building square footage was available for each of the 106 homes. Length, width, and building footprint shape are not specified, so we model Pecan Street homes as

Table 2
Appliance type allocations for occupancy estimation.

Accompanied loads	Activated loads
Electric vehicles, laundry machines, dishwashers, in-sink disposals, microwaves, ovens, electric ranges	Bathroom circuits, bedroom circuits, dining room circuits, garage circuits, kitchen circuits, lighting circuits, living room circuits, office circuits, outdoor lighting circuits, pool lighting, utility room circuits, ventilation hoods

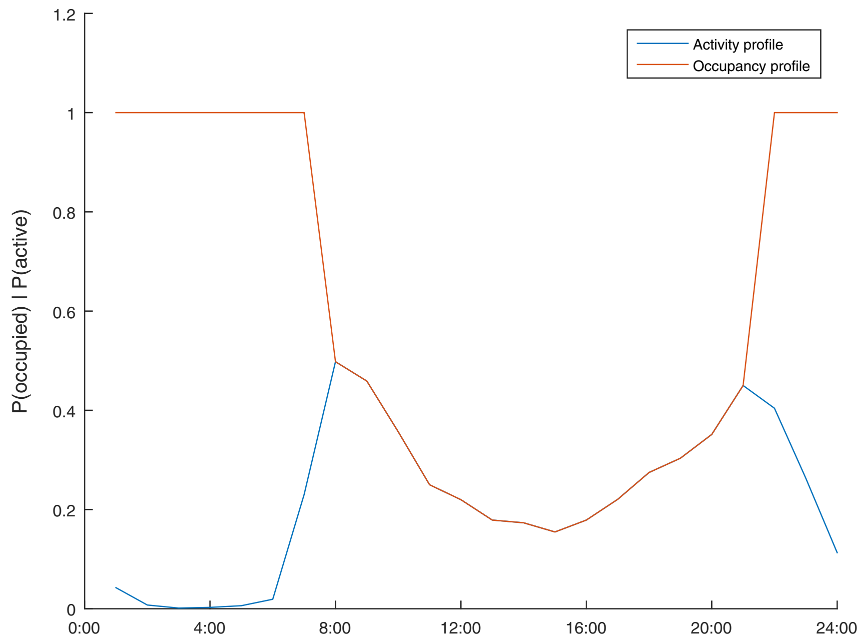


Fig. 1. Estimated activity and occupancy profiles for a single home where morning activity peaked at 8 AM and evening activity peaked at 9 PM.

rectangular homes with width 1.4 times the length. This is a reasonable assumption, as Rhodes et al. showed in [20] that simplifying the geometry of homes does not significantly impact EnergyPlus modeling performance.

- Attic insulation R-value – Attic insulation value was recorded in energy audit records. When data was missing, it was replaced with the average value from the remaining homes.
- Ceiling height – When ceiling height was not reported for an individual home, it was replaced with the average value from the remaining homes.
- AC condensing unit capacity – Nameplate condensing unit capacity is known for all 106 homes in the sample.
- AC condensing unit efficiency – In addition to the nameplate efficiency of each home's condensing unit, the age of each unit was known. To account for performance degradation over time, we estimate actual operating efficiency according to the method described in [26].
- Water heater fuel – Water heater fuel was determined based on energy audit reports and the availability of monitored data.
- Building shell infiltration – Building shell infiltration was reported in Pecan Street energy audits as the result of a blower door test conducted on each home. Results were reported in units of air changes per hour at 15 Pa. PNNL's prototypes specify infiltration in terms of equivalent leakage area. To convert between the two units, we use the method described in [27]. The opening of windows and doors – either as an intentional means of temperature control or simply as occupants enter and leave the home – effectively increases infiltration rates. No data is provided to describe how often windows and doors are left open in the Pecan Street homes, so the default assumptions from the PNNL prototypes are used.
- Occupancy schedule – Home occupancy was estimated as described above.
- Building orientation – The orientation of each home was reported for all homes as one of the 16 cardinal, intercardinal, or secondary-intercardinal directions.
- Number of residents – When number of residents was not reported for an individual home, it was replaced with the average value from the remaining homes.
- Heating and cooling setpoints – Heating and cooling setpoints were reported in one of Pecan Street's annual surveys. Residents reported their heating and cooling season thermostat setpoints for their sleeping, morning, workday, and evening hours without reporting actual hours. Because explicit hours for these periods were not identified, we assign sleeping hours as midnight to 7 AM, morning hours as 7 AM to 9 AM, workday hours as 9 AM to 6 PM, and evening hours as 6 PM to midnight. When setpoints were not reported, we assign these values the average of the reporting homes. Homes without programmable thermostats were assigned the same value for all periods.
- Window area per wall – Window area for each external wall, by orientation, was reported in Pecan Street audit records. When window area was not reported, values were assigned as the average of all reporting homes.
- Window type – Windows were described in Pecan Street energy audit records as a combination of frame material and number of panes. EnergyPlus describes windows by their overall U-value, which is primarily a function of these two factors. Overall U-factor for each window type was determined according to [28]. When window type was not reported, values were assigned as the average of all reporting homes.

The final input required to simulate the Pecan Street homes in EnergyPlus is weather data. To ensure that simulated conditions match actual conditions for the monitored period as closely as possible, recorded weather data for Austin in 2015 is taken from [29].

To measure the degree to which the EnergyPlus simulations match monitored data, accuracy will be reported in terms of the hourly coefficient of variation of the root mean square error (CVRMSE). This metric is used in ASHRAE Guideline 14 to define a model as calibrated if its hourly CVRMSE is less than 30% [30].

2.6. Modeling assumptions and limitations

In addition to the assumptions described above, we assume that any changes made by homeowners based on recommendations from the energy audits are minor and do not significantly affect 2015 energy consumption. We also rely on the accuracy of these energy audit records to describe the physical characteristics of the home and its appliance stock. Similarly, we assume the results of the homeowner surveys provide an accurate representation of the thermostat setpoints and number of occupants in each home.

Surveys offer little additional information on the behavior of the simulated homes' occupants. Occupant behavior determines how often lighting, HVAC, and other loads operate, and is therefore well-known to be a significant determinant of a building's energy consumption [31,32]. Additional data gathering would obviously improve simulation accuracy, as numerous important characteristics of the modeled homes and their occupants' behavior are missing. In using the PNNL prototypes, we assume that these missing fields are relatively accurately represented by the default values and assumptions used in these models.

2.7. Sensitivity analysis

Once the prototype home models have been modified with the actual monitored homes' characteristics and occupant behaviors, sensitivity analysis will be conducted to determine which factors have the greatest impact on the accuracy of the simulations. This is done by specifying all fields but one with the actual homes' reported characteristics. The remaining field is assigned the default value used in the PNNL prototype models. The effect of this modification is measured as the change in CVRMSE between the fully specified baseline case and the less specified sensitivity analysis case.

$$\Delta CVRMSE = CVRMSE_{SensAnalysis} - CVRMSE_{Baseline} \quad (1)$$

For sensitivity analysis of floor area, condensing unit efficiency, building infiltration, heating and cooling setpoints, home orientation, ceiling height, window type, window area, attic insulation R-value, occupancy pattern, and number of residents, the actual values are simply reset to the default values used in the PNNL prototype models. In the baseline simulations, AC condenser capacities were specified with the actual homes' nameplate capacities. For the sensitivity analysis, we allow EnergyPlus to autosize the condenser capacity. We test sensitivity to correcting nameplate condensing unit efficiency for its age by simply using the nameplate efficiency, without the correction factor described in Section 2.5. For lighting, we change the modeled homes' lighting power density from the default value in the IECC 2006 prototype to the updated IECC 2012 value, thereby greatly increasing the efficiency of each home's lighting array. Finally, we vary the exterior construction material from its default stucco construction to standard brick masonry construction.

3. Results

3.1. Baseline model accuracy – whole-home consumption

Fig. 2 shows a scatter plot with the EnergyPlus simulated annual energy consumption of the modeled homes on the x-axis and measured annual energy consumption on the y-axis. The black diagonal shows where simulated energy consumption exactly matches actual consumption. The two red lines show ±50% relative annual error bounds.

