Y. Lou, Y. Ye, Y. Yang, W. Zuo, G. Wang, M. Strong, S. Upadhyaya, C. Payne. 2022. "Individualized Empirical Baselines for Evaluating the Energy Performance of Existing Buildings". *Science and Technology for the Built Environment*, pp. 1-15. https://doi.org/10.1080/23744731.2022.2134680

Individualized Empirical Baselines for Evaluating the Energy Performance of Existing Buildings

Yingli Lou[a], Yunyang Ye[a*], Yizhi Yang[a], Wangda Zuo[b,c,**], Gang Wang[d], Matthew Strong[e], Satish Upadhyaya[e], Chris Payne[f]

[a] Department of Civil, Environmental and Architectural Engineering, University of Colorado Boulder, Boulder, CO 80309, USA.

[b] Department of Architectural Engineering, Pennsylvania State University, University Park, PA 16802, USA.

[c]National Renewable Energy Laboratory, Golden, CO 80401, USA.

[d] Civil, Architectural and Environmental Engineering Department, University of Miami, Coral Gables, FL 33146, USA.

[e] Department of Computer Science, University of Colorado Boulder, Boulder, CO 80309, USA [f] Fairview High School, Boulder, CO, 80305, USA.

*Current working in Pacific Northwest National Laboratory, Richland, WA 99354, USA.

** Corresponding author. Email address: wangda.zuo@psu.edu (Wangda Zuo)

Abstract

The evaluation of building energy performance requires a baseline for comparison. Common empirical baselines are usually used for existing buildings since they are fast and convenient. However, the same type of building at the same location will receive the same baseline despite their difference in usage. Individualized baselines by creating building energy models are possible solutions, but it is labor intensive and time-consuming. To fill the gap, this study is to develop individualized empirical baselines for existing buildings in a fast way. First, common empirical baselines are created based on survey data. Then, to get training samples, building energy models for large-scale existing buildings are created and simulated. Finally, based on simulation results, mathematical models to get individualized empirical baselines in a fast way are created. U.S. medium office buildings were used as an example to demonstrate the method. We developed 30 mathematical models for medium office buildings in two vintages (constructed before 1980 and after 1980) and 15 climate zones. The mean absolute percentage errors (MAPE) between the individualized empirical baselines and the modeled baselines for those 30 mathematical models are all lower than 5.5%. An engineer can obtain the individualized empirical baseline for an existing building in a few seconds by using the open-source tool we developed.

Keywords: Building energy rating; Empirical baseline; Existing buildings; Building energy model; Large-scale simulation.

Nomenclature

$CV(RMSE)$	Coefficient of variation of the root-mean-square error
D	Dataset for all building samples
D^v	One subset of dataset D
$\|D\|$	Total number of samples in dataset D
$\|D^v\|$	Total number of samples in subset D^v
k	Key model input k that has a significant impact on building energy consumption
Emp	Individualized empirical baseline site EUI
EUI	Empirical baseline site EUI
EUI	Modeled site EUI from prototypical building energy models
$Gini\ (D^v)$	Gini value
$Gini\_impurity(D, j)$	Gini impurity
i	Building i
INP	Model input
$INP_{m,k}$	The value of key model input k for prototypical building energy models m
$INP\_r_{i,j}$	Ranked value of model input j for building i in all building samples
j	Neutral building characteristic j
m	Prototypical building energy models m
$MAPE$	Mean absolute percentage error
Mod	Modeled baseline site EUI
Mod_r_i	Ranked value of modeled baseline site EUI for building i in all building samples
n	Number of building samples
p_t	The probability that output type t occurs in subset D^v
q	Label q of building site EUI
Q	The number of classes in the label of building site EUI
$RMSE$	Root-mean-square error
T	Total number of the output types
V	Total number of the subsets
ρ	Spearman's correlation coefficient

1. Introduction

The evaluation of building energy performance requires a baseline for comparison. Historically, common empirical baselines are usually used for existing buildings. The empirical baseline represents the measured median energy use intensity (EUI) for a group of similar buildings. The measured EUI of a candidate building is then compared with the empirical baseline. For example, the median EUI of medium office buildings in the hot climate zone is 700 MJ/m^2-yr. If a medium office building is in the hot climate zone and its measured EUI is 600 MJ/m^2-yr, this building can be labeled as an energy-efficient building. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Building Energy Quotient In Operation rating (ASHRAE 2022), the United States (U.S.) Environmental Protection Agency ENERGY STAR program (ENERGY STAR 2021), and ASHRAE Standard 100 (ASHRAE 2015) adopt empirical baselines. They create empirical baselines based on the median EUI of comparable buildings based on the U.S. Commercial Building Energy Consumption Survey (CBECS). However, the same type of buildings at the same location has the same empirical baseline despite their difference in usage. For examples, buildings with different total floor areas have the same empirical baseline.

Developing individualized modeled baselines for existing buildings is a possible solution. However, it requires creating building energy models for existing buildings, which is labor intensive and time-consuming. Because the creation of building energy models highly relies on detailed building physical feature information. A typical building energy model requires more than 1,000 model inputs, such as construction material property of each part of buildings, detailed information for HVAC systems, and occupancy schedules.

One existing approach to address this problem is adding adjustments to the empirical baseline to account for some neutral building characteristics, such as operating hours, plug loads, and other factors that are meant to be neutral in the comparison of energy performance. For example, a building with higher operating hours has a higher empirical baseline. The existing adjustments were made based on measured data of existing building samples. However, this can only investigate limited neutral building characteristics because the existing large-scale building energy survey data (e.g., CBECS) has limited information on buildings. For example, the number of people in the building and the space temperature is difficult to collect for extensive building samples. As a result, many existing buildings still have a common empirical baseline.

To further develop individualized empirical baselines in an efficient way, this paper developed mathematical models for existing buildings based on a large-scale building energy simulation result. The rest of this paper is organized as follows: Section 2 introduces the methodology including the calculation of empirical baselines, the generation of training samples, key neutral building characteristics, and the method to develop individualized empirical baselines in a fast way. Section 3 shows the case study on U.S.

medium office buildings. Section 4 discusses the application of this research. Finally, Section 5 concludes the findings of this research.

2. Methodology

This research developed a method to obtain individualized empirical baselines for existing buildings in a fast way, as shown in Fig.1. First, empirical baselines are calculated based on survey data. Then, training samples are generated by creating and simulating a large sample of building energy models. Next, we need to identify key building characteristics that are meant to be neutral when evaluating energy performance. The definition of neutral building characteristics is further explained in subsection 2.3.1. Finally, individualized empirical baselines are developed using the key neutral building characteristics as input variables. The individualized empirical baseline for a candidate building can be obtained in a few seconds. Following four subsections will introduce these four steps in detail.

