



Cluster analysis of simulated energy use for LEED certified U.S. office buildings



Mohammad Heidarinejad^{a,*}, Matthew Dahlhausen^a, Sean McMahon^b,
Chris Pyke^b, Jelena Srebric^a

^a Department of Mechanical Engineering, University of Maryland, College Park, MD, USA

^b The U.S. Green Building Council (USGBC), Washington, DC, USA

ARTICLE INFO

Article history:

Received 23 July 2014

Received in revised form 19 August 2014

Accepted 5 September 2014

Available online 16 September 2014

Keywords:

Building classification

Office buildings

LEED certification

Energy simulations and modeling

Building-site energy utilization index

Internally-load dominated

ABSTRACT

This study uses cluster analysis to examine simulated energy consumption of 134 U.S. LEED NC office buildings to classify buildings into high, medium, and low energy use intensity clusters. The analysis uses energy simulation results from the LEED database, as a comparably large data set of energy end uses from sub-meter data does not yet exist. The difference between the low energy use cluster and other clusters is explained mostly by lower process loads and lower heating energy intensity, and partly by lower intensities of other HVAC related end-uses. Lighting energy use shows the least variation between clusters. The lower heating energy intensity in the low intensity cluster is largely explained by lower roof U-values, lower window-to-wall ratio, and smaller building size. Unregulated process loads are the most significant contributor to total building energy use, accounting for 36%, 33%, and 31% of the energy use in the high, medium, and low intensity clusters, respectively. This analysis provides a quantitative evaluation of the large difference in energy intensities in high-performance office buildings, showing that these buildings are dominated by internal loads, especially unregulated process loads.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Office buildings are a common space type in building classification systems, including Leadership in Energy & Environmental Design (LEED), Commercial Building Energy Consumption Survey (CBECS), and Energy Star Portfolio Manager [1–3]. Offices account for 17% of the total floor area and 19% of primary energy consumption for commercial buildings in the U.S. [4]. Office buildings in aggregate are typically the most energy intensive class of buildings in developed countries [5]. Consequently, the literature recommends to consider office buildings as one of the primary space types due to (1) prevalence of office buildings compared to other commercial building types, (2) increase in the floor area of office buildings in the recent years, and (3) existence of unique energy use pattern of HVAC, lighting, and plug load end-uses [5]. In order to determine energy end-use composition, opportunities to improve energy simulation accuracy, and characteristics that yield low energy use for office buildings, this study analyzes the energy simulation results for 134 U.S. office buildings that attained Leadership

in Energy & Environmental Design for New Construction & Major Renovations (LEED NC) certification, representing new or major-renovated high-performance buildings.

This study does not report metered energy data because such a large and detailed dataset still does not exist from building sub-metering. Prior comparisons of metered energy data to energy simulation results for LEED buildings have found that metered data may vary from the energy simulation results [6–8], hence the present study does not aim to make claims about energy use in particular buildings, but rather to focus on energy use characteristics of the sector overall. Considering the simulation models are accurate, though imprecise measure of average energy use consumption [7], this study assumes the models to be a valid means of providing insights into energy use for a group of buildings in aggregate. The major benefit of using simulated LEED NC office buildings is to quantify the contributions of each energy end-use, including interior lighting, plug load, cooling, heating, and fans, to the total building energy consumption. This technique is used in the absence of sub-metered interval data for a large building set. Sub-metering at one office building requires expensive installation of end-use sensors [9–11], costing around \$100,000 for a medium sized office building [10], and still contains significant uncertainties due to sensor failures and other errors in sensor accuracy

* Corresponding author. Tel.: +1 301 405 6247; fax: +1 301 314 9477.
E-mail address: muh182@umd.edu (M. Heidarinejad).

or installation [12]. Large datasets of sub-metered energy use do not exist due to the costs and effort associated with data collection, such as normalizing for operation, maintenance, and controls, as well as remaining significant uncertainties in comparing data sets from different investigators. Therefore, in absence of such a dataset, energy simulations can provide necessary data to disaggregate energy end-uses and evaluate performance due to inherent building characteristics.

Use of building energy simulations allows assessment of building energy performance throughout different phases in the building lifecycle, including building design, operation, and retrofit. Two types of commonly used simulation studies are (1) case studies of individual buildings comparing simulation to metered performance and (2) reference building models that model typical performance for different building types and climate zones. The U.S. Department of Energy (DOE)'s Commercial Reference Building Models are examples of energy simulation models used extensively in the building industry and scientific publications to benchmark and analyze building performance [13]. For example, to evaluate energy and Indoor Air Quality (IAQ) performance of the buildings, the DOE Reference Building Models are usually one of the commonly used models to draw conclusions on (1) sensitivity analysis of the input and output variables for the new construction building energy models [14], (2) effectiveness of the Energy Efficiency Measures (EEMs) for the building retrofits [15,16], (3) benchmark building energy consumptions [17], (4) improvement of infiltration rates for the building energy models [18], (5) assessment of IAQ [19], (6) optimal cost of the building operation [20], (7) model-predictive control of the building operation [21], (8) assessment of the weather data on the building energy models [22], (9) evaluation of building integrated photovoltaic window with semi-transparent solar cells [23], and (10) quantification of energy performance for the existing buildings [24]. This indicates the state-of-the-art in using building energy simulation models with typical inputs applicable to a wide range of applications in the building industry to make conclusions on a broad range of research areas. Therefore, this study aims to extend the use of a large scale building energy simulation results to develop clusters of buildings for the LEED NC certified U.S. office buildings based on the energy end-use type. These clusters can help building benchmarking efforts by comparing an individual building to the high performance buildings designed by LEED NC guidelines and requirements.

Cluster analysis determines groups of buildings in the building stock based on a few energy use variables. Specifically, supervised clustering is a suitable technique to determine groupings of buildings by energy use. Clustering techniques alone are limited in their usefulness when applied to building energy use patterns with a large number of potential variables. Clusters rely on minimization of distances between groups of variables, so a large number of variables can result in sphericity, an instance where the cumulative variances in multi-dimensional datasets fully fill lower-order dimensional space, making clusters hard to distinguish in a two- or three-dimensional visualization. Recently, cluster analyses are used as a useful statistical tool to analyze building energy consumption to identify (i) two clusters of buildings based on EUI (Energy Use Intensity) for office buildings in China using the agglomerative hierarchical algorithm [25], (ii) five clusters based on the heating energy consumptions in the school building sector in Greece using the k-means clustering algorithm [26], and (iii) eight clusters of multifamily buildings based on the building fuel type, age, and size in NY City Benchmarking program using the model-based clustering algorithm [27]. This study uses cluster analysis to classify LEED NC office buildings based on energy use intensity using the k-means clustering algorithm.

2. Research methodology

LEED NC versions 2 (V2) and 3 (V3) award credits for buildings that can demonstrate a simulated energy use reduction for a proposed building as compared to a baseline building of the same size that meets code minimum requirements according to ASHRAE Standard 90.1 or Title 24 California for buildings in California [28–31]. The proposed building must have the same amount of unregulated energy use, energy not subject to code requirements, as the baseline building. The unregulated energy use in both the baseline and proposed energy models must account for at least 25% of the annual building energy cost (not energy use), though there are exemption opportunities if savings can be justified in the model. None of the studied office buildings used this exemption, meaning plug loads accounted for at least 25% of energy costs in the models, and there were no plug load reductions in the proposed compared to baseline models.

This study used proposed energy simulations for 134 U.S. office buildings that attained LEED NC certification from the U.S. Green Building Council (USGBC). These buildings cover 13 climate zones in the U.S., and vary in size from 185 m² (~1,991 ft²) to 18,580 m² (~199,999 ft²). The data analyses use the results of energy simulations from 40 projects submitted under LEED NC 2.2, represented with “V2” and 94 projects submitted under LEED NC 2009, represented with “V3”, resulting in the total of 134 office building projects. The majority of the projects comprise LEED certified office buildings from 2008–2011. Accessibility to the LEED Energy and Atmosphere (EA) forms and availability of the submitted information for the building energy end-uses and building envelope properties are one of the primary reasons to select these representative buildings. This study selects majority of the accessible LEED V3 office buildings till 2012, and selects the LEED V2 projects with a random selection out of the entire LEED V2 office buildings. These energy simulations constitute the office buildings sampled from several hundred projects. These projects use different energy simulation software approved for use by LEED [32], unique energy modelers, and cover a variety of building sizes and locations. Energy simulation results are submitted to USGBC and reviewed before the credits are approved [32,33].