EnergyPlus simulations are seen to generally overestimate energy consumption in the simulated homes. Of the 106 homes simulated, 46 homes saw their annual energy consumption overestimated by over 50%, 45 homes were overestimated by 0–50%, and only 15 homes were underestimated. Average monitored whole-home consumption is around 10,700 kWh/yr, while average EnergyPlus simulated consumption is around 14,400 kWh/yr. Nearly all of the >50% overestimates occurred in homes whose actual annual consumption is less than 10,000 kWh/yr, indicating that EnergyPlus simulations are less accurate in accounting for the physical characteristics and behaviors seen in these exceptionally efficient homes.

3.2. Baseline model accuracy – condensing unit consumption

Fig. 3 shows the simulated and actual annual energy consumption of these condensing units. Of the 106 homes simulated, only

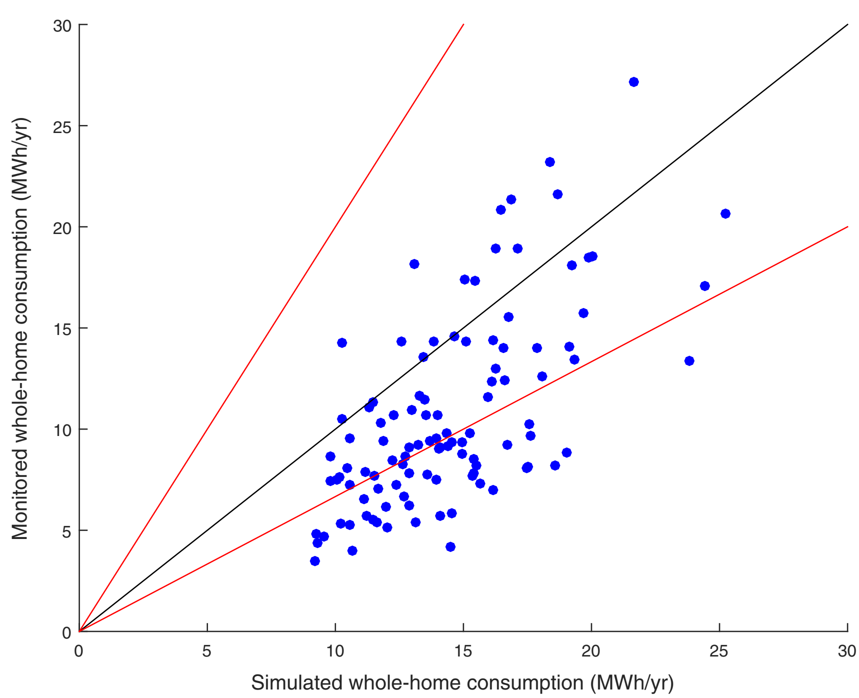


Fig. 2. Scatter plot showing simulated whole-home energy consumption on the x-axis and actual measured energy consumption on the y-axis. The black diagonal shows perfect agreement, while red diagonals show ±50% error bounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

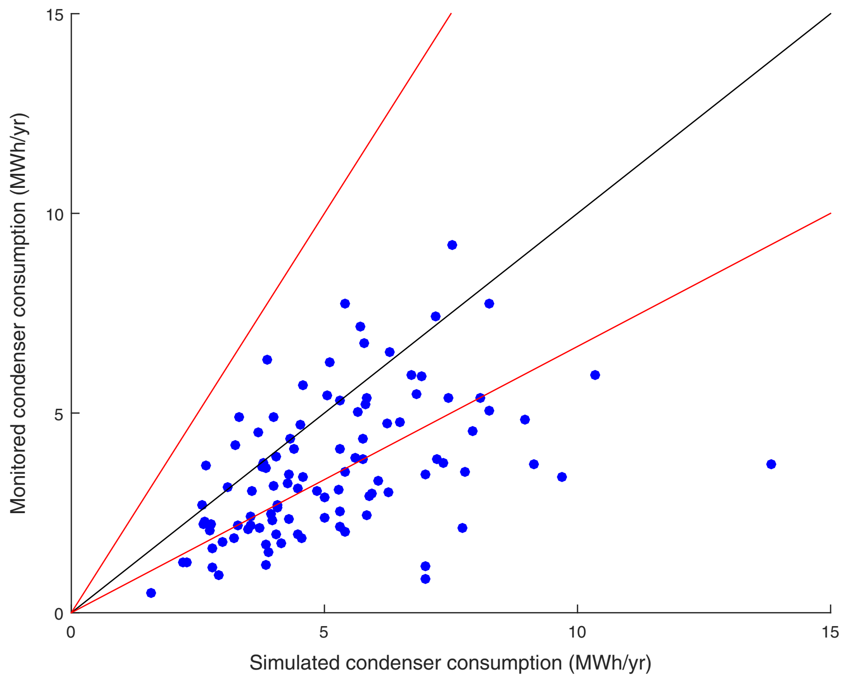


Fig. 3. Scatter plot showing simulated condensing unit energy consumption on the x-axis and monitored condensing unit consumption on the y-axis. The black diagonal shows perfect agreement, while red diagonals show ±50% error bounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

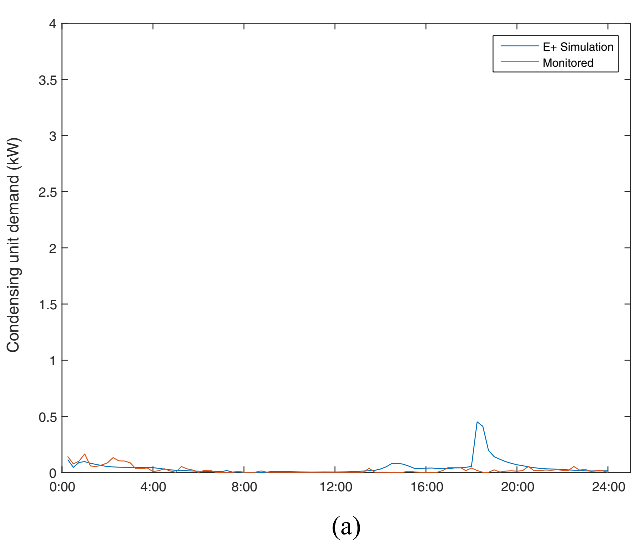
102 had monitored data available for their condensing unit, so only this smaller sample is presented.

This figure shows annual condensing unit energy consumption being overestimated by >50% in nearly half of simulated homes. Average simulated condenser consumption is around 1,500 kWh/yr more than monitored condenser consumption, which explains around 40% of whole-home simulation error. To determine the source of this error, 15-min monitored and simulated interval data profiles were plotted for homes where condensing unit consumption was overestimated by >50% in EnergyPlus.

Fig. 4 shows the average monitored and simulated condensing unit demand for the 52 homes in which EnergyPlus simulated energy consumption was >50% more than actual monitored consumption. The left figure shows average simulated and monitored demand profiles for a day with negligible heating or cooling energy. This shows EnergyPlus correctly identifying no significant cooling load and very closely matching the monitored condenser use on this day on average across all 52 homes. The right figure shows a peak cooling day. This shows the EnergyPlus simulation overestimating both the hours where cooling is required and the condensing unit demand during cooling hours. Interval data profiles for homes with <50% error and underestimated condensing unit consumption can be found in Appendix B.

3.3. Baseline model accuracy – central air fan consumption

Fig. 5 shows the simulated and actual annual energy consumption of central air supply fans. Of the 106 homes simulated, only 75



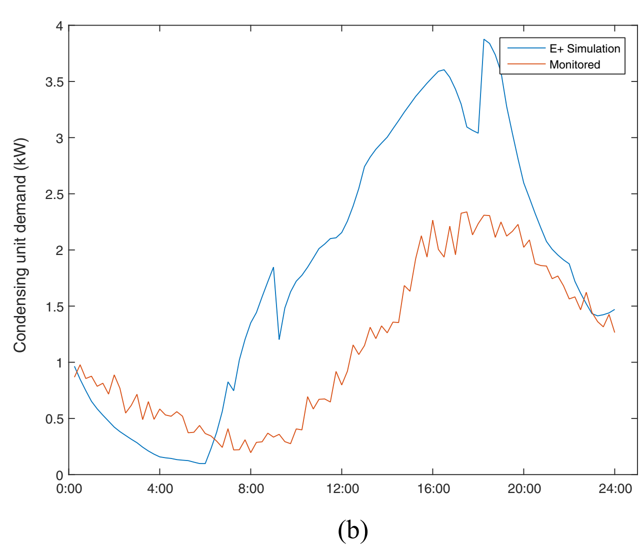


Fig. 4. Average condensing unit demand for homes where EnergyPlus overestimated annual condensing unit energy consumption by >50%. (a) Shows average condensing unit demand for a day with negligible heating or cooling energy. (b) Shows average condensing unit demand for a peak cooling day.