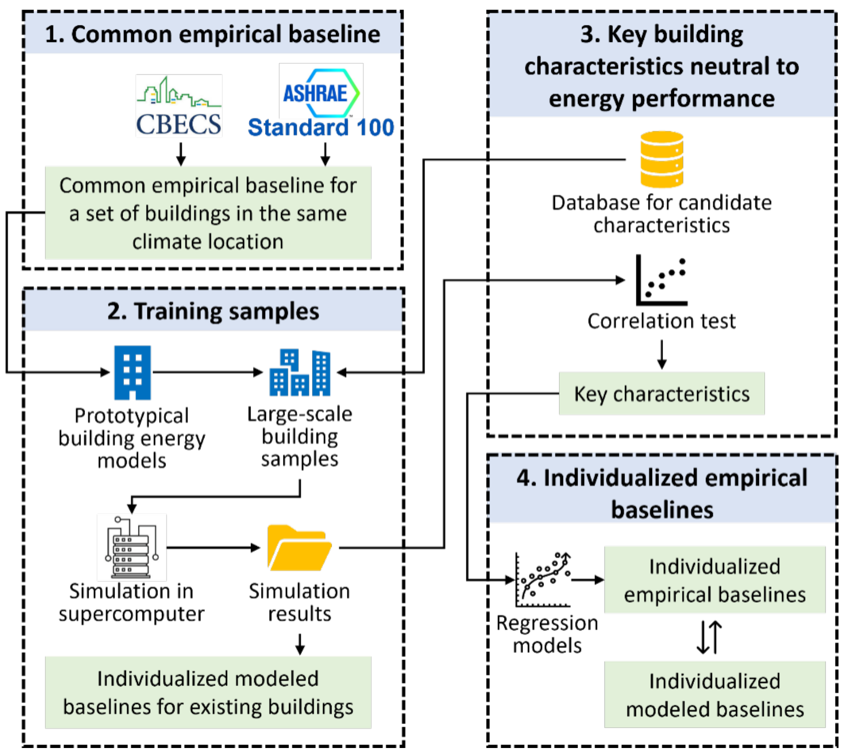

Fig.1. Methodology of developing individualized empirical baselines for existing buildings.

2.1. Calculation of common empirical baseline

To identify empirical baseline EUIs in various climate zones, ASHRAE standard 100 (ASHRAE 2017) provided climate zonal EUI ratios. These ratios were used to derive climate zonal EUIs for each building type by multiplying them with the CBECS's national median EUIs. This method produces representative total EUIs by building type and climate zone. We adopted this method to calculate empirical baseline EUIs for 15 climate zones in two vintages (pre-1980 and post-1980). Vintage pre-1980 means that the building was constructed before 1980 and vintage post-1980 means that the building was constructed in or after 1980.

2.2. Generation of training samples

To get a large-scale training sample, prototypical building energy models that can represent empirical data first need to be created. Then, subsection 2.2.2 introduces the generation of individualized modeled baselines for a large sample of buildings using the prototypical building energy models as a starting point.

2.2.1. Prototypical building energy models

Fig. 2. shows the process of creating prototypical building energy models that can represent empirical data. First, we need to identify the ranges of key model inputs and relations of key model input values in different climate zones and vintages. For example, the insulation of the building should be more in colder climate zones. The relation types of key model inputs are summarized in Table 1. Then, the values of various key model inputs are defined through the model calibration. The calibration's goal is to minimize the difference between modeled energy consumption and empirical baselines. More detailed information about the creation of prototypical building energy models can be found in our previous research (Ye et al. 2020).

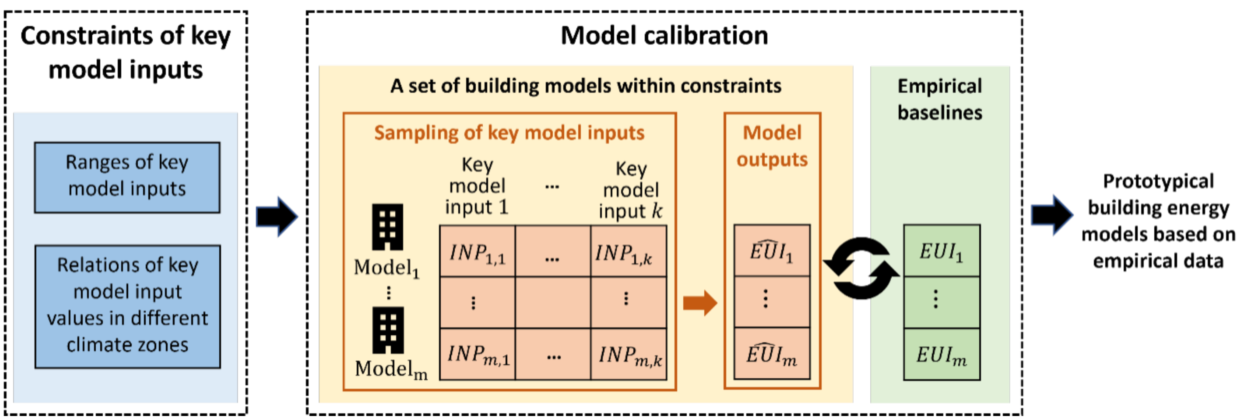

Fig. 2. Workflow of creating prototypical building energy models based on empirical data.

Table 1. Relation of model inputs in a set of building models

Relation Index	Model Inputs Relation Description		Model Input Example
	Climate	Vintages	
Type 1	Values in all climate zones are same	Values for post-1980 and pre-1980 models are same	Weekly operation hours
Type 2	Values in all climate zones are same	Values for post-1980 models are not higher than pre-1980 models	Electric equipment power density
Type 3	Values in all climate zones are same	Values for post-1980 models are not lower than pre-1980 models	Rated cooling COP
Type 4	Values in climate zones 5~8 are not higher than the other climate zones.	No constraint	Window U-factor
Type 5	Values in climate zones 5~8 are not lower than the other climate zones	No constraint	Exterior wall insulation R-value

Prototypical building energy models are evaluated using the coefficient of variation of the root-mean-square error (CV(RMSE)). According to ASHRAE Guideline 14, when the CV(RMSE) is lower than 0.15, the modeled energy consumption is consistent with the empirical data (ASHRAE 2014). To calculate the CV(RMSE), we must calculate the root-mean-square error (RMSE) first, as shown in the following equation:

$$RMSE = \sqrt{\frac{\sum_{m=1}^{30} (EUI_m - EUI_m)^2}{30}}, \tag{1}$$

where m is the prototypical building energy model, which has 30 models in total (15 climate zones×2 vintages); EUI is the empirical baseline site EUI; EUI is the modeled site EUI of the prototypical building energy model.