2.1. Building mean weighted and GFA mean weighted calculation

The clustering uses total site energy use intensity to group buildings into clusters with similar properties. Cluster results are given in terms of medians, building-weighted means defined in Eq. (1), and GFA-weighted means defined in Eq. (2) as follows:

$$\text{Building – weighted mean EUI} = \frac{1}{N} \sum_{i=1}^N \frac{E_i}{A_i} \quad (1)$$

$$\text{GFA – weighted Mean EUI} = \frac{\sum_{i=1}^N E_i}{\sum_{i=1}^N A_i} \quad (2)$$

where N is the total number of buildings, E_i is the building energy use, A_i is the building area, and EUI is the resulting mean Energy Use Intensity. Calculating the standard error for GFA-weighted EUI requires a Taylor expansion of the variance of E/A given by the following equation:

$$\begin{aligned} \text{GFA – weighted mean EUI standard error} &= \frac{\sqrt{\text{var}\left(\frac{E}{A}\right)}}{\sqrt{N}} \\ &= \frac{\sqrt{\frac{\text{var}(E)}{A^2} - \frac{2\bar{E}}{A^3} \text{cov}(E, A) + \frac{\bar{E}^2}{A^4} \text{var}(A)}}{\sqrt{N}} \end{aligned} \quad (3)$$

Table 1
Six variable categories selected from LEED NC certification documents.

(1) Project criteria	(2) Area	(3) Climate	(4) Building envelope	(5) Occupancy	(6) Energy end-use
LEEDLv1 BuildType	GFA CFA	Zone AvgDB AvgDWP WINDSPD HDD65 CDD50	WallU RoofU WinU WinSHGC WWR	AvgFTE AvgTrans TotOcc	Heat Cool FanPump ExtLight IntLight SHW Process

Median property values represent the most typical building in the cluster. The building-weighted means show the most common properties for buildings in this cluster. The GFA-weighted means show the overall importance a property weighted by the largest buildings in that cluster. As an example, the GFA-weighted mean heating EUI in the high energy use cluster is significantly larger than that of the median due to a few large buildings with high heating load. This difference shows that a few buildings are responsible for the GFA-weighted results, whereas if the building-weighted mean and GFA-weighted mean are similar, it suggests greater similarity between the buildings within a cluster.

2.2. Framework

This study developed a framework to analyze building energy end use clusters for the portfolio of buildings with detailed inputs. For the cluster analysis, variables are selected and categorized based on their usefulness [34]. The building data are collected from the submitted design documents for the LEED certification, and the weather data used the typical meteorological year (TMY3) from the weather station closest to the project location [35].

2.2.1. Selected variables

Project variables are grouped in six categories: (1) Project information, (2) Area, (3) Climate variables, (4) Building envelope, (5) Occupancy rates [33,36], and (6) Energy end use values. Renewable energy production is not included in the analysis because it does not contribute to energy end use. Mechanical equipment efficiencies are available in some projects, but calculating overall equipment efficiency is not possible across different equipment types with unknown run-times, so they are not included in the analysis. Plug and occupancy density and schedules are not available on the reporting documents, but are set to ASHRAE 90.1 standards, meaning they do not contribute to improvement in energy performance demonstrated by the model, and are theoretically equivalent across the buildings. However, authors did notice a variation in plug load density, caused partly by variations in sub-space classification as a portion of building area and inclusions of multiple types of process loads. Table 1 summarizes the following project variables:

1. Project criteria include LEED certification level (LEEDLv1) and building type (BuildType) according to LEED V3 building type classification. Please, note that this analysis only considers office buildings.
2. Area variables include Gross Floor Area (GFA) and Conditioned Floor Area (CFA).
3. Climate variables include American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) climate zones (Zone), average annual dry bulb temperature (AvgDB), average annual dew point temperature (AvgDWP), average annual wind speed (WINDSPD), heating degree-days referenced to 18.3 °C (65 °F) (HDD65), and cooling degree-days referenced to 10 °C (50 °F) (CDD50). Climate variables use typical meteorological year data (TMY3) from the weather station closest to the project location [35].

4. Building envelope variables include wall assembly U-value (WallU), roof assembly U-value (RoofU), window U-value (WinU), window solar heat gain coefficient (WinSHGC), and window-to-wall ratio (WWR). If certification documents list more than one U-value for an assembly, an area-weighted average provides the overall U-value.
5. Water use or occupant use documentation provides occupancy rates including average full time employees (AvgFTE), average visitors and transients (AvgTrans), and total occupancy (TotOcc) as the summation of both employees and transients.
6. Energy simulation results provide annual energy use, broken down by energy end-use and summarized into seven major categories. Heating (Heat) includes space heating provided by natural gas, electric or other means. Cooling (Cool) includes space cooling and heat rejection. Fans and pumps (FanPump) include interior fans, pumps, and HVAC controls. Heating, cooling, and fan and pumps constitute all Heating, Ventilation and Air-Conditioning (HVAC) related energy end uses. Exterior lighting (ExtLight), Interior lighting (IntLight) and Service Hot Water heating (SHW) are regulated loads, kept distinct. Process loads (Process) are unregulated energy use including process or industrial equipment, receptacle equipment, data servers, refrigeration, cooking, individual task or process lighting and other uncategorized loads. This analysis does not consider fans used in parking garages, which are only present in four of the office buildings. All lighting, hot water, and process loads constitute non-HVAC related loads (non-HVAC).

2.2.2. Outlier omission

Several office buildings contain extremely high process energy consumption for dedicated data centers, laboratory space, or industrial manufacturing included in these office buildings. These space uses do not constitute typical office buildings. These outliers have a total EUI, calculated as the total energy use over GFA, which exceeds the upper limit of the data, $Q3 + 1.5 \times IQR$, or are below the lower limit of the original data, $Q1 - 1.5 \times IQR$ for the original dataset. Here, IQR, Q3, and Q1 are the interquartile range, and the medians of the upper and lower half of total EUI, as suggested in the literature [8]. Finally, out of 134 LEED NC office buildings, 11 office buildings are excluded based on the results of this data cleaning process. Fig. 1 shows the median and standard deviation of the EUI with and without outliers. Interestingly, there are no outliers on the low end of the data, which would require a building to have total site intensity less than 47.7 kWh/m² (15.1 kBtu/ft²). A possible explanation is that for the lowest energy use buildings, as the energy use in the building decreases, further energy savings becomes significantly more difficult and expensive to a point that renewable energy is preferred to further load reduction. Renewable energy is available for projects, but was not included in the overall building site EUI in order to avoid comparing high-performing buildings to low-performing buildings with significant renewable energy production. All upper-end outliers had very high process loads, such as server equipment, which contributed to high HVAC energy use. This suggests that accurate accounting for unregulated process loads is critical to energy simulation accuracy.

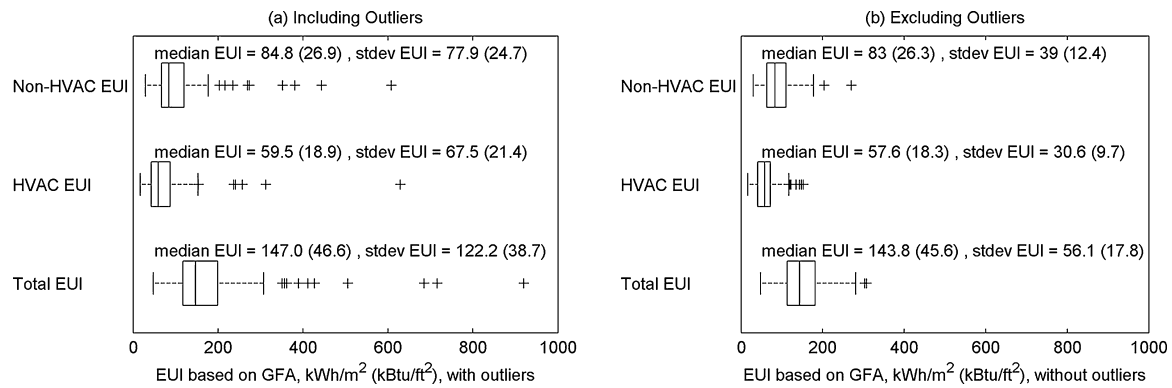


Fig. 1. Total, HVAC, and non-HVAC EUIs; (a) including outliers and (b) excluding outliers. Crosses indicate the data that are greater than the upper limit of the 3rd quartile plus $1.5 \times$ the interquartile range. (Note: EUI is calculated as the annual sum of all energy end-uses, excluding parking fans, divided by the GFA of the buildings.)

Table 2 provides a summary of the analyzed LEED office buildings with and without outliers. After omission of outlier buildings, this study selected 123 buildings to analyze from the originally selected 134 office buildings. Omitting outliers significantly reduces the standard deviation, standard error of the EUI, and deviation of median from mean. For example, in Table 2, for the GFA-weighted mean EUI, the standard deviation reduces from 194.0 kWh/m² (61.5 kBtu/ft²) to 81.8 kWh/m² (25.9 kBtu/ft²) after the omission. Table 3 provides a summary for the analyzed LEED NC office buildings based on the certification level. For the selected LEED NC certified office buildings, Silver certification has the most contribution, and the Platinum certified office buildings have the lower median EUI compared to the other certification levels.