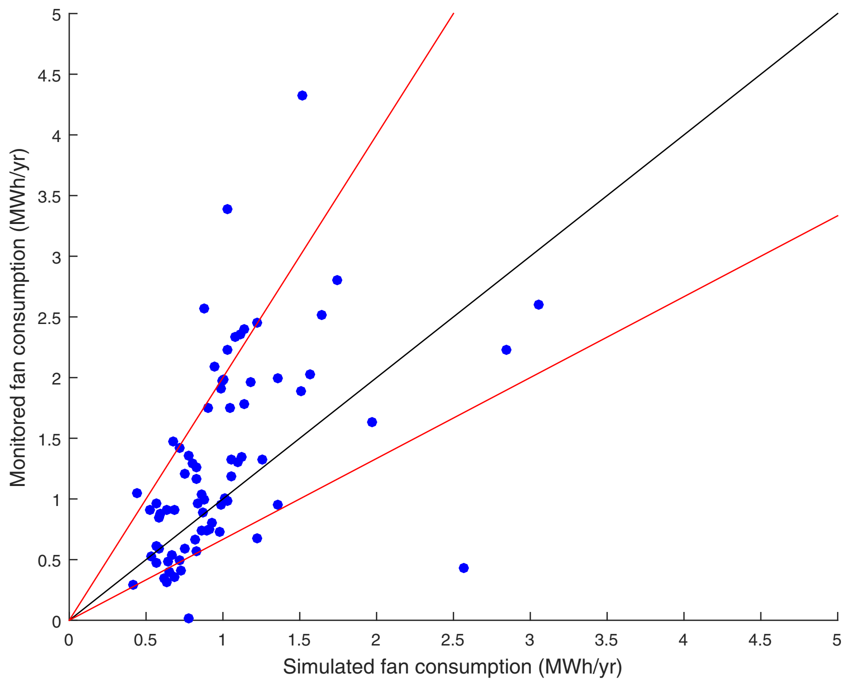


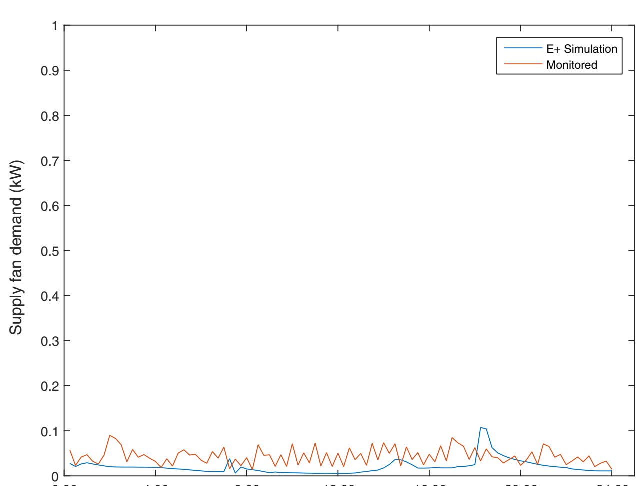
Fig. 5. Scatter plot showing simulated central air supply fan energy consumption on the x-axis and monitored consumption on the y-axis. The black diagonal shows perfect agreement, while red diagonals show ±50% error bounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

had monitored data available for their supply fan, so we only present this smaller sample.

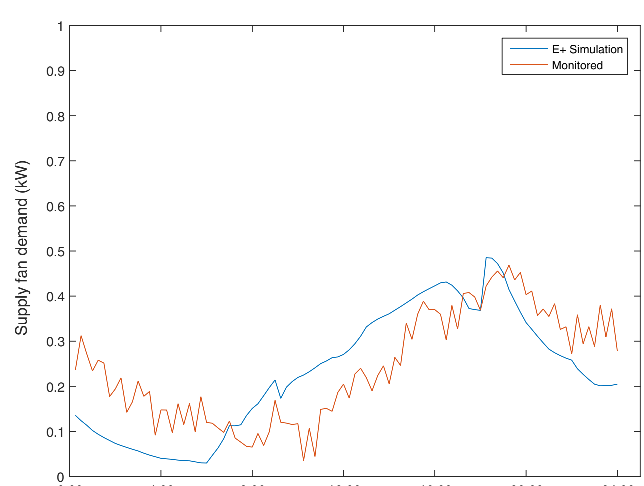
Central air supply fan energy is more accurately simulated than condensing unit energy. Only 19 of the 75 homes included are outside of the 50% error bounds, with average annual consumption around 1,300 kWh/yr in the monitored sample and around 1,000 kWh/yr in EnergyPlus simulations. In general, fan energy consumption is underestimated for homes with high fan consumption and overestimated for homes with low fan consumption. This indicates that the factors which describe the high variance in actual consumption have not been accounted for in the models.

Fig. 6 shows the average monitored and simulated central air supply fan demand for the 21 homes in which EnergyPlus simulated energy consumption was from 0% to 50% more than actual monitored consumption. The left figure shows average simulated and monitored demand profiles for a day with negligible heating or cooling energy. This shows EnergyPlus correctly identifying no significant fan load and very closely matching the monitored fan use on this day in all homes. The right figure shows a peak cooling day. This shows EnergyPlus simulates fan demand earlier in the day than in the monitored homes. Peak fan loads roughly coincide both in timing and in actual kW. Interval data profiles for homes with >50% error and underestimated fan consumption can be found in Appendix C.

Finally, Fig. 7 shows annual energy consumption errors of condensing units on the x-axis and supply fans on the y-axis.



(a)



(b)

Fig. 6. Average central air supply fan demand for homes where EnergyPlus overestimated annual central air supply fan energy consumption by <50%. (a) Shows average demand for a day with negligible heating or cooling energy. (b) Shows average demand for a peak cooling day.

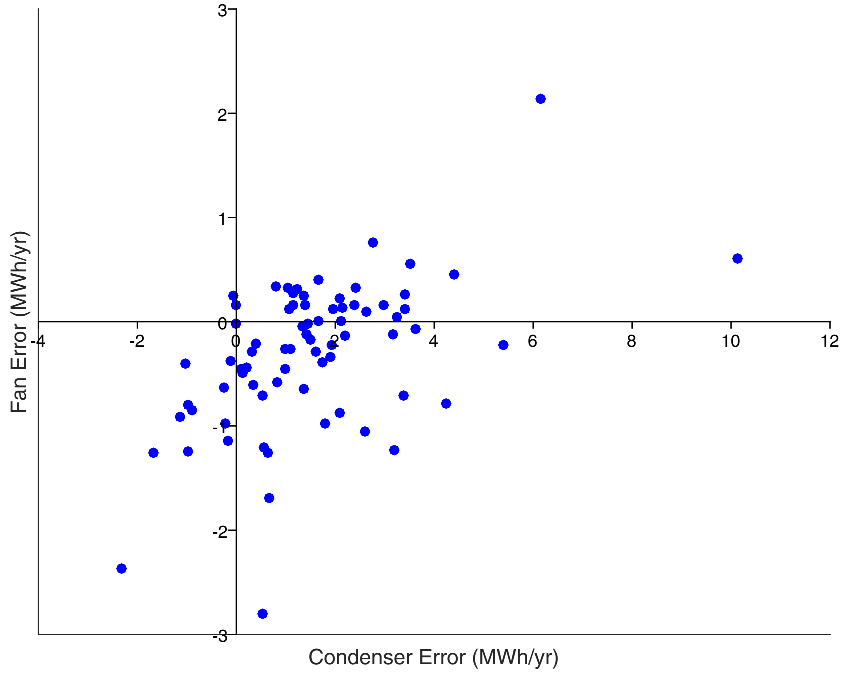


Fig. 7. Scatter plot of EnergyPlus simulation condenser errors on the x-axis and fan errors on the y-axis.

This shows that fan and condenser errors are positively correlated (r = 0.52), indicating that fan and condenser energy errors generally track together, and neither is compensating for the other.

3.4. Baseline model accuracy – non-HVAC energy use

Fig. 8 shows the simulated and actual annual energy consumption of all other end uses in the simulated homes. This includes interior and exterior lighting, a refrigerator, miscellaneous plug loads, kitchen appliances, a dishwasher, and washer and dryer. To properly account for these loads in the monitored data, we require monitored data for the whole home, the condensing unit, and the central air fan. This leaves 74 homes of the original 106 with the necessary data for this comparison.