Based on the results of RMSE, we can calculate the CV(RMSE) by using the following equation:

$$CV(RMSE) = \frac{RMSE}{avg(EUI_m)}, \tag{2}$$

where $avg(EUI_m)$ is the average value of the empirical baseline site EUI in 15 climate zones and two vintages, which can be calculated using the following equation:

$$avg(EUI_m) = \frac{\sum_{m=1}^{30} EUI_m}{30}, \tag{3}$$

2.2.2. Training samples

Building model inputs can be classified as neutral inputs and non-neutral inputs (or building assets). Neutral inputs are inputs that do not affect the energy efficiency rating. For example, climate and occupied hours should not affect the building performance evaluation. The definition of neutral building characteristics is further explained in subsection 2.3.1. Non-neutral inputs are inputs with a direct impact on energy consumption and affect the energy efficiency rating, such as exterior wall insulation and HVAC system efficiency.

To make the evaluation of building energy performance focus on the building's assets, when creating the baseline model for a candidate building, neutral model inputs should be the same as the candidate building while non-neutral model inputs should be generalized values. In this research, the generalized values of non-neutral model inputs are the values of the prototypical building energy models created in subsection 2.2.1. The neutral model inputs and their value range for candidate buildings will be introduced in subsection 2.3.1. Following introduces the steps to generate modeled baselines.

First, neutral model input combinations are sampled using the Latin Hypercube Sampling (LHS) (McKay et al. 2000), which is usually used in the sampling of building energy model inputs (Chen et al. 2019; Lim and Zhai 2017, 2018). LHS is a statistical method for generating a near-random sample of parameter values from a multidimensional distribution. It is difficult to reveal the distributions of the neutral model inputs because most neutral inputs are not included in existing survey data. We assume that the distributions of the neutral model inputs are all near-random and LHS is adopted for sampling.

Then, large-scale baseline models are generated by modifying the neutral model inputs of prototypical building energy models. Python coding and building component library (NREL 2022a) are adopted to automize this process. To reduce the simulation time, this research adopts parallel simulation, and the RMACC Summit Supercomputer (CU Boulder 2022) at the University of Colorado Boulder is used for parallel simulation.

Finally, the simulation results are post-processed to extract useful information. The values of neutral model inputs and the simulation results of energy use intensity for all candidate buildings are saved in a single CSV file.

2.3. Key neutral building characteristics

Building characteristics that should be neutral for building energy rating are first introduced in subsection 2.3.1. Then, subsection 2.3.2 introduces the method of identifying key neutral building characteristics.

2.3.1. Neutral building characteristics

Neutral building characteristics are those that do not affect the energy rating of a building. Physical characteristics that are dictated by the building's architectural design and functions are generally neutral. For example, the total floor area is designed by the building's functions. Building operation characteristics, such as the number of occupants in the building and operation hours, should also be considered as neutral building characteristics. The location and the weather conditions that the building is exposed to are neutral building characteristics. A building should not get a poor rating because it is in a colder climate. This research will propose neutral building characteristics based on the literature review and engineering judgment.

2.3.2. Identification of key neutral building characteristics

Key neutral building characteristics are identified by conducting the correlation test between the proposed neutral building characteristics and site EUIs of building samples generated in subsection 2.2.2. Spearman's correlation coefficient (ρ) is adopted for the correlation test. It is a nonparametric measure of rank correlation. It assesses how well the relationship between two variables can be described using a monotonic function. The calculation of ρ is expressed as follows:

$$\rho_j = \frac{n \times \sum_{i=1}^{n} INP\_r_{i,j} \times Mod\_r_i - \sum_{i=1}^{n} INP\_r_{ij} \times \sum_{i=1}^{n} Mod\_r_i}{\sqrt{n \times \sum_{i=1}^{n} INP\_r_{i,j}{}^2 - (\sum_{i=1}^{n} INP\_r_{i,j})^2}\sqrt{n \times \sum_{i=1}^{n} Mod\_r_i{}^2 - (\sum_{i=1}^{n} Mod\_r_i)^2}}, \quad (4)$$

where j is one of the neutral building characteristics; n is the total number of building samples; i is the index for each building sample; $INP\_r_{i,j}$ is the ranked value of model input j for building i; Mod_r_i is the ranked value of the modeled baseline site EUI for building i. $|\rho| \leq 0.1$ means that the correlation between two variables is negligible (Schober et al. 2018). Therefore, $|\rho| > 0.1$ is used as a threshold to identify the key building characteristics that need to be considered when developing individualized empirical baselines in subsection 2.4.

2.4. Individualized empirical baselines

This research obtains individualized empirical baselines in a fast way by developing mathematical models. The key neutral building characteristics identified in subsection 2.3.2 are used as input variables for the mathematical models. Extra Trees (Geurts et al. 2006; Scikit-Learn 2022) is adopted to estimate the relations between key neutral building characteristics and building site EUIs. The structure of Extra Trees is summarized in Fig. 3.

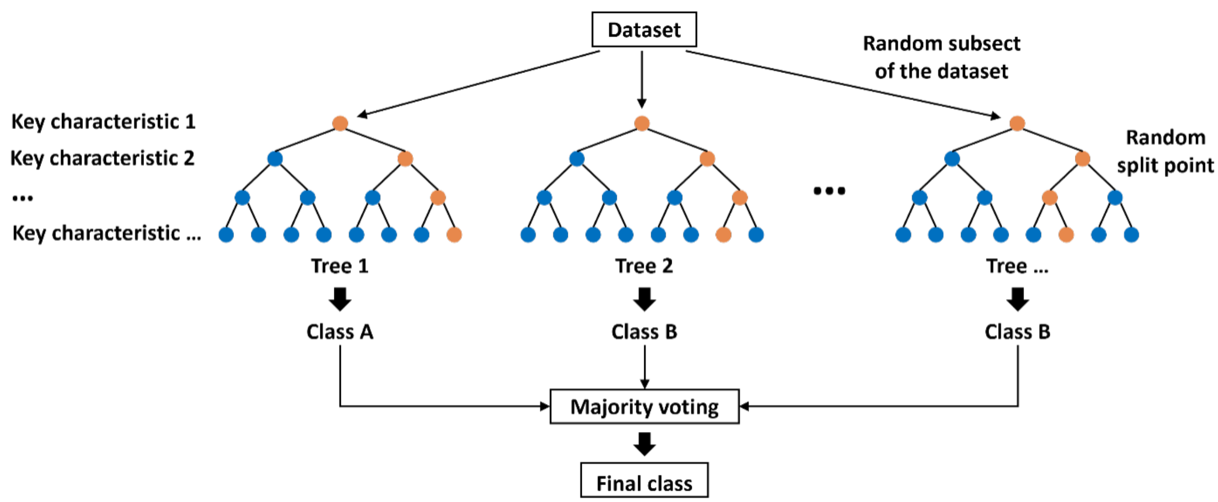

Fig. 3. Structure of Extra Trees (Kapoor 2020).

Extra Trees constructs the set of decision trees by randomly selecting a subset of the dataset. In the training of each decision tree (Wikipedia 2022), the split point to divide the tree at a particular node is randomly selected. (Chu et al. 2021). Each decision tree generates one prediction, and the final prediction is based on the majority prediction.