2.3. Cluster analysis

Cluster analysis is a method to find structures in the data by identifying natural groups, clusters, in datasets. Although there is a wide range of algorithms for the clustering analysis, including partitioning, hierarchical, model-, density-, and grid-based methods, typically, there is no measure of accuracy to select the clustering algorithm. Usefulness of the clusters is the only measure to

interpret the results with the post-classification processing [37]. For the building energy consumptions, the k-means that is the commonly used partitioning algorithm, agglomerative hierarchical, and model-based algorithms are used to determine number of clusters for different datasets of energy consumptions [25–27,38]. For the selected database of LEED NC office buildings, k-means cluster analysis is successfully used to analyze the energy end-use type of the LEED NC office buildings to provide insight into the LEED buildings by showing the differences in total energy use for three buildings clusters. Instead of analyzing a building dataset by breaking it into energy use classes by total energy use intensity so that there are an equal number of buildings in each class, or organizing along quartiles, clustering can determine groups that have a fuzzy boundary to distinguish between clusters [38]. The advantage of this approach is that the building differentiation by total energy use intensity maximizes difference between clusters to allow for a meaningful comparison of the differences between building clusters.

3. Clustering results

Traditionally, buildings are classified as three different groups of small-, medium-, and large-sized buildings based on their area [13,39]. This study considers the energy use patterns of the buildings to develop three clusters based on the building energy use, enabling to correlate the building energy use with the key influential variables. Interestingly, the clustering results show that the three clusters that emerged based on building energy use patterns are also directly correlated to the building size. The buildings were clustered using Matlab's k-means clustering function with three clusters [40]: a high energy intensity, medium energy intensity, and low energy intensity cluster, by the square Euclidian distance of total building-site energy utilization index on both axes, shown in Fig. 2(a). Three clusters generated meaningful differences in energy consumption with enough nuance to separate clusters of buildings. Fig. 2(b)–(d) shows the distributions of HVAC EUI, non-HVAC EUI, and total EUI for the three clusters. A comparison between clusters show that non-HVAC EUI is more directly correlated with the total

Table 2
Total EUI of the LEED NC office buildings, with and without outliers, expressed in kWh/m² (kBtu/ft²).

Variables	With outliers	Without outliers
Number of buildings	134	123
Mean EUI	181.0 (57.4)	152.1 (48.2)
Standard deviation of mean EUI	122.2 (38.7)	56.1 (17.8)
Standard error of mean EUI	10.6 (3.3)	5.1 (1.6)
Median EUI	147.0 (46.6)	143.8 (45.6)
GFA-weighted mean EUI	197.8 (62.7)	166.5 (52.8)
Standard deviation of GFA-weighted mean EUI	194.0 (61.5)	81.8 (25.9)
Standard error of GFA-weighted mean EUI	16.8 (5.3)	7.4 (2.3)

Table 3
Total EUI of the LEED NC office buildings, by certification level, excluding outliers, expressed in kWh/m² (kBtu/ft²).

Variables	Platinum	Gold	Silver	Certified
Number of buildings	25	34	37	27
Mean EUI	142.7 (45.2)	147.0 (46.6)	157.9 (50.0)	159.1 (50.4)
Standard deviation of mean EUI	68.7 (21.8)	43.0 (13.6)	60.0 (19.0)	53.4 (16.9)
Standard error of mean EUI	13.7 (4.4)	7.4 (2.3)	9.9 (3.1)	10.3 (3.3)
Median EUI	131.3 (41.6)	141.4 (44.8)	145.9 (46.3)	146.9 (46.6)
GFA-weighted mean EUI	174.3 (55.2)	142.9 (45.3)	171.7 (54.4)	177.3 (56.2)
Standard deviation of GFA-weighted mean EUI	78.6 (24.9)	48.0 (15.2)	76.4 (24.2)	89.9 (25.5)
Standard error of GFA-weighted mean EUI	15.7 (5.0)	8.2 (2.6)	12.6 (4.0)	17.3 (5.5)

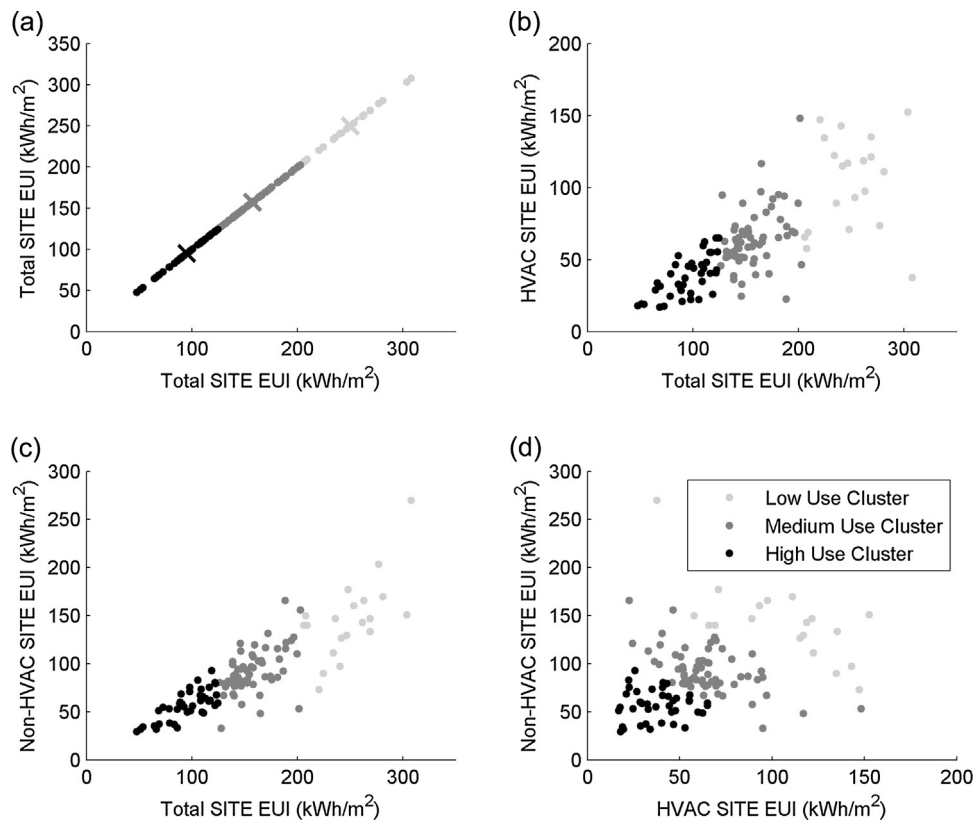


Fig. 2. Cluster results showing High (light gray), Medium (gray), and Low (dark gray) energy use intensity clusters, showing (a) comparison of total site energy, which generates the clusters, and correlation between (b) HVAC related energy uses, (c) Non-HVAC related energy use, and (d) the relation between HVAC and Non-HVAC related energy uses. (Note: The axis for total site EUI, non-HVAC EUI, and HVAC EUI are (0, 350), (0, 300), and (0, 200), respectively.)

EUI, rather than the HVAC EUI. The buildings in the first cluster, shown in light gray color, are distinctly separated from the rest of the buildings, indicating different building characteristics that lead to greater energy use.

Table 4 summarizes building size characteristics for the three clusters. The high and medium intensity clusters have similar, but still distinguishable, sizes within a standard error of each other and a median size of 2,323–2,787 m² (25,000–30,000 ft²). However, buildings in the low intensity cluster are significantly smaller, about half the size of those in the other clusters, with a median size of 1,091 m² (11,744 ft²).

Fig. 3 provides the number and percent of LEED NC office buildings for each cluster based on the certification level. All certification levels are present in all clusters, with the percentage of higher certification levels (Gold and Platinum) increasing with decreasing energy intensity. However, this correlation is not very strong, and the proportion of buildings with a specific certification level does not explicitly follow energy intensity. Gold certified buildings are significantly more prevalent in the medium energy use cluster than in the low energy use cluster. Platinum and Silver buildings are least

presented in the medium use cluster. Little can be concluded about a building's energy use knowing certification level alone.

Fig. 4 shows the breakdown by energy end-use for each cluster. Key findings are that buildings in the low energy use cluster consume roughly 1/3 of the energy as the average building in the high energy use cluster. Process energy use is the most significant energy end-use, followed by heating, then lighting. Cooling energy use intensity in the high use cluster is 3 times greater than that of the low use cluster, fan and pump energy use is 2.5 times greater, and process energy use is 3 times greater. These end-uses have a similar percentage of energy use across all clusters, suggesting that they are codependent, and may serve as key indicators for a building's total energy use.