EnergyPlus simulations again overestimate energy consumption for these loads in nearly all included homes. Average simulated consumption is around 8,600 kWh/yr, and monitored consumption around 5,900 kWh/yr, meaning around 73% of the whole-home error is due to these other loads.

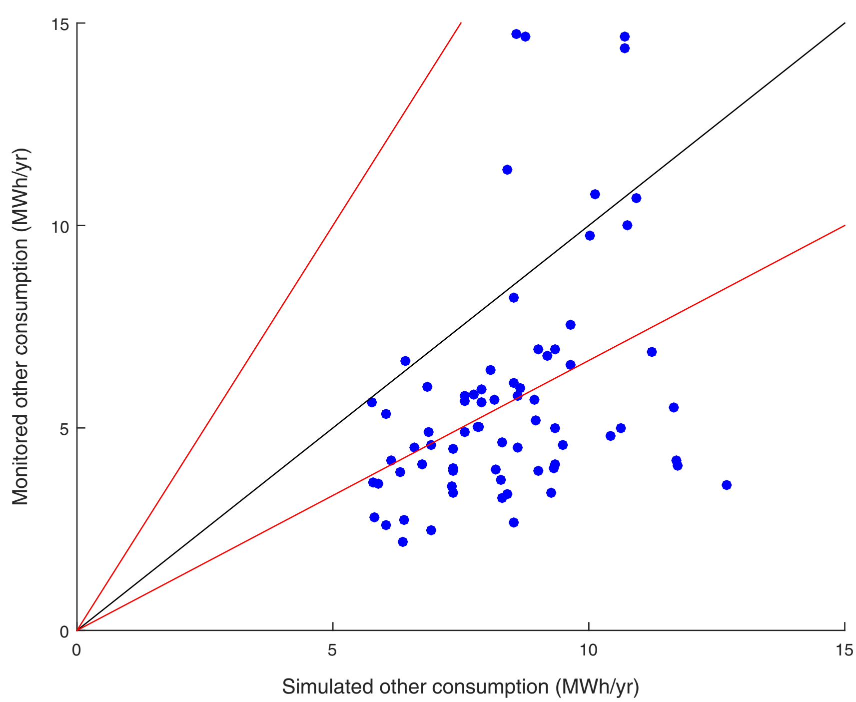


Fig. 8. Scatter plot showing simulated non-HVAC energy consumption on the x-axis and measured consumption on the y-axis. The black diagonal shows perfect agreement, while red diagonals show ±50% error bounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This other energy use also contributes to the overestimated condensing unit energy consumption discussed above. Extra internal energy use increases the homes' cooling loads during Austin's long cooling season. Because most homes are heated by gas and Austin's heating season is short, the corresponding reduction in heating energy use is minimal.

3.5. Sensitivity analysis

Fig. 9 below shows the change in CVRMSE of whole-home energy consumption when individual factors are changed from their actual values as described in Section 2.7. Positive values indicate that CVRMSE increased when a home characteristic was replaced with a default value.

These results show that the CVRMSE between EnergyPlus simulations and actual consumption at the whole-home level is most sensitive to home square footage, condensing unit age and nameplate efficiency, building shell infiltration, heating and cooling setpoints, and window area. When these factors are changed in EnergyPlus from their actual value back to the PNNL prototype default value, CVRMSE changes by upwards of 40% in some homes.

Results are slightly less sensitive to correcting for attic insulation R-value, which change CVRMSE by over 30% in some cases. Building orientation, ceiling height, window frame material and number of panes, lighting power density, and exterior construction material change CVRMSE by over 10% each. Finally, changes in occupancy schedule and number of residents result in a change in CVRMSE of less than 10% in all cases.

In these figures, a negative change in the error term – indicating that model accuracy improved when actual values were replaced with defaults – does not necessarily mean the default value is more accurate, or EnergyPlus is handling these values incorrectly. Instead, it reflects the fact that the baseline simulations do not match metered consumption. Most homes' total energy consumption is overestimated, so any time these homes have an actual value replaced with a default that makes the EnergyPlus model more efficient, the error term will decrease.

In the sensitivity analysis shown in Fig. 9(o), EnergyPlus's autosize method was used to estimate the cooling capacity of each home's condensing unit. This provides a convenient and important measure of a simulated home's estimated cooling load based on building characteristics, occupant behaviors, and internal loads. Fig. 10 shows a scatterplot of these autosized condenser capacities on the x-axis and the nameplate capacities of the condensers in the actual homes on the y-axis.

This shows the autosized condenser capacities are nearly all between 50% and 100% of the installed condenser capacities. This could indicate that the EnergyPlus calculated cooling load is lower than the actual cooling load, or the actual condensers could be oversized by design. Note that installed condenser capacities

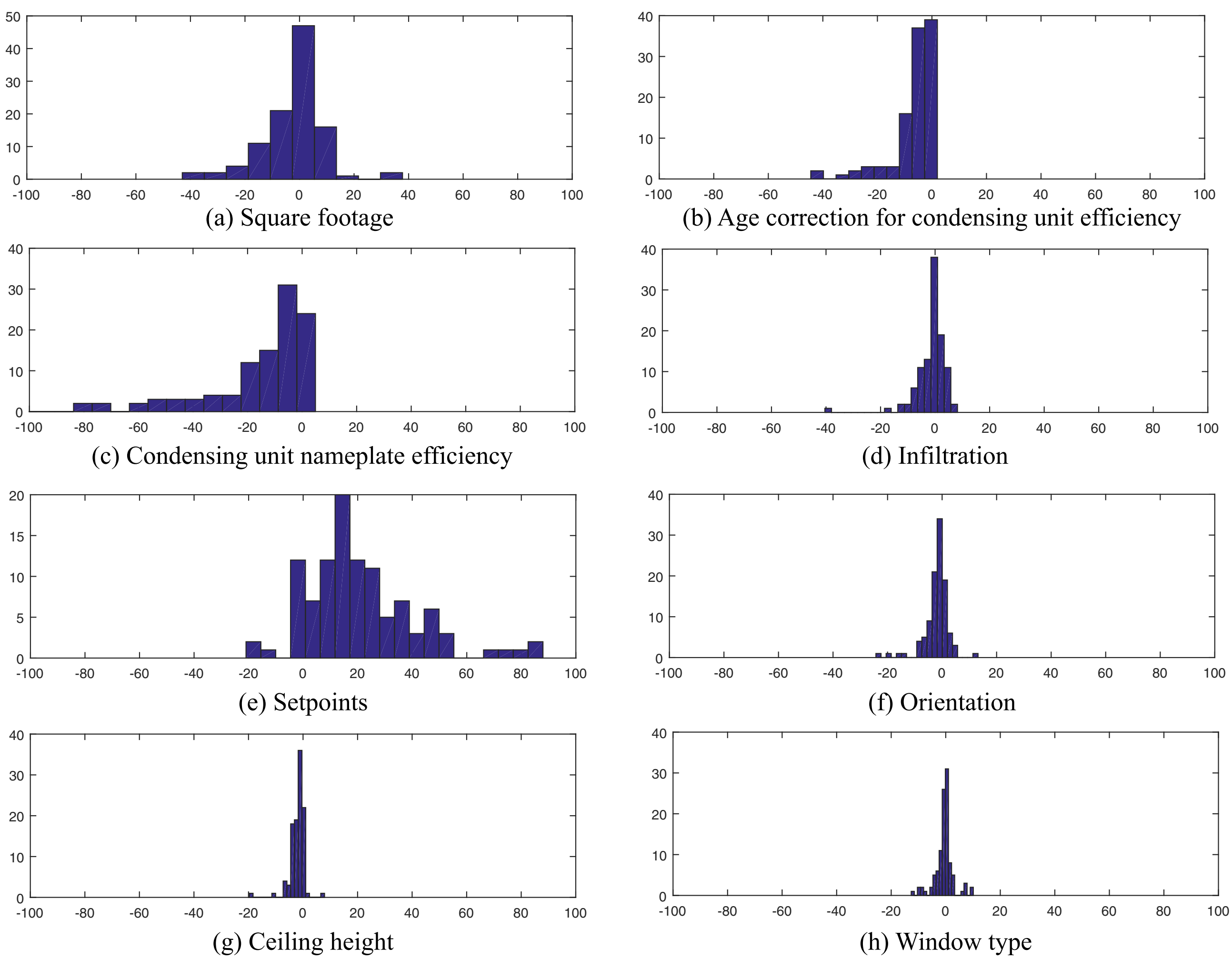


Fig. 9. Histograms showing change in whole-home energy consumption CVRMSE resulting from a change in the factors listed below each figure. The x-axis shows the change in CVRMSE, and the y-axis shows the number of homes in each bin.