Extra Trees is a machine learning techniques and was developed as an extension of random forest algorithm, and is less likely to overfit a dataset (Geurts et al. 2006). Extra Trees is very similar to Random Forest. Both are composed of a large number of decision trees, where the final decision is obtained taking into account the prediction of every tree. Furthermore, when selecting the partition of each node, both of them randomly choose a subset of the dataset. The difference between Extra Trees and Random Forest is the selection of cut points in order to split nodes. Random Forest chooses the optimum split while Extra

Trees chooses it randomly (John et al. 2015). Therefore, in terms of computational time, the Extra Trees algorithm is faster because it randomly chooses the split point and does not calculate the optimal one.

The method to measure the quality of the split for one characteristic is the Gini impurity (Scikit-Learn 2022; Yuan et al. 2021), as shown in the following equation:

$$Gini\_impurity(D, j) = \sum_{v=1}^{V} \frac{|D^v|}{|D|} Gini(D^v) , \tag{5}$$

where, D^v refers to one subset of dataset D classified based on characteristic j, V refers to the total number of the subsets, $|D^v|$ and $|D|$ refer to the total number of samples in subset D^v and in dataset D respectively. $Gini\ (D^v)$ is the Gini value of subset D^v, which can be expressed in the following equation:

$$Gini(D^v) = 1 - \sum_{t=1}^{T} {p_t}^2 , \tag{6}$$

where T is the total number of the output types in subset D^v, p_t is the probability that output type t occurs in subset D^v. The less the Gini value is, the higher the purity of the dataset is.

Gini impurity has a maximum value of 0.5, which is the worst we can get. This means that under the split point for one characteristic, the outputs of samples are evenly distributed. Gini impurity has a minimum value of 0 is the best we can get. This means that under split point for one characteristic, samples have a same output.

This research used 80% of building samples to train the Extra Trees' models, and the rest of the 20% of building samples were used to validate the developed mathematical models. The mean absolute percentage error (MAPE) between individualized empirical baselines and modeled baselines is adopted to validate the mathematical models, as shown in the following equation:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{Emp_i - Mod_i}{Mod_i} \right| , \tag{7}$$

where n is the number of building samples; Emp_i is individualized empirical baseline site EUI for building i; Mod_i is the modeled baseline site EUI. If the MAPE value is lower than 10%, the prediction is accurate (Setiawan et al. 2021).

3. Case Study: Medium Office Buildings

U.S. medium office buildings were used as an example to illustrate the methodology of developing individualized empirical baselines. Subsection 3.1 presents the common empirical baselines of U.S. medium office buildings in 15 climate zones and two vintages. There is one common empirical baseline for buildings in the same climate zone and vintage. Individualized modeled baselines of medium office buildings are presented in subsection 3.2. Based on these simulation results, key building characteristics that should be neutral in the comparison of energy performance are identified in subsection 3.3. Finally, using these key neutral building characteristics as input variables, subsection 3.4 develops mathematical models for each climate zone and vintage to get the individualized empirical baseline in a fast way.

3.1. Calculation of common empirical baseline

As described in subsection 2.1, median EUIs of medium offices are needed to derive empirical baselines in each climate zone and vintage. Because the CBECS 2012 only has the label for offices (CBECS 2016), we proposed the criteria to filter medium office buildings from all office buildings, as shown in Fig. 4. Total floor area and the number of floors is used as two indicators to filter medium office buildings. Because the total floor area of the small office reference model is 511 m^2 and the total floor area of the large office reference model is 46,320 m^2 (NREL 2022b), we only consider buildings whose total floor areas are between 511 m^2 and 46,320 m^2, as candidates for medium office buildings. Then, we considered three situations to select medium office buildings among these candidates: (1) buildings whose total floor areas are between 1,000 m^2 and 10,000 m^2 are considered medium office buildings; (2) buildings whose total floor areas are between 511 m^2 and 1,000 m^2, if they have more than one floor, are considered medium office buildings; (3) buildings whose total floor areas are between 10,000 m^2 and 46,320 m^2, if they have less than five floors, are considered medium office buildings.

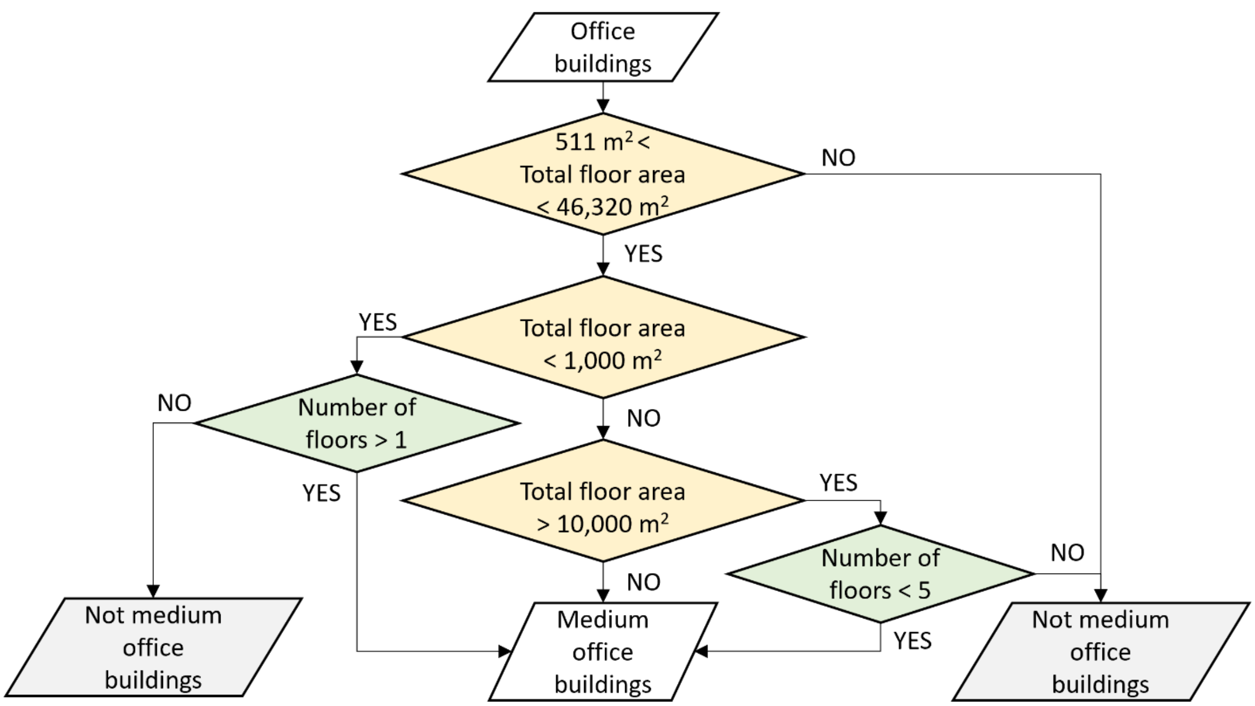

Fig. 4. Criteria to filter medium office buildings from all office buildings

Based on the filtered medium office building samples from the CBECS 2012, the median site EUI is 746.26 MJ/m^2-yr for pre-1980 medium office buildings and is 700.21 MJ/m^2-yr for post-1980 medium office buildings. According to the ratio provided by ASHRAE standard 100 (ASHRAE 2017), empirical baselines of medium office buildings in 15 climate zones and two vintages were calculated, as shown in Table 2. The division of climate zones is shown in Fig. 5. The buildings in climate zone 8 have a significantly higher empirical baseline than the other climate zones due to their large heating needs. The buildings in the post-1980 vintage have a lower empirical baseline than the buildings in the pre-1980 vintage because the newly constructed buildings have better system efficiency (e.g., the efficiency of the cooling coil) than the older buildings.