Process energy use accounts for at least 30% of building energy use across clusters, and up to 40% of primary energy use in the high energy use cluster. Energy simulations for LEED use a default W/ft² or W/m² value for unregulated process loads, accounting for much of the total annual energy use in low intensity building cluster. This result is comparable to the metered energy use in high-performance buildings, where plug loads may account for

Table 4
Building size characteristics of the high, medium, and low energy use clusters.

Cluster	High	Medium	Low
Number of buildings (#)	20	63	40
Percentage of buildings (%)	16.3	51.2	32.5
Percentage of total building area (%)	22.9	56.3	20.8
Percentage of total building energy use (%)	34.4	53.1	12.5
Median size in m ² (ft ²)	2,605 (28,040)	2,369 (25,502)	1,091 (11,744)
Mean size in m ² (ft ²)	4,994 (53,754)	3,891 (41,878)	2,263 (24,355)
Standard deviation of size in m ² (ft ²)	4,892 (52,655)	4,113 (44,272)	3,345 (36,010)
Standard error of size in m ² (ft ²)	1,094 (11,774)	518 (5,578)	529 (5,694)

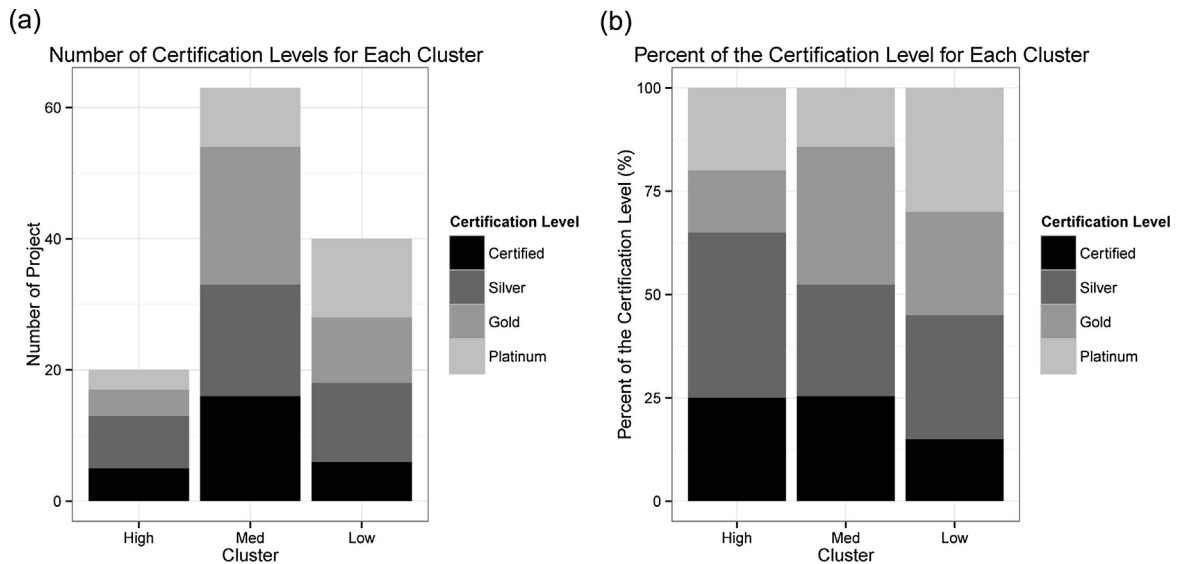


Fig. 3. Distribution of the LEED buildings: (a) number of LEED NC office buildings for each cluster and (b) percent of the LEED NC office buildings for each cluster.

more than 50% of the building energy use [41,42]. Recent research studies showed that using plug load management programs and occupant behavior modifications could lead to significant plug load reduction [41–43]. This suggests that further development of high performance building energy simulations should focus on improving the accuracy of unregulated process and internal lighting load inputs. Furthermore, certification programs, such as LEED, may consider allowing verifiable plug load reduction strategies in energy simulations.

Internal lighting energy use intensity varied the least of all end-uses across the clusters, but increased as a percent of load in the lower intensity building cluster. The energy use intensity of lighting in the low use cluster is three-quarters that of the high energy

use cluster. Lighting goes from around 10% of energy use in the high energy use cluster to over 20% in the low energy use cluster. This result indicates that the marginal improvements in lighting reduction available in these buildings are much smaller than for other end-uses, suggesting that further energy use reductions in best-practice highest performing buildings will need to come from passive lighting strategies. The results of a review on the shading devices used in the projects showed about half of the buildings for each cluster did not utilize any shading devices while about 30% of the buildings in all clusters only used overhangs. Table 5 provides a statistics of the shading devices used in all three clusters. Other shading devices used in these projects include exterior fins, windows film, louver sun screen, cover canopy, balcony, perforated

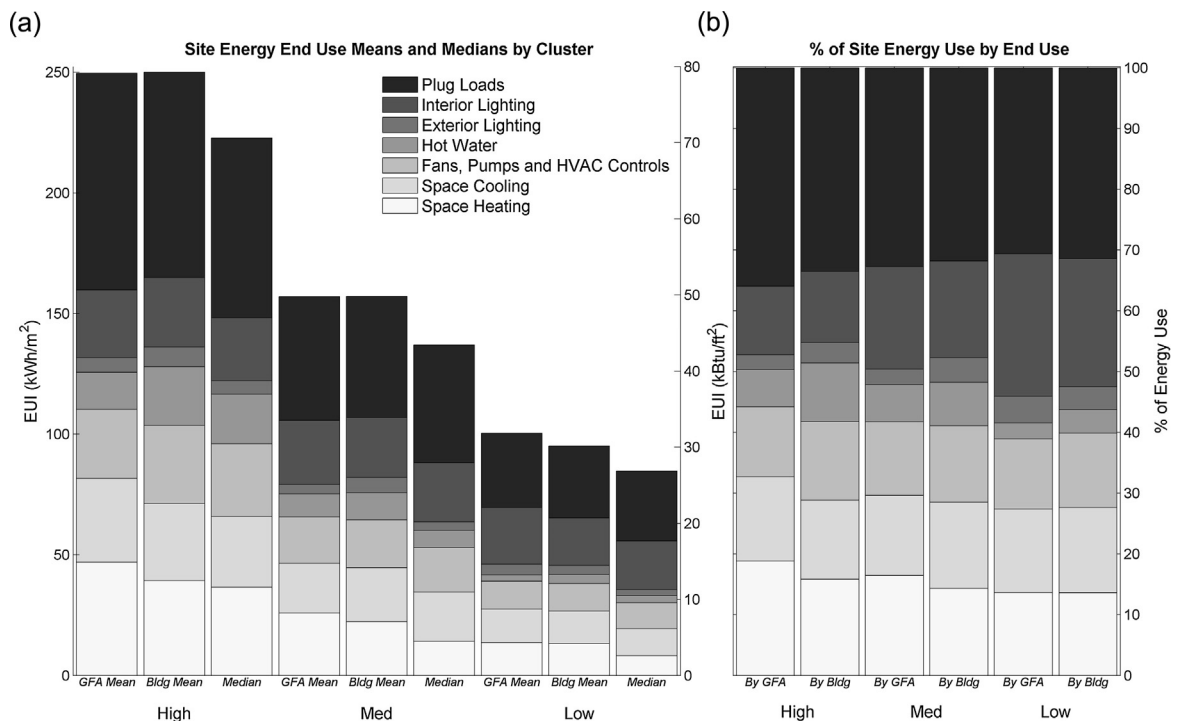


Fig. 4. (a) Site energy end use intensities contributing to total site energy use, by cluster, and shown by GFA-weighted mean, building-weighted mean, and median values, (b) shows site energy end uses as a percent of total site energy use.

Table 5
Percentage of shading devices specified in the EA credit forms for the reviewed office buildings.

High			Med			Low		
None	Overhangs	Other	None	Overhangs	Other	None	Overhangs	Other
45%	30%	25%	54%	34.9%	11.1%	55%	32.5%	12.5%

fixed panels, or combination of these strategies. The lighting control strategies are not directly specified in the submission forms. Therefore, as one of the enhancement in new version of LEED, LEED V4, the new registered projects need to comply with ASHRAE 90.1 2010 instead of ASHRAE 90.1 2007 that requires more advanced lighting control options [44]. This upgrade for the lighting code compliance will increase potential for additional energy savings. Furthermore, the high variability in energy intensity across the various end-uses suggests that building energy efficiency codes that focus on the component level of systems, rather than the building as an integrated system, are not fully capturing best-practice designs and their resulting available energy savings. In particular, unregulated process loads that drive total energy use and much HVAC use as well are not a part of energy codes.