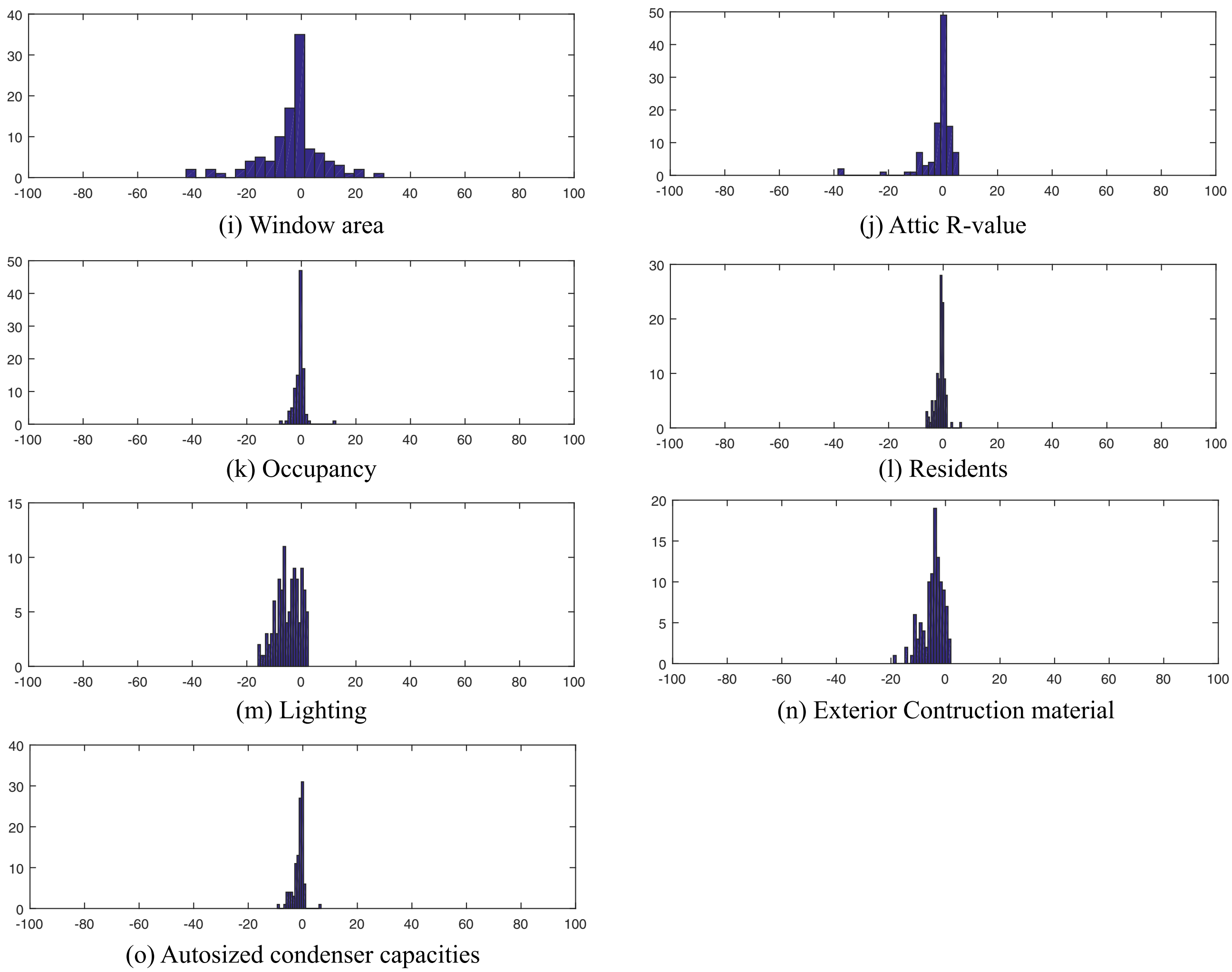


Fig. 9 (*continued*)

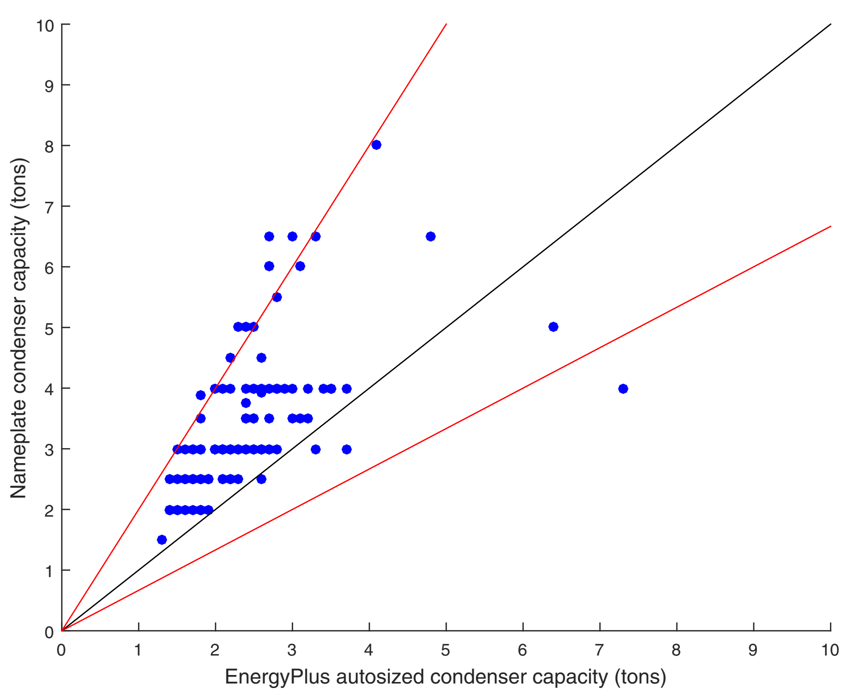


Fig. 10. Scatterplot of autosized condenser capacities on the x-axis and nameplate capacities on the y-axis. The black diagonal shows perfect agreement, while red diagonals show ±50% error bounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Summary of model accuracies in terms of relative annual errors and hourly CVRMSEs.

	Relative error (annual)			CVRMSE (hourly)		
	Min (%)	Mean (%)	Max (%)	Min (%)	Mean (%)	Max (%)
Whole-home	<1	56	240	41	104	360
AC condensing unit	<1	76	730	62	206	1400
Central air supply fan	<1	97	4100	58	207	5700
Other	2	74	250	44	113	270

follow manufacturer's nominal sizes that typically specify condensers in half-ton increments, while EnergyPlus autosized capacities can be assigned any value. If autosized capacities are rounded up to the nearest half-ton increment, they more closely match nameplate capacities. See Appendix B for this figure.

4. Conclusions

A summary of simulation accuracy results, in annual relative error and hourly CVRMSE terms, is shown in Table 3.

These results show that EnergyPlus simulations of single-family homes, as has been reported in commercial buildings, do not consistently or accurately predict actual energy consumption at either the whole-home or device level when specified as described above. As with any model, the quality of inputs determines the quality of the result and more comprehensive energy audit records would allow for more accurate modeling of all systems.

Despite the considerable extent to which the PNNL prototypes were modified with characteristics of the actual monitored homes, whole-home annual energy consumption was consistently overestimated, in many instances by more than 100%. When accuracy is measured by annual relative error, some models appear to be well calibrated as shown in Table 3. But when those same models' accuracies are measured by hourly CVRMSE, none of the models meet ASHRAE's tolerance of 30% CVRMSE. Much of this error can be attributed to the fact that the Pecan Street homes consume far less electricity than average, and the PNNL prototypes were intended to model average code-compliant homes. Any field that was not included in the energy audit records was not changed from the PNNL default, so the models simulated here still have many characteristics in common with the prototypes.

At the device level, condensing unit energy consumption was generally overestimated. Central air supply fan energy is fairly accurately simulated, with the remainder of the whole-home overestimate coming from other end uses, including lighting and all non-HVAC loads. To better understand the source of these errors and how various home and occupant characteristics affect them, the sensitivity analysis conducted here identifies the factors that are most crucial to developing accurate models in the future.