Table 2. Common empirical baselines of U.S. medium office buildings

Climate Zone	**Weather Feature**	**Representative City**	**Empirical Baselines (MJ/m^2-yr)**	
			Pre-1980	**Post-1980**
1A	Very hot	Miami, FL	731.34	686.20
2A	Hot humid	Tampa, FL	723.87	679.20
2B	Hot dry	Tucson, AZ	731.34	686.20
3A	Warm humid	Atlanta, GA	723.87	679.20
3B	Warm dry	El Paso, TX	694.02	651.19
3C	Warm marine	San Diego, CA	574.62	539.16
4A	Mixed humid	New York, NY	783.58	735.22
4B	Mixed dry	Albuquerque, NM	679.10	637.19
4C	Mixed marine	Seattle, WA	694.02	651.19
5A	Cool humid	Buffalo, NY	828.35	777.23
5B	Cool dry	Denver, CO	716.41	672.20
6A	Cold humid	Rochester, MN	925.37	868.26
6B	Cold dry	Great Falls, MT	813.43	763.23
7	Very cold	International Falls, MN	992.53	931.28
8	Subarctic/arctic	Fairbanks, AK	1388.05	1302.39

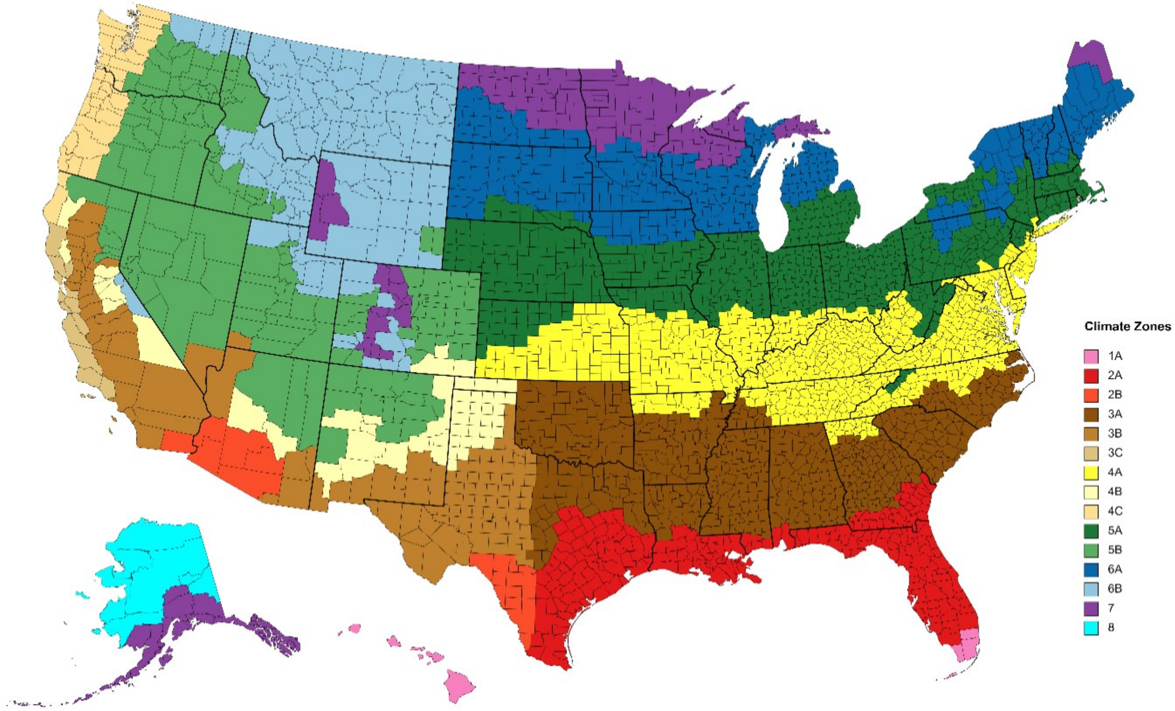

Fig. 5. International Energy Conservation Code climate regions (International Code Council 2022).

3.2. Generation of training samples

Thirty prototypical building energy models (15 climate zones × 2 vintages) for U.S. medium office buildings were created based on the CBECS 2012 (subsection 3.2.1). Using the 30 models as a starting point, subsection 3.2.2 generated 42,000 training samples.

3.2.1. Prototypical building energy models

The initial models used in this research are DOE Reference models generated by OpenStudio Standards (NREL 2022b), which include 30 models (15 climate zones×2 vintages). Key model inputs were calibrated to let the modeled energy consumption match the empirical baselines calculated in subsection 3.1. These key model inputs were defined based on our previous research (Ye et al. 2020, 2021).

For the key model inputs provided in the CBECS 2012, the median value is adopted for prototypical building energy models. For example, the median value of the total floor area for medium office buildings is 3,437m^2, and the median value of the weekly operation is 54 hours. The key model inputs that are not provided in the CBECS 2012 were calibrated using the method described in Fig. 2. The ranges of the model inputs were determined by referring to the 2012 CBECS and publications (Deru et al. 2011; Griffith et al. 2008; Huang and Franconi 1999; NREL 2022b; Sharp 1996; N Wang et al. 2015; Na Wang and Gorrissen 2013; Winiarski et al. 2006, 2007; Ye et al. 2020), as shown in Table 3.