Heating energy use varies the most between clusters, with heating energy use in the high energy cluster four times that of the low energy use cluster, though it remains between 14% and 19% of the total site energy use. One possible explanation is that this reflects the bias in comparing site energy use for buildings with different fuel sources for heating, as the low and medium energy clusters have a much greater percentage of buildings that use electricity for heat, shown in Table 6. Buildings that use on-site gas will show a greater site heating energy intensity than those that run an electric heat pump, because of the coefficient of performance (COP) advantage from the heat pump, typically >2, to the efficiency of gas systems, typically <0.96. Smaller buildings are more likely to rely on a heat pump system because of their low capital cost and ease of installation, thus accounting for some of the difference in heating energy intensity. This suggests that heating energy may

Table 6
Heating characteristics of clusters.

Cluster	High	Medium	Low
Number of buildings	20	63	40
Buildings with only gas heat	3	16	8
Buildings with only electric heat	6	25	24
Buildings with both gas and electric heat	9	19	7
Buildings with other heat	2	2	0
% Cluster heat energy use from gas	65.5	81.1	41.8
% Cluster heat energy use from electricity	30.9	16.3	58.2
% Cluster heat energy use from other	3.6	2.6	0

be more appropriately compared using source EUI rather than site EUI. However, adjusting for source energy conversions also shows similar relative differences in heating energy use intensity between the clusters, meaning that other factors, namely building enclosure components, are responsible for this difference.

Fig. 5 shows U-values and window components for the three building clusters. There is a clear trend towards lower roof U-values in the lower energy use cluster. The window and wall U-values do not show this trend. The lower energy use cluster has a lower window-to-wall ratio than the high and medium clusters. As a result, the area-weighted U-value for the exterior surface is much lower than that of the other clusters. Buildings in the low energy use cluster are smaller and therefore likely to have fewer stories, so the roof surface area is a greater proportion of the overall building area compared to other buildings. SGHC has the least variation among the clusters although the SGHC standard deviation is within the variation of SGHC for different climates in ASHRAE Standard 90.1. These

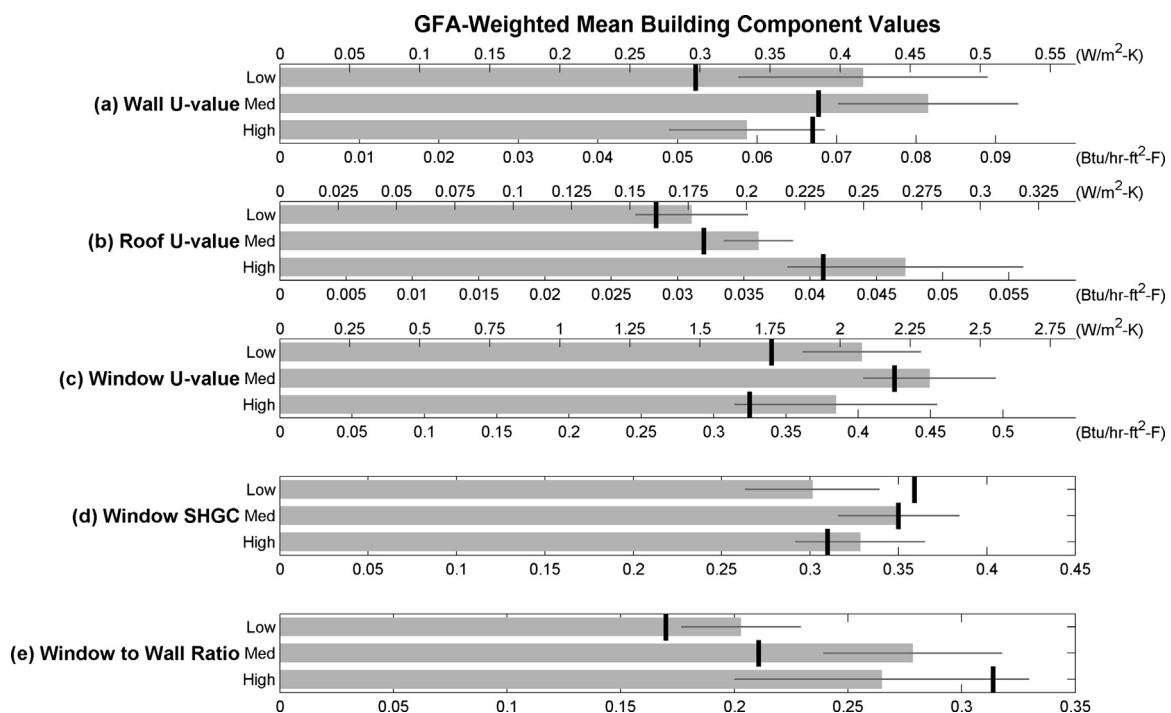


Fig. 5. Building component insulation; (a) Wall U-values, (b) Roof U-values, and (c) Window U-values, and window characteristics (d) Window Solar Heat Gain Coefficient (SHGC) and (e) Window-to-Wall Ratio (WWR) for each cluster. Horizontal gray lines show the 95% confidence interval, and thick vertical black lines show cluster medians.

Table 7
Heating and cooling degrees days for each cluster. Heating degree days are referenced to 18.3 °C (65 °F), and cooling degree days to 10 °C (50 °F) with 95% confidence interval.

Cluster	HDD 18.3 °C (HDD 65 °F)			CDD 10 °C (CDD 50 °F)		
	High	Med	Low	High	Med	Low
Median	2,911 (5,239)	2,250 (4,050)	3,096 (5,572)	2,178 (3,920)	2,289 (4,121)	1,965 (3,537)
GFA-weighted mean	2,962 ± 247 (5,332 ± 444)	2,449 ± 141 (4,407 ± 254)	2,704 ± 179 (4,867 ± 322)	2,361 ± 211 (4,249 ± 379)	2,340 ± 130 (4,211 ± 235)	2,307 ± 135 (4,153 ± 244)

Table 8
Unique features of each cluster.

Cluster	Unique features	
	Building	Energy Use Pattern
Low intensity	Smaller size, lower WWR, lower roof U-values compared to two other clusters.	Higher percent of interior lighting energy end-use type compared to two other clusters.
Medium intensity	Located in milder climates, and higher window and wall U-values compared to the high intensity cluster, with similar building size.	The percent of interior lighting and plug load end-use types are within the range of low and high intensity clusters, and SHW end-use is the only end-use that has relatively higher percent than the low intensity cluster.
High intensity	Lower wall U-value compared to two other clusters.	Higher percent of plug load, lower percent of lighting end-use type compared to two other clusters.

differences explain the majority of the variations between heating loads in the high, medium, and low energy use clusters.

Table 7 shows that the clusters have similar CDD and HDD profiles, with the exception being that the medium energy cluster has significantly fewer HDDs than the other clusters. Buildings located in low and high intensity clusters have similar climate distributions, suggesting the difference in heating energy use intensity between these clusters is not explained by climate differences.

In summary, the clusters have unique features compared to each other resulting in distinct energy use patterns. The largest differences between clusters are in process loads and heating energy use, followed by other HVAC uses, and little difference in lighting energy end-use. The heating energy use difference is explained by the building components insulation values. Low intensity cluster benefited from the lower roof U-values, lower window-to-wall ratio, and smaller size. Buildings in the medium intensity cluster are relatively smaller than the high intensity cluster and are located more in milder climates. Buildings in the high intensity cluster have lower wall U-values compared to buildings in the two other clusters and share a similar climate distribution with the low energy intensity cluster. Table 8 shows the unique features of each cluster, and Table 9 shows the similarities between the clusters. Further case study intervention studies on the low intensity buildings are necessary to provide recommendations for building owners and designers, to explore what is possible in terms of realistic low energy intensity building design for new construction buildings and potential energy retrofit savings for existing buildings.

Table 9
Similarities between clusters.

Cluster	Similarities
Process and lighting	Process and lighting end-use type account for approximately over 50% of the building total energy use.
HVAC	HVAC end uses (heating, cooling, ventilation) share a similar percentage of building energy use across clusters.
Climate	High and low intensity use cluster share similar heating and cooling degree days.
Building envelope	No significant variation exists in SHGC among different clusters.
Certification level	Decreasing cluster energy intensity coincides with greater composition of higher certification levels.