Finally, the modeling of these homes demonstrates the difficulty of generating accurate simulations, even when provided with considerable building and occupant characteristic data. The relative inaccuracy of the models developed here goes to show that many determinants of home energy consumption are not captured during a traditional energy audit and survey, and many more appliance stock, appliance use, and occupant behavioral characteristics are needed to generate accurate residential building simulations.

5. Policy discussion and recommendations

These results provide additional context for the growing use of EnergyPlus in single-family homes. Results here, and previous research in the commercial sector, show that simulations do not accurately estimate actual energy consumption in occupied buildings. Whether these discrepancies between measured energy consumption and simulation results are due to imperfect input data or EnergyPlus algorithms cannot be determined here. Additional research using complete, ground-truth input data and detailed measured end-use data is needed to better understand the sources of these errors. What can be said is that there are too many variables affecting energy use in occupied homes that cannot be accurately included in building simulations. Simulation tools do likely provide a reasonable estimate of as-built building performance under default operational settings, device stocks, and occupancy and behavioral assumptions. But these tools should not be used to estimate or predict actual occupied building energy consumption.

The DOE Building Energy Codes Program should consider the inaccuracies seen here and in previous research as their work continues to use EnergyPlus as a tool for evaluating future energy codes. The current method of simulating incremental changes to building codes and estimating energy savings and lifecycle costs can be a valuable tool. But the fact that simulations typically do not accurately predict actual energy consumption once homes are occupied means that these simulations should not be used to predict actual realized energy consumption or savings in future homes.

Finally, details of RESNET's EnergyPlus-based compliance tool have not yet been released, but it can be assumed that it will likely simulate a designed homes' performance over a year, and compare that to some baseline code-compliant version of the same home. This would reflect a major transition to a systems-level approach to code compliance, as any whole-building simulation model would consider interactions between the building envelope, internal loads, and the heating and cooling systems. If this is the case, a set of assumptions and standard conditions will have to be established that fairly value the future occupants' levels of efficiency, but that also limit the effects of model inaccuracies. The results of our sensitivity analysis can serve as a guide for establishing these baseline states. The most sensitive factors need to be considered most carefully, as default values and assumptions do not accurately represent actual homes. Further, the device-level measures of the accuracy of the 106 models of Pecan Street homes highlight the fact that far more information than was described here is necessary to build accurate models. Whole-building simulation can be a powerful decision-making tool, but care is needed to ensure that decision-makers are aware of their limitations, and not let the relative ease of simulating building energy performance get ahead of the capabilities of the tool.

Acknowledgments

This work is supported by the Center for Climate and Energy Decision Making (CEDM), through a cooperative agreement between Carnegie Mellon University and the National Science Foundation [grant number SES-1463492].

Appendix A. Pecan Street survey, audit, and intervention summary

Table A.1 shows the participation in Pecan Street programs of the 106 homes included in final simulations. These programs are detailed below:

Intervention 1 indicates that the home is located in Texas but is outside of the Greater Austin area.

Intervention 2 was a study in which energy use in each home in the study was monitored for one year and then summarized. At the end of that year, this summary was provided to the homeowner and another year of data was monitored to determine if energy use changed.

Table A.1
Summary of simulated homes' participation in Pecan Street programs.

Intervention/program	Homes included
Intervention 1	0
Intervention 2	81
Intervention 3	19
Intervention 4	9
Intervention 5	27
Intervention 6	15
Intervention 7	18
Intervention 8	106
Intervention 9	7
Intervention 10	0

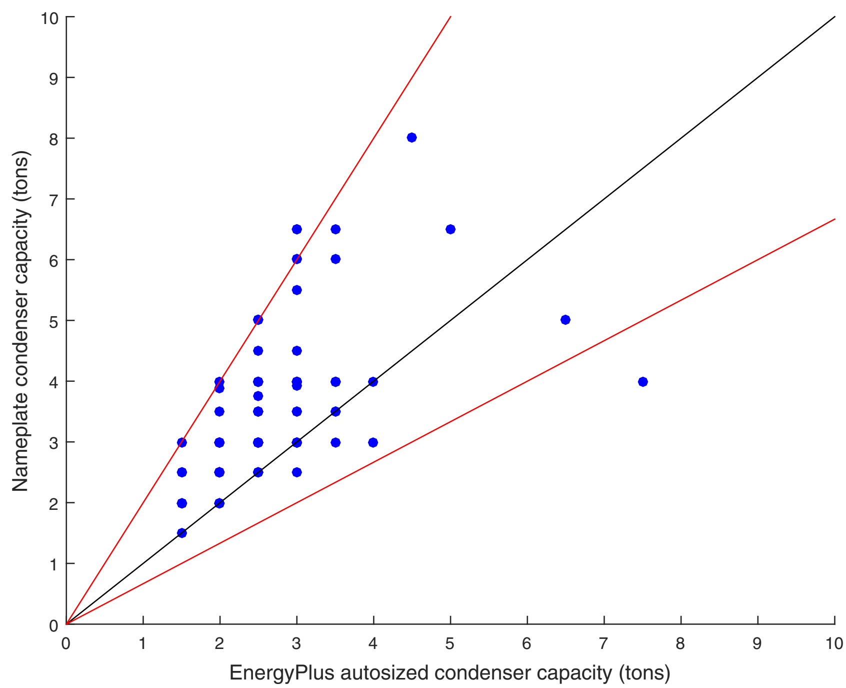


Fig. B.1. Scatterplot of autosized condenser capacities rounded up to the nearest half-ton on the x-axis and nameplate capacities on the y-axis. The black diagonal shows perfect agreement, while red diagonals show ±50% error bounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

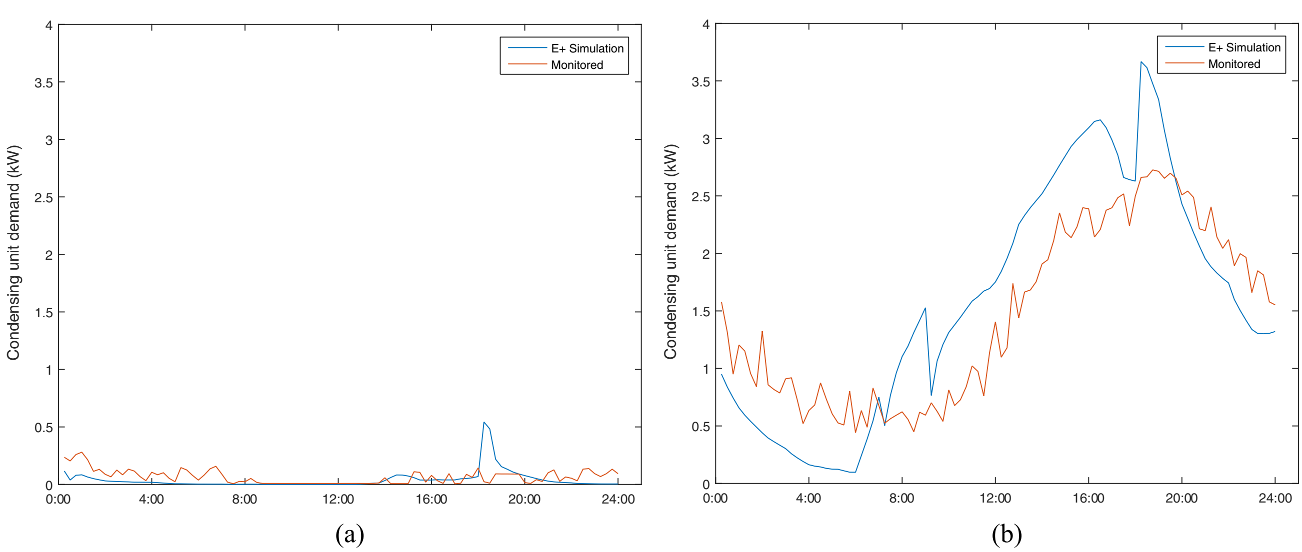


Fig. B.2. Average condensing unit demand for homes where EnergyPlus overestimated annual condensing unit energy consumption by <50%. (a) Shows average condensing unit demand for a day with negligible heating or cooling energy. (b) Shows average condensing unit demand for a peak cooling day.