Table 3. Ranges of model inputs for creating prototypical medium office building energy models

Model Input	Unit	Range	Type of Relation*
Aspect ratio	-	[1.5, 2.4]	Type 1
Floor-to-floor height	m	[3.96, 5.69]	Type 1
Window-to-wall ratio	%	[11, 25]	Type 1
Exterior wall insulation R-value	m^2-K/W	Pre-1980: Climate zones 1~4: [0.38, 1.18] Climate zones 5~8: [0.44, 1.69] Post-1980: Climate zones 1~4: [0.18, 2.26] Climate zones 5~8: [0.81, 4.69]	Type 5
Roof insulation R-value	m^2-K/W	Pre-1980: Climate zones 1~4: [1.56, 5.30] Climate zones 5~8: [1.60, 6.10] Post-1980: Climate zones 1~4: [1.80, 3.67] Climate zones 5~8: [1.60, 5.68]	Type 5
Window U-factor	W/m^2-K	Pre-1980: Climate zones 1~4: [4.09, 7.00] Climate zones 5~8: [2.82, 7.00] Post-1980: Climate zones 1~4: [3.27, 7.00] Climate zones 5~8: [1.99, 6.72]	Type 4
Window solar heat gain coefficient (SHGC)	-	Pre-1980: Climate zones 1~4: [0.22, 0.67] Climate zones 5~8: [0.40, 0.77] Post-1980: Climate zones 1~4: [0.25, 0.65] Climate zones 5~8: [0.35, 0.62]	Type 5
Infiltration rate	m^3/s-m^2	[0.00031, 0.00113]	Type 1
People density	person/m^2	[0.0229, 0.0538]	Type 1
Lighting power density	W/m^2	Pre-1980: [10.76, 23.68] Post-1980: [8.61, 18.30]	Type 2
Electric equipment power density	W/m^2	Pre-1980: [5.38, 14.81] Post-1980: [5.38, 13.34]	Type 2
Rated cooling COP	-	Pre-1980: [2.52, 3.39] Post-1980: [2.61, 3.50]	Type 3
Efficiency for heating system	-	Pre-1980: [0.65, 0.80] Post-1980: [0.65, 0.80]	Type 3
Ventilation	m^3/s-person	[0.0066, 0.0261]	Type 1
Efficiency for service water heating equipment	-	Pre-1980: [0.65, 0.80] Post-1980: [0.75, 0.83]	Type 3
Indoor heating setpoint temperature	ºC	[20, 22]	Type 1
Indoor cooling setpoint temperature	ºC	[22, 25]	Type 1

* Type of relation is defined in Table 1.

The geometry of prototypical medium office building energy models is shown in Fig.6. The values of key model inputs are shown in Table 4 and Table 5. The total floor area of the prototypical medium office building energy models is 3,437 m^2 with an 18% window-to-wall ratio. It has steel-frame exterior walls and insulation entirely above deck roofs. The operation time of this building is from 07:45 am - 6:30 pm on weekdays. Operation hour in this research is defined as the time period during which the value of people density is larger than 50% maximum value. These 30 prototypical medium office building energy models are provided in the GitHub repository (Lou 2022b).

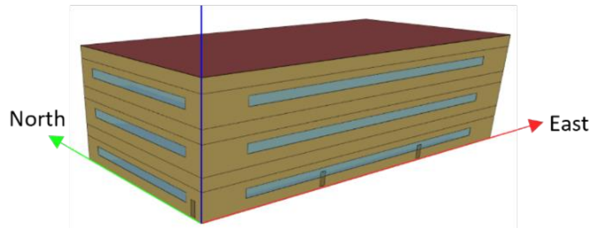

Fig.6. Geometry of prototypical medium office building energy models based on the CBECS 2012.

Table 4. Values of key model inputs of prototypical medium office building energy models based on the CBECS 2012

Category	**Name**	**Value**	
Weather condition	Climate zone	1A, Miami, FL 2A, Tampa, FL 2B, Tucson, AZ 3A, Atlanta, GA 3B, El Paso, TX 3C, San Diego, CA 4A, New York, NY 4B, Albuquerque, NM 4C, Seattle, WA 5A, Buffalo, NY 5B, Denver, CO 6A, Rochester, MN 6B, Great Falls, MT 7, International Falls, MN 8, Fairbanks, AK	
Geometry	Total floor area	3,437m^2	
	Aspect ratio	2.07	
	Floor-to-floor height	4.53 m	
	Window-to-wall ratio	18%	
Envelope	Exterior wall insulation R-value	Table 5	
	Roof insulation R-value	Table 5	
	Window U-factor	Table 5	
	Window SHGC	Table 5	
Schedule	Occupancy schedule	07:45 am - 6:30 pm on weekdays	
	System schedule	05:45 am - 10:30 pm on weekdays	
Internal load	People density	0.035 person/m^2	
	Lighting power density	Pre-1980: 18.00 W/m^2	Post-1980: 15.50 W/m^2
	Electric equipment power density	Pre-1980: 13.08 W/m^2	Post-1980: 8.19 W/m^2
	Infiltration rate	0.0010m^3/s-m^2 for the whole building	
System	Rated cooling COP	Pre-1980: 3.39	Post-1980: 3.50
	Efficiency for heating system	Pre-1980: 0.65	Post-1980: 0.80
	Ventilation	0.012 m^3/s-person for the whole building	
	Efficiency for service water heating equipment	Pre-1980: 0.70	Post-1980: 0.83
	Indoor heating setpoint temperature	20 ºC	
	Indoor cooling setpoint temperature	25 ºC	

Table 5. Model inputs for envelopes of prototypical medium office building energy models based on the CBECS 2012

Name of Input	Unit	Vintage	1A	2A	2B	3A	3B	3C	4A	4B	4C	5A	5B	6A	6B	7	8
Exterior wall insulation R-value	m^2-K/W	Pre-1980	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.19	1.19	1.19	1.19	1.19	1.69
		Post-1980	0.38	1.34	1.34	2.26	2.26	2.26	2.26	2.26	2.26	4.41	4.41	4.41	4.41	4.46	4.69
Roof Insulation R-value	m^2-K/W	Pre-1980	4.87	4.87	4.87	4.87	4.87	4.87	4.87	4.87	4.87	4.91	4.91	4.91	4.91	4.91	5.90
		Post-1980	2.26	3.17	3.17	3.67	3.67	3.67	3.67	3.67	3.67	4.31	4.31	4.31	4.31	5.36	5.68
Window U-factor	W/m^2-K	Pre-1980	4.85	4.85	4.85	4.85	4.85	4.85	4.85	4.85	4.85	4.73	4.73	4.11	4.11	3.99	3.41
		Post-1980	5.25	3.86	3.86	3.29	3.29	3.29	3.29	3.29	3.29	2.41	2.41	2.41	2.41	1.99	1.99
Window SHGC	-	Pre-1980	0.38	0.65	0.65	0.67	0.67	0.67	0.67	0.67	0.67	0.76	0.76	0.77	0.77	0.77	0.77
		Post-1980	0.51	0.51	0.51	0.55	0.55	0.55	0.59	0.59	0.59	0.60	0.60	0.62	0.62	0.62	0.62

Fig. 7 shows the energy performance of prototypical medium office building energy models. The CV(RMSE) between empirical baselines and modeled energy consumptions for these 30 models is 0.05, which meets the requirement of ASHRAE Guideline 14. Most modeled energy consumptions are very close to the empirical baselines. The relative difference between the modeled energy consumption and the empirical baseline is even lower than 1% for some climate zones, like 1A, 3A, 4C, 6B, and 8 in pre-1980 and 2A and 4A in post-1980.