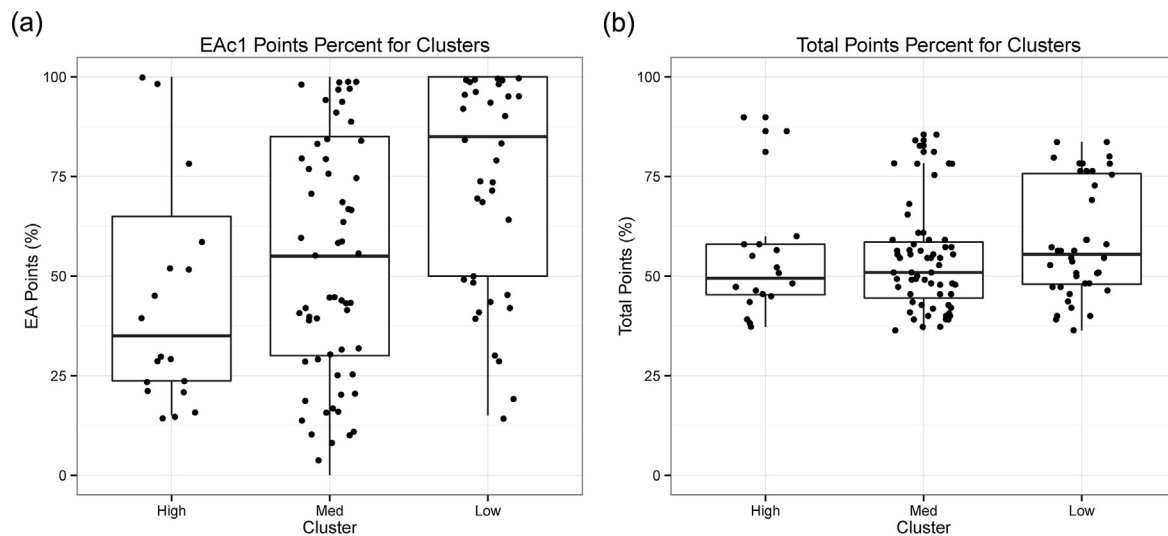


Fig. 6. Distribution of: (a) the EAc1 and (b) total points percent for the high, medium, and low energy intensity clusters.

4. Cluster characteristics

This study examined whether the identified clusters are influenced by the LEED credits, simulation tools, certification level, and year of application. Furthermore, the dataset was used to reduce its dimensionality in order to perform factor analysis. Finally, the identified clusters are compared with the Reference Building energy models.

4.1. Compare the LEED points and certification level

This study compares the Energy and Atmosphere Credit 1 (EAc1): optimize energy performance and total points achieved by the project, certification level, energy simulation tools used for the energy modeling, and year of application for the three identified clusters. Since there are both V2 and V3 projects, percentage of EAc1 points and the total points attained by the projects are normalized with the maximum achievable EAc1 and total points for each version, respectively. The following equations are used to normalize the points:

$$EAc1_p(\%) = \frac{EAc1_{p-project}}{EAc1_{p-max}} \quad (4)$$

where $EAc1_{p-max}$ is 10 for V2 and 19 for V3, and $Total_{p-max}$ is 69 for V2 and 110 for V3.

Fig. 6 shows distribution of the EAc1 and total points for the three identified clusters, indicating the cluster attained a descending distribution of EAc1 and total points. Low intensity cluster has a higher percentage of EAc1 points following medium and high intensity clusters. Compared to the EAc1 point distribution, the total percentage of the points shows similar descending pattern with a lower range of variation. Low intensity cluster has attained the highest points compared to the medium and high intensity clusters; however, there is no difference between the medium and high intensity clusters.

Fig. 7 illustrates distribution of energy simulation tools used in each cluster. 97% of the buildings used Trace, eQuest, HAP, EnergyPro, while the other 3% of the buildings used other simulation tools including EnergyPlus, Energy Gauge Summit Premier, and DOE-2. In general, this finding indicates a strong user preference towards simpler energy simulation tools. Among the clusters, the medium intensity cluster comprises a fairly similar distribution of Trace, eQuest, HAP, and EnergyPro simulations, while buildings within the low intensity cluster used eQuest and Trace and the

high intensity cluster buildings mostly used Trace to design the buildings.

Fig. 8 shows the year when the applications for the reviewed projects were submitted. The results indicate the patterns in the design of building over the five years have not been changed. Medium intensity cluster has the highest number of application per year with the low intensity and high intensity clusters are following. In addition, the results for the reviewed LEED NC office buildings illustrate that most of the projects selected in this study are from years of 2009 and 2010.

In summary, the results show the EAc1 point is a better indicator to specify the building energy use pattern rather than total points of the building, although the certification level is also correlated with the identified clusters. Distribution of the energy simulation tools have influenced the design of the building clusters, while the year of application submission do not show any significant changes for the identified clusters.

4.2. Use of factor analysis

The goal of the factor analysis is to reduce the dimensionality of the examined associations between the total energy use and correlated variables into key factors that explain most of the variance. Building size is the most significant predictor of building energy use, so this study used GFA to normalize occupant and energy use variables to investigate other variable effects. Factor analysis requires two characteristics to ensure meaningful data factorability. First, there need to be a sufficient number of cases, generally over 100, and at least five cases per variable of interest [45]. The characteristic is that there is sufficient correlation with several variables, above 0.3 [46]. While the data are sufficient in number, there are not significant correlations outside of sub-groups of highly correlated weather variables. Several combinations of variables did not produce an adequate matrix for factor analysis that had significant enough correlations and was reliably different from an identity matrix [47]. This study discourages factor analysis as a tool in analyzing sets of energy simulation data.

4.3. Comparison with the DOE's commercial reference buildings

An important comparison is how these LEED energy simulation results differ from other energy simulations of office buildings, including the Reference Buildings based on ASHRAE standard 90.1-2004. In general, most of the buildings in the three clusters are

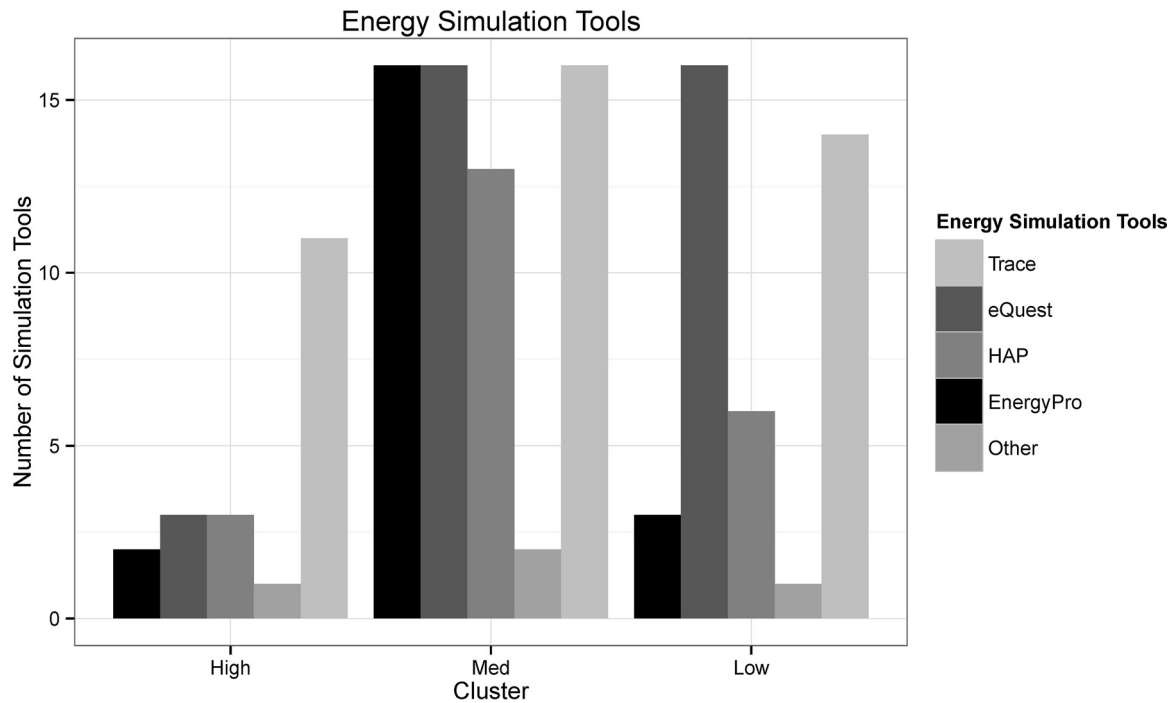


Fig. 7. Distribution of energy simulation tools used for the three identified clusters.

within the medium-sized office buildings. Fig. 9 shows a comparison of the floor area of the reference buildings and the floor area for office buildings in the LEED dataset for different climate zones. Colder climate zones are relatively more populous in the LEED dataset and warmer climate zones less so, relative to the reference buildings.

A comparison of the annual site EUI of the reference office buildings, of three different sizes, and across all U.S. climate zones, indicate the LEED buildings have greater EUI variability across climate zones than the reference buildings. The EUIs are within

the standard error of the EUIs of the reference buildings with the exception of zones 4A and 5A, which are higher than the reference values, and zones 5B and 6A which are lower. These differences indicate that standardized building simulations with reference building characteristics can provide best practices for energy simulation professionals, but they also can be very different from actual building energy consumption due to specific features of individual buildings.