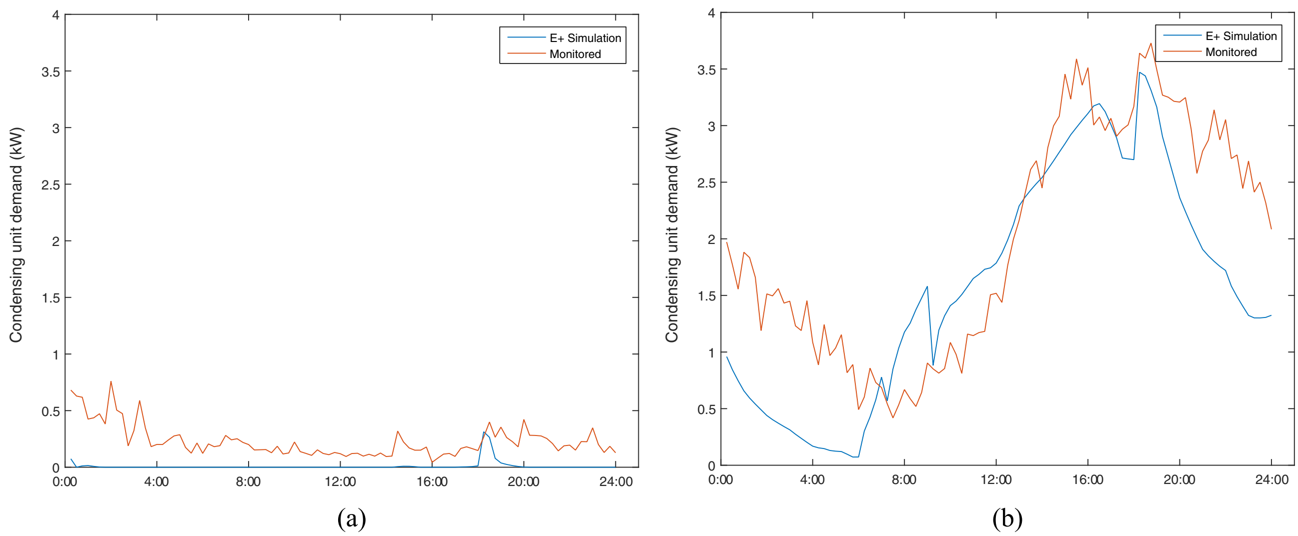


Fig. B.3. Average condensing unit demand for homes where EnergyPlus underestimated annual condensing unit energy consumption. (a) Shows average condensing unit demand for a day with negligible heating or cooling energy. (b) Shows average condensing unit demand for a peak cooling day.

Intervention 3 indicates the home was part of a control group for the CCET trial. This was a study which made home energy use data available to homeowners via an online portal and established an experimental time-of-use pricing scheme.

Intervention 4 indicates that homeowners only had access to the CCET home energy reporting portal. While these participants were made aware of how the experimental pricing scheme would affect them, they did not receive any actual financial incentive.

Intervention 5 indicates that homeowners both had access to the CCET online energy portal and received financial incentives in accordance with the experimental pricing scheme established to reduce peak demand.

Intervention 6 indicates that homeowners in the CCET trial received text messages asking them to reduce their energy consumption on peak days.

Intervention 7 was similar to Intervention 6, but text messages provided information on which appliances should be curtailed.

Intervention 8 indicates participation in the energy internet demonstration program. Most of these homes are in Austin's Mueller neighborhood.

Intervention 9 was a program that delivered new LG washers, dryers, and some refrigerators to participating homeowners.

Intervention 10 is a program that gives tablets to residents of low-income apartment complexes to provide access to their online energy portal.

Appendix B. Additional condensing unit results

See Figs. B.1–B.3.

Appendix C. Additional central air supply fan results

See Figs. C.1–C.3.

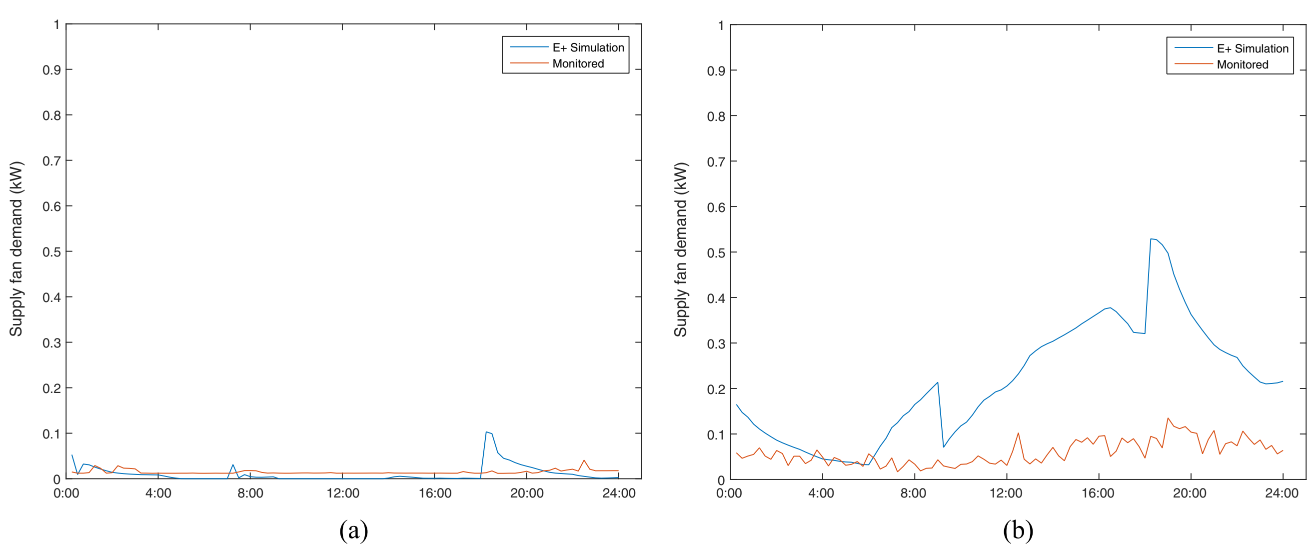


Fig. C.1. Average central air supply fan demand for homes where EnergyPlus overestimated annual central air supply fan energy consumption by >50%. (a) Shows average demand for a day with negligible heating or cooling energy. (b) Shows average demand for a peak cooling day.

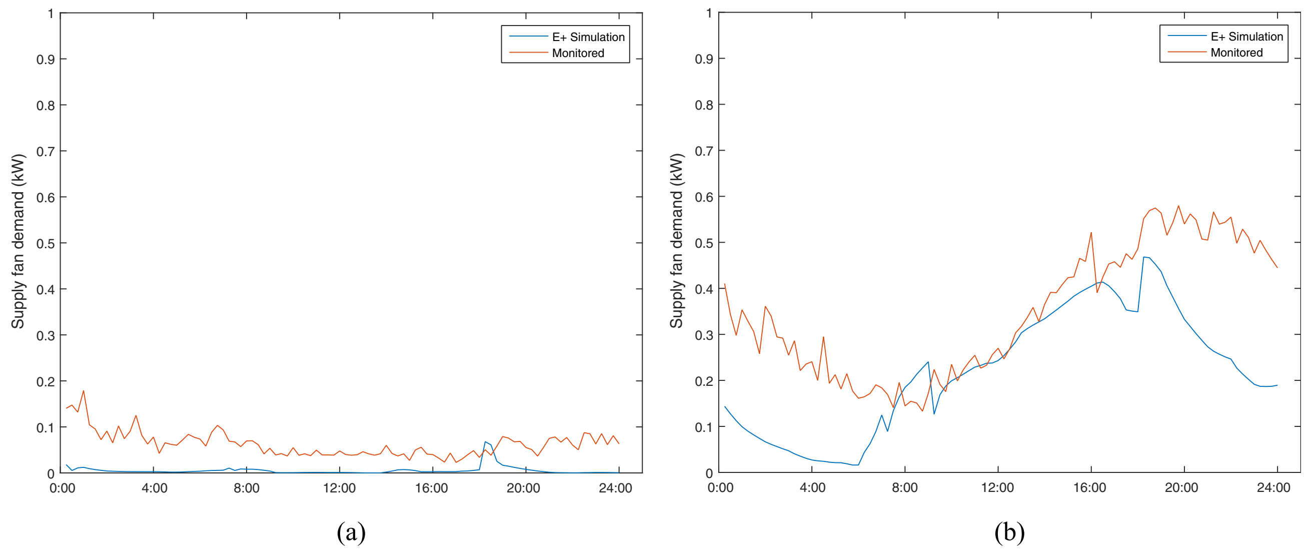


Fig. C.2. Average central air supply fan demand for homes where EnergyPlus underestimated annual central air supply fan energy consumption by <50%. (a) Shows average demand for a day with negligible heating or cooling energy. (b) Shows average demand for a peak cooling day.