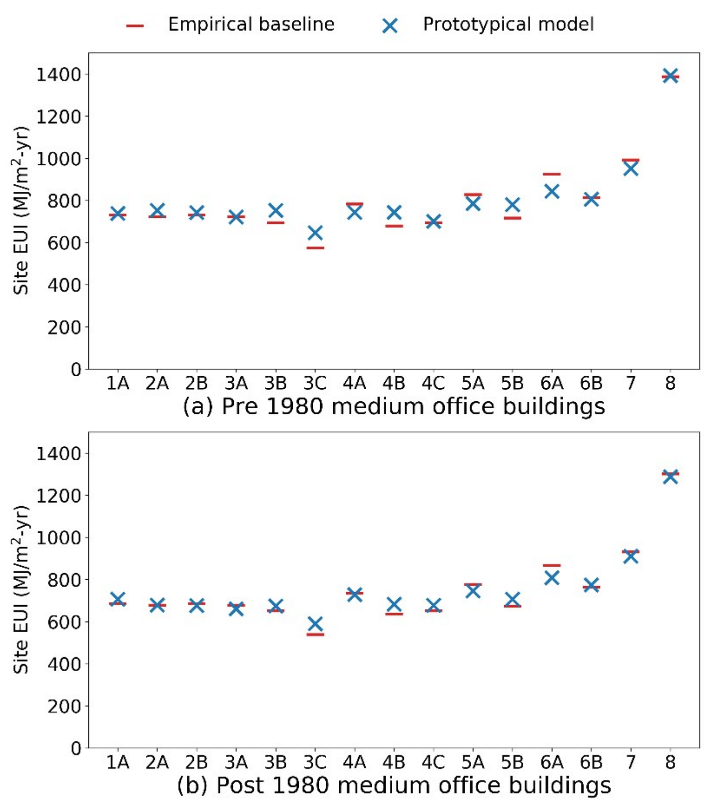

Fig. 7. Energy performance prototypical medium office building energy models based on the CBECS 2012.

3.2.2. Training samples

Using the prototypical building energy models as a starting point, the neutral model inputs were modified to be the same as the candidate buildings. The number of building samples in each climate zone and vintage depends on the number of neutral building characteristics (will be introduced in subsection 3.3.1). To get a large sample of modeled baselines, this study created and simulated 42,000 building samples. The distribution of modeled baselines of building samples is shown in Fig. 8 using violin plots. The upper line in the violin plot represents the upper quartile site EUI in that climate zone; the middle line in the violin plot represents the median site EUI in that climate zone; the lower line in the violin plot represents the lower quartile site EUI in that climate zone. The width of each curve corresponds with the approximate frequency of that site EUI. Buildings in climate 3C have the most concentrated distribution on site EUI while their site EUIs are most discrete in climate 8.

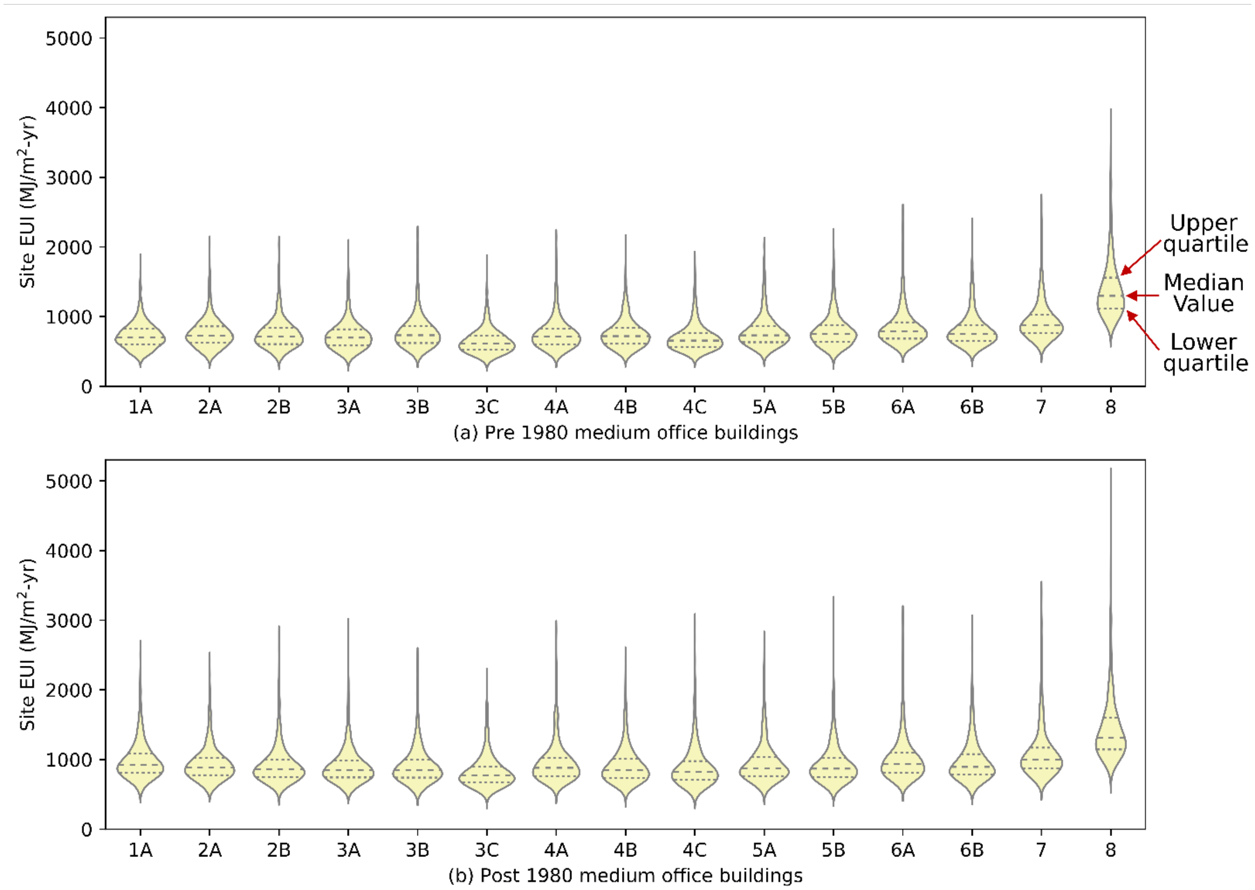

Fig. 8. Modeled baselines for medium office buildings in the U.S.

3.3. Key neutral building characteristics

Building characteristics that should be neutral in the comparison of energy performance for medium office buildings are proposed in subsection 3.3.1. Then, subsection 3.3.2 identifies key neutral building characteristics by conducting the correlation test between these proposed neutral building characteristics and site EUIs of building samples.

3.3.1. Neutral building characteristics

The climate zone and the year of building construction have already been considered as neutral building characteristics when developing prototypical building energy models in subsection 3.2.1. Furthermore, this research proposed 14 other neutral building characteristics related to building geometry and operation, as listed in Table 6. For the model inputs provided by the CBECS 2012 (total floor area, weekly operation hours, and window-to-wall ratio), the minimum and the maximum values were defined by excluding outliers (Galarnyk 2018). The ranges of the other 11 model inputs were defined by referring to the existing literature and engineering judgment.