In summary, an energy simulation of the medium size reference office building in Chicago, built to ASHRAE Standard 90.1-2004

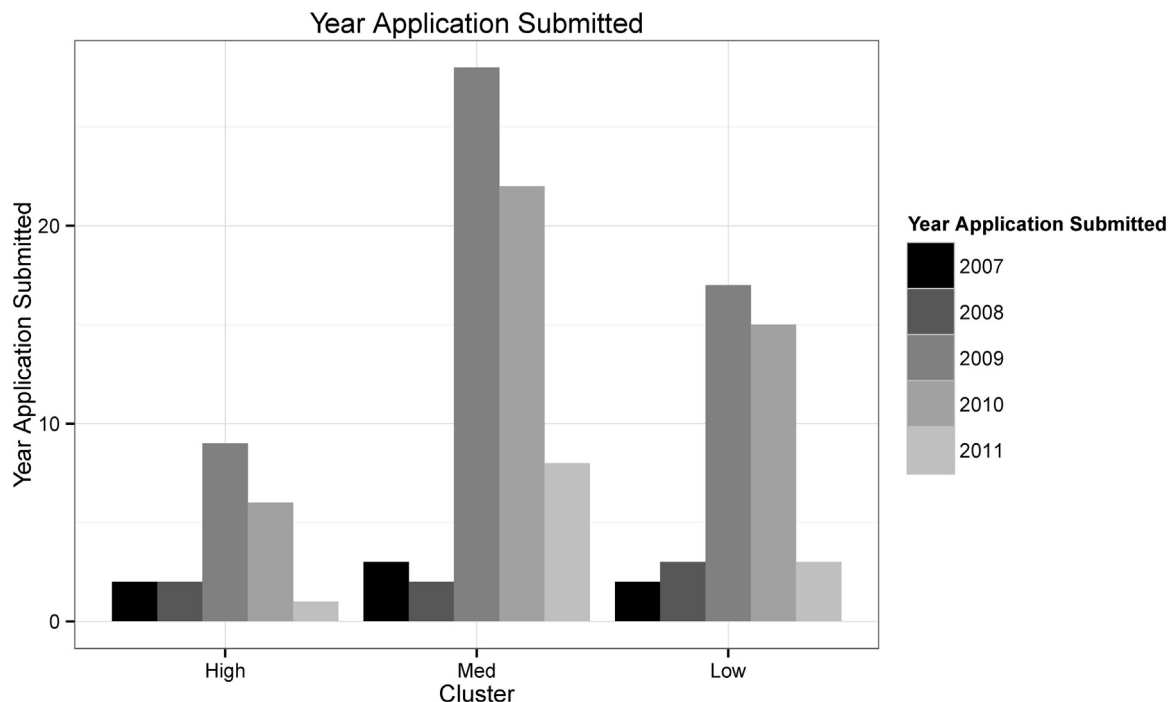


Fig. 8. Distribution of year application was submitted for the LEED certification for different clusters.

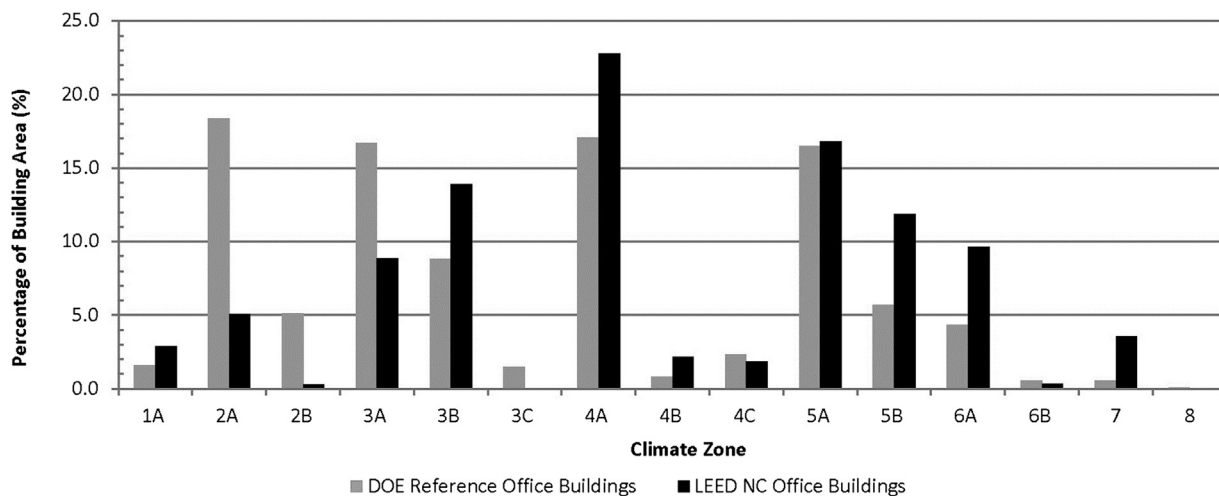


Fig. 9. Comparison of the area between simulated LEED NC Offices and for the DOE new construction reference buildings.

provides a comparison for buildings in the most frequent climate zone for the LEED buildings, climate zone (5A). The default process loads are 7.5 W/m^2 (0.7 W/ft^2) and lighting loads are 10.8 W/m^2 (1.0 W/ft^2), giving the energy simulation output EUIs of 59 kWh/m^2 (18.8 kBtu/ft^2) for process EUI, comparable to a building in the medium intensity cluster. The simulation output for lighting EUI is 34 kWh/m^2 (10.7 kBtu/ft^2), which is greater than the EUIs of all the clusters, suggesting a need for more stringent lighting efficiency standards. The total site EUI is within a standard deviation of the mean site EUI of a medium intensity building in the LEED dataset. Most of the LEED buildings, at the whole building level, have a similar energy intensity as the energy intensity of the NREL buildings designed to meet ASHRAE Standard 90.1-2004 at the component level. With an exception of the buildings in the low intensity cluster that show significantly improved performance compared to the NREL building averaged performance.

5. Conclusions

This study used cluster analysis to determine energy use characteristics and develop three clusters for 134 LEED NC certified office buildings. These buildings cover 13 climate zones in the U.S., and vary in size from 185 m^2 ($\sim 1,991 \text{ ft}^2$) to $18,580 \text{ m}^2$ ($\sim 199,999 \text{ ft}^2$). Three clusters with unique similarities and differences include (1) low, (2) medium, and (3) high energy use intensity. The difference between the low energy use cluster and other clusters is explained mostly by lower process loads and lower heating energy intensity, and partly by lower intensities of other HVAC related end-uses. Lighting energy use shows the least variation between clusters. The lower heating energy intensity in the low intensity cluster is largely explained by lower roof U-values, lower window-to-wall ratio, and smaller building size. The roof U-values are significantly lower than is required by ASHRAE 90.1-2004. Compared to buildings in the high intensity cluster, buildings in the medium intensity cluster are relatively smaller, are located in milder climates, and have higher window and wall U-values. Buildings in the high intensity cluster have lower wall U-values compared to buildings in the two other clusters, and share a similar climate distribution with the low energy intensity cluster. No significant variation exists in window SHGC among clusters.

In general for all of the three clusters, HVAC end uses (heating, cooling, and ventilation) share a similar percentage of building energy use across clusters. Unregulated process loads are the most significant contributor to total building energy use, accounting for

36%, 33%, and 31% of the high, medium, and low intensity clusters, respectively. This suggests that improving assumptions, accuracy, and granularity of internal process loads is necessary to accurately predict energy performance in buildings. Energy simulations guidelines and rating programs should consider allowing and promoting common techniques for reducing internal loads as valid methods for reducing total building energy use.

This study benefited from a large data set of energy simulation results that quantified each energy end-use contribution from interior lighting, process loads, cooling, heating, and fans. Datasets of energy simulations are currently the only way to isolate energy end-uses and performance due to inherent building characteristics and location, in absence of a similarly sized dataset of sub-metered data. For the reviewed LEED certified buildings, 97% of the projects used Trace, eQuest, HAP, EnergyPro simulation tools. Buildings in the medium intensity cluster have similar distribution of Trace, eQuest, HAP, and EnergyPro. Buildings within the low intensity cluster buildings used eQuest and Trace and the high intensity cluster buildings used Trace. This indicates majority of LEED NC office buildings used Trace and eQuest. The Energy and Atmosphere Credit 1 (EAc1) to optimize the building energy performance of the buildings has a direct correlation with the building clusters. In general, low intensity buildings benefit from higher points. Similarly, the buildings with higher certification levels are major contributors to the lower energy intensity clusters. 30%, 33%, and 40% of the buildings in the low, medium, and high intensity clusters are Platinum, Gold, and Silver certified office buildings, respectively.