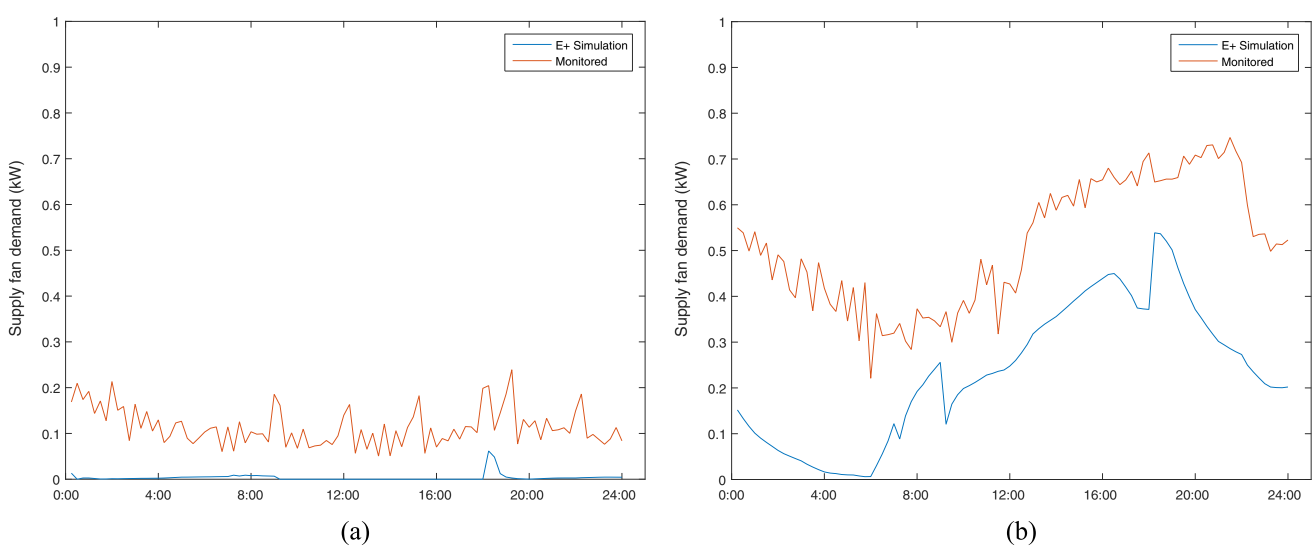


Fig. C.3. Average central air supply fan demand for homes where EnergyPlus underestimated annual central air supply fan energy consumption by >50%. (a) Shows average demand for a day with negligible heating or cooling energy. (b) Shows average demand for a peak cooling day.

References

[1] Augenbroe G. Trends in building simulation. Build Environ 2002;37:891–902. http://dx.doi.org/10.1016/S0360-1323(02)00041-0.

[2] Crawley DB, Lawrie LK, Winkelmann FC, Buhl WF, Huang YJ, Pedersen CO, et al. EnergyPlus: creating a new-generation building energy simulation program. Energy Build 2001;33:319–31. http://dx.doi.org/10.1016/S0378-7788(00)00114-6.

[3] Taylor T, Fernandez N, Lucas R. Methodology for evaluating cost-effectiveness of residential energy code changes; 2012.

[4] Turner C, Frankel M. Energy performance of LEED for new construction buildings. WA: Vancouver; 2008.

[5] DOE Office of Energy Efficiency and Renewable Energy. EnergyPlus is going residential; 2016. <http://energy.gov/eere/buildings/articles/energyplus-going-residential> [accessed September 27, 2016].

[6] Li N, Yang Z, Becerik-Gerber B, Tang C, Chen N. Why is the reliability of building simulation limited as a tool for evaluating energy conservation measures? Appl Energy 2015;159:196–205. http://dx.doi.org/10.1016/j.apenergy.2015.09.001.

[7] Mustafaraj G, Marini D, Costa A, Keane M. Model calibration for building energy efficiency simulation. Appl Energy 2014;130:72–85. http://dx.doi.org/10.1016/j.apenergy.2014.05.019.

[8] Yang T, Pan Y, Mao J, Wang Y, Huang Z. An automated optimization method for calibrating building energy simulation models with measured data: orientation and a case study. Appl Energy 2016;179:1220–31. http://dx.doi.org/10.1016/j.apenergy.2016.07.084.

[9] Yang Z, Becerik-Gerber B. A model calibration framework for simultaneous multi-level building energy simulation. Appl Energy 2015;149:415–31. http://dx.doi.org/10.1016/j.apenergy.2015.03.048.

[10] Pan Y, Huang Z, Wu G. Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai. Energy Build 2007;39:651–7. http://dx.doi.org/10.1016/j.enbuild.2006.09.013.

[11] Royapoor M, Roskilly T. Building model calibration using energy and environmental data. Energy Build 2015;94:109–20. http://dx.doi.org/10.1016/j.enbuild.2015.02.050.

[12] Karlsson F, Rohdin P, Persson M-L. Measured and predicted energy demand of a low energy building: important aspects when using building energy simulation. Build Serv Eng Res Technol 2007;28:223–35. http://dx.doi.org/10.1177/0143624407077393.

[13] Loutzenhiser PG, Maxwell GM, Manz H. An empirical validation of the daylighting algorithms and associated interactions in building energy simulation programs using various shading devices and windows. Energy 2007;32:1855–70. http://dx.doi.org/10.1016/j.energy.2007.02.005.

[14] RESNET Residential Energy Services Network; 2016. <https://www.resnet.us/> [accessed September 27, 2016].

[15] US Department of Energy. States that can use REScheck to show compliance; 2015. <https://www.energycodes.gov/images/states-can-use-rescheck-show-compliance> [accessed September 29, 2016].
[16] RESNET. National registry of accredited rating software programs; 2016. <http://www.resnet.us/professional/programs/energy_rating_software> [accessed September 29, 2016].
[17] Lin HW, Hong T. On variations of space-heating energy use in office buildings. Appl Energy 2013;111:515–28. http://dx.doi.org/10.1016/j.apenergy.2013.05.040.
[18] Coakley, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renew Sustain Energy Rev 2014;37:123–41. http://dx.doi.org/10.1016/j.rser.2014.05.007.
[19] Sun K, Hong T, Taylor-Lange SC, Piette MA. A pattern-based automated approach to building energy model calibration. Appl Energy 2016;165:214–24. http://dx.doi.org/10.1016/j.apenergy.2015.12.026.
[20] Rhodes JD, Gorman WH, Upshaw CR, Webber ME. Using BEopt (EnergyPlus) with energy audits and surveys to predict actual residential energy usage. Energy Build 2015;86:808–16. http://dx.doi.org/10.1016/j.enbuild.2014.10.076.
[21] DOE Building Energy Codes Program. Residential prototype building models; n.d. <https://www.energycodes.gov/development/residential/iecc_models> [accessed October 11, 2016].
[22] DOE. Residential prescriptive requirements: 2009 International Energy Conservation Code (IECC); n.d. <https://energycode.pnl.gov/EnergyCodeReqs/index.jsp?state=Texas> [accessed October 11, 2016].
[23] Pecan Street Dataport 2016. <https://dataport.pecanstreet.org>.
[24] US Energy Information Administration. Residential energy consumption survey 2009. Washington, DC; 2012.
[25] Austin Energy. Corporate reports and data library – energy use and sales; 2016. <http://austinenergy.com>.
[26] Hendron R. Building America performance analysis procedures for existing homes; 2006.
[27] Sherman MH, Dickerhoff DJ. Airtightness of U.S. dwellings. ASHRAE Trans 1998;104:1359–67.
[28] DOE Office of Energy Efficiency and Renewable Energy. Selecting windows for energy efficiency new 2013. http://dx.doi.org/10.1093/rheumatology/kep450.
[29] NOAA. Quality controlled local climatological data; 2016. <https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/quality-controlled-local-climatological-data-qclcd> [accessed January 1, 2016].
[30] ASHRAE. ASHRAE guideline 14-2002, measurement of energy and demand savings. Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers; 2002.
[31] Hong T, Lin H-W. Occupant behaviour: impact on energy use of private offices. Berkeley, CA; 2013.
[32] Brandemuehl MJ, Field KM. Effects of variations of occupant behavior on residential building net zero energy performance. Proc Build Simul 2011.