Disclaimer

This paper was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Acknowledgments

This material is based upon work supported by the Energy Efficient Buildings Hub (EEB Hub), an energy innovation hub sponsored by the U.S. Department of Energy under Award Number DE-EE0004261. Authors would like to thank Liusheng Yan and Yang-Seon Kim at the Pennsylvania State University for their support in data entry and analyses.

References

- [1] Commercial Building Energy Consumption Survey (CBECS), 2003. U.S. Energy Information Administration (EIA). Available from: <http://www.eia.gov/consumption/commercial/data/2003/> (accessed December 2013).
- [2] Portfolio Manager, US Environmental Protection Agency (EPA). Available from: http://www.energystar.gov/index.cfm?c=evaluate_performance.bus_portfolio.manager (accessed December 2013).
- [3] LEED 2009 for New Construction and Major Renovations. EA Prerequisite 2: Minimum Energy Performance. U.S. Green Building Council, 2009.
- [4] 2010 Buildings Energy Data Book. Table 3.2.2. U.S. Department of Energy, 2011.
- [5] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, *Energy and Buildings* 40 (3) (2008) 394–398.
- [6] J.H. Scofield, Do LEED-certified buildings save energy? Not really. . . , *Energy and Buildings* 41 (12) (2009) 1386–1390.
- [7] C. Turner, M. Frankel, Energy Performance of LEED for New Construction Buildings, New Buildings Institute, Vancouver, WA, 2008.
- [8] J.H. Scofield, Efficacy of LEED-certification in reducing energy consumption and greenhouse gas emission for large New York City office buildings, *Energy and Buildings* 67 (0) (2013) 517–524.
- [9] D. Harris, C. Higgins, Key performance indicators and analysis for commercial buildings with system level data, fueling our future with efficiency, 2012 ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA, 2012.
- [10] M.C. Lorek, F. Chraim, K.S.J. Pister, S. Lanzisera, COTS-based stick-on electricity meters for building submetering, *IEEE Sensors* 13 (2013) 1–4.
- [11] N. Brown, R. Bull, F. Faruk, T. Ekwevugbe, Novel instrumentation for monitoring after-hours electricity consumption of electrical equipment and some potential savings from a switch-off campaign, *Energy and Buildings* 47 (0) (2012) 74–83.
- [12] P. Delgoshai, K. Xu, S. Wagner, R. Sweetser, J. Freihaut, Hourly plug load measurements and profiles for a medium office building – a case study, in: AEI 2013, American Society of Civil Engineers, University Park, PA, 2013, pp. 827–836.
- [13] M. Deru, K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, M. Yazdani, J. Huang, D. Crawley, U.S. Department of Energy Commercial Reference Building Models of the National Building Stock, NREL, Golden, CO, 2011.
- [14] K. Xu, Assessing the Minimum Instrumentation to Well Tune Existing Medium Sized Office Building Energy Models, Architectural Engineering, The Pennsylvania State University, 2012.
- [15] M. Dahlhausen, M. Heidarinejad, J. Srebric, Managing energy retrofits, in: IFMA's World Workplace Conference 2013, IFMA, Philadelphia, PA, USA, 2013.
- [16] M. Dahlhausen, Staging Building Energy Retrofits, Architectural Engineering, The Pennsylvania State University, 2014.
- [17] N. Fumo, P. Mago, R. Luck, Methodology to estimate building energy consumption using EnergyPlus Benchmark Models, *Energy and Buildings* 42 (12) (2010) 2331–2337.
- [18] L.C. Ng, A. Musser, A.K. Persily, S.J. Emmerich, Multizone airflow models for calculating infiltration rates in commercial reference buildings, *Energy and Buildings* 58 (0) (2013) 11–18.
- [19] L.C. Ng, A. Musser, A.K. Persily, S.J. Emmerich, Indoor air quality analyses of commercial reference buildings, *Building and Environment* 58 (0) (2012) 179–187.
- [20] S.P. Corgnati, E. Fabrizio, M. Filippi, V. Monetti, Reference buildings for cost optimal analysis: method of definition and application, *Applied Energy* 102 (0) (2013) 983–993.
- [21] P. May-Ostendorp, G.P. Henze, C.D. Corbin, B. Rajagopalan, C. Felsmann, Model-predictive control of mixed-mode buildings with rule extraction, *Building and Environment* 46 (2) (2011) 428–437.
- [22] M. Bhandari, S. Shrestha, J. New, Evaluation of weather datasets for building energy simulation, *Energy and Buildings* 49 (0) (2012) 109–118.
- [23] Y.T. Chae, J. Kim, H. Park, B. Shin, Building energy performance evaluation of building integrated photovoltaic (BIPV) window with semi-transparent solar cells, *Applied Energy* 129 (0) (2014) 217–227.
- [24] S. Wang, C. Yan, F. Xiao, Quantitative energy performance assessment methods for existing buildings, *Energy and Buildings* 55 (0) (2012) 873–888.
- [25] H. Xiao, Q. Wei, Y. Jiang, The reality and statistical distribution of energy consumption in office buildings in China, *Energy and Buildings* 50 (0) (2012) 259–265.
- [26] N. Gaitani, C. Lehmann, M. Santamouris, G. Mihalakakou, P. Patargias, Using principal component and cluster analysis in the heating evaluation of the school building sector, *Applied Energy* 87 (6) (2010) 2079–2086.
- [27] D. Hsu, Characterizing energy use in New York city commercial and multifamily buildings, 2012 ACEEE Summer Study on Energy Efficient in Buildings, Pacific Grove, CA, 2012.
- [28] ASHRAE Standard 90. 1-2007, Energy Standard for Buildings except Low-Rise Residential Buildings, American Society of Heating, Refrigeration Air Conditioning Engineers, Inc., Atlanta, GA, 2007.
- [29] LEED 2009 For New Construction and Major Renovations, USGBC, Washington, DC.
- [30] California Energy Commission, Building energy efficiency standards for residential and non-residential buildings, California Code of Regulations, Title 24, Part 6, 2005.
- [31] LEED 2009 for new construction and major renovations. EA Prerequisite 2: Minimum Energy Performance, U.S. Green Building Council, 2009.
- [32] Advanced Energy Modeling For LEED, Technical Manual v2.0, U.S. Green Building Council, Washington, DC, 2011.
- [33] Advanced Energy Modeling For LEED, Technical Manual v1.0, U.S. Green Building Council, Washington, DC, 2010.
- [34] M. Heidarinejad, Relative Significance of Heat Transfer Processes to Quantify Tradeoffs Between Complexity and Accuracy of Energy Simulations with a Building Energy use Patterns Classification, Mechanical Engineering, Pennsylvania State University, 2014.
- [35] S. Wilcox, W. Marion Users Manuals for TMY3 Data Sets, 2008.
- [36] C.E. Commission, Nonresidential Alternative Calculation Method (ACM) Approval Manual for the 2005 Building Energy Efficiency Standards for Residential and Nonresidential Buildings, 2005.
- [37] A.H. Fielding, Cluster and Classification Techniques for the Biosciences, Cambridge University Press, New York, NY, 2007.
- [38] M. Santamouris, G. Mihalakakou, P. Patargias, N. Gaitani, K. Sfakianaki, M. Papaglastra, C. Pavlou, P. Doukas, E. Primikiri, V. Geros, M.N. Assimakopoulos, R. Mitoula, S. Zerefos, Using intelligent clustering techniques to classify the energy performance of school buildings, *Energy and Buildings* 39 (1) (2007) 45–51.
- [39] M.S. Al-Homoud, Optimum thermal design of office buildings, *International Journal of Energy Research* 21 (10) (1997) 941–957.
- [40] K-means clustering, MathWorks.
- [41] S. Poll, C. Teubert, Pilot study of a plug load management system: preparing for sustainability base, 2012 IEEE Green Technologies Conference, 2012, pp. 1–6.
- [42] C. Lobato, S. Pless, M. Sheppy, T. Paul, Reducing plug and process loads for a large scale, low energy office building: NREL's research support facility, in: ASHRAE Winter Conference, ASHRAE, Las Vegas, Nevada, USA, 2011.
- [43] GSA plug load reduction checklist. U.S. General Services Administration.
- [44] Advanced Control Options for LEED v4. Available from: <http://www.ies.org/Ida/HotTopics/LED/15.cfm> (accessed December 2013).
- [45] F.B. Bryant, P.R. Yarnold, Principal-components analysis and exploratory and confirmatory factor analysis, in: L.G. Grimm, P.R. Yarnold (Eds.), *Reading and Understanding Multivariate Statistics*, American Psychological Association, Washington, DC, 1995, pp. 99–136.
- [46] B.G. Tabachnick, L.S. Fidell, *Using Multivariate Statistics*, Allyn & Bacon, Inc., Needham Heights, MA, USA, 2006.
- [47] H. Kaiser, An index of factorial simplicity, *Psychometrika* 39 (1) (1974) 31–36.