Table 6. Building characteristics that are neutral in energy rating

Building Characteristics	**Unit**	**Default Value**	**Range**
Total floor area	m^2	3,437	[520, 15,130]
Aspect ratio	-	2.07	[1.45, 2.69]
Floor-to-floor height	m	4.53	[3.96, 5.69]
Window-to-wall ratio	-	0.18	[0, 0.65]
Building orientation	degree	0	[0, 360) *
People density	person/m^2	0.035	[0.023, 0.054]
Indoor heating setpoint temperature	^{o}C	20	[20, 22]
Indoor cooling setpoint temperature	^{o}C	25	[22, 25]
Service water usage	lpm	1.86	[1.33, 2.43]
Indoor heating design supply air humidity ratio	kg-water/kg-air	0.0080	[0.0056, 0.0104]
Indoor cooling design supply air humidity ratio	kg-water/kg-air	0.0085	[0.0060, 0.0111]
Weekly operation hours	hours/week	53.75	[25, 95]
Electric equipment power density	W/m^2	Pre-1980: 13.08 Post-1980: 8.19	Pre-1980: [5.40, 21.30] Post-1980: [3.37, 13.34]
Ventilation	m^3/s-person	0.0123	[0.0066, 0.0236]

* 0 degree means that the direction of the building is south (the direction illustrated in Fig.6); 90 degree means that the direction of the building is west; 180 degree means that the direction of the building is east; 270 degree means that the direction of the building is north.

3.3.2. Identification of key neutral building characteristics

The correlation coefficient (ρ) between the proposed neutral building characteristics and site EUIs of building samples are shown in Table 7. The positive value means that the building characteristic has a positive correlation to the energy consumption of the building, while the negative value means that the building characteristic has a negative correlation to the energy consumption of the building. Total floor area, window-to-wall ratio, and weekly operation hours are key neutral building characteristics for pre-1980 medium office buildings in all climate zones. Floor-to-floor height, people density, indoor cooling setpoint temperature, electric equipment power density, and ventilation are key neutral building characteristics for pre-1980 medium office buildings in several climate zones. Total floor area, floor-to-floor height, window-to-wall ratio, cooling setpoint temperature, weekly operation hours, and electric equipment power density are key neutral building characteristics for post-1980 medium office buildings in all climate zones. People density and ventilation are key neutral building characteristics for post-1980 medium office buildings in several climate zones.

Table 7. Key building characteristics that are neutral in energy rating for medium office buildings

Note: light red shading means key neutral building characteristics

Neutral building characteristics	Climate Zone														
	1A	2A	2B	3A	3B	3C	4A	4B	4C	5A	5B	6A	6B	7	8
Pre-1980 medium office buildings															
Total floor area	-0.50	-0.51	-0.57	-0.53	-0.56	-0.53	-0.63	-0.60	-0.63	-0.67	-0.63	-0.72	-0.69	-0.72	-0.75
Aspect ratio	0.04	0.00	-0.02	0.06	0.05	0.01	0.02	0.02	0.05	0.00	0.01	-0.01	0.04	0.09	0.03
Floor-to-floor height	0.07	0.11	0.15	0.09	0.14	0.04	0.10	0.11	0.16	0.16	0.17	0.15	0.16	0.19	0.20
Window-to-wall ratio	0.26	0.30	0.29	0.30	0.30	0.31	0.26	0.25	0.21	0.24	0.30	0.22	0.23	0.17	0.20
Building orientation	0.02	0.02	-0.01	-0.02	-0.03	0.01	0.01	0.01	0.01	0.01	-0.05	0.00	0.00	0.04	-0.01
People density	0.01	0.02	0.01	0.06	0.02	0.06	0.08	0.03	0.09	0.15	0.05	0.10	0.14	0.20	0.26
Indoor heating setpoint temperature	0.00	-0.04	0.00	-0.03	0.02	0.05	0.01	0.06	0.00	0.03	0.02	-0.01	0.05	0.02	-0.03
Indoor cooling setpoint temperature	-0.15	-0.13	-0.13	-0.15	-0.16	-0.15	-0.09	-0.07	-0.05	-0.08	-0.08	-0.05	-0.04	-0.07	0.03
Service water usage	0.03	0.01	0.03	-0.01	-0.01	0.02	-0.03	0.04	0.07	-0.04	0.06	0.05	-0.03	0.03	0.01
Indoor heating design supply air humidity ratio	0.01	-0.02	-0.02	0.03	0.01	-0.01	-0.04	-0.04	-0.02	0.00	0.01	-0.05	0.02	0.00	0.01
Indoor cooling design supply air humidity ratio	-0.01	-0.01	-0.02	-0.03	0.01	0.03	-0.02	0.02	-0.03	0.03	0.00	-0.02	0.04	-0.03	0.02
Weekly operation hours	0.64	0.59	0.58	0.60	0.56	0.59	0.55	0.57	0.49	0.50	0.55	0.45	0.47	0.39	0.28
Electric equipment power density	0.39	0.36	0.31	0.38	0.35	0.31	0.27	0.31	0.25	0.20	0.24	0.16	0.20	0.13	0.07
Ventilation	0.02	0.06	0.06	0.03	0.01	-0.01	0.14	0.05	0.10	0.21	0.13	0.19	0.15	0.28	0.28
Post-1980 medium office buildings															
Total floor area	-0.55	-0.58	-0.64	-0.58	-0.62	-0.60	-0.67	-0.66	-0.67	-0.67	-0.63	-0.69	-0.70	-0.74	-0.75
Aspect ratio	0.06	-0.02	0.07	0.01	0.02	0.03	0.04	-0.01	0.00	0.06	0.07	0.03	-0.01	0.01	0.06
Floor-to-floor height	0.13	0.15	0.13	0.12	0.14	0.11	0.14	0.16	0.18	0.12	0.12	0.19	0.12	0.20	0.17
Window-to-wall ratio	0.52	0.44	0.39	0.46	0.45	0.43	0.39	0.46	0.48	0.40	0.43	0.39	0.44	0.32	0.29
Building orientation	0.00	0.00	-0.01	0.02	0.06	0.00	-0.04	-0.03	-0.03	0.03	-0.01	0.01	-0.01	-0.02	0.02
People density	0.12	0.08	0.05	0.05	0.02	0.05	0.09	0.06	0.03	0.10	0.03	0.10	0.03	0.11	0.19
Indoor heating setpoint temperature	-0.02	0.04	0.02	0.00	0.07	-0.01	0.00	0.02	0.00	0.03	0.03	0.00	-0.01	0.02	0.00
Indoor cooling setpoint temperature	-0.28	-0.29	-0.23	-0.29	-0.25	-0.26	-0.31	-0.23	-0.28	-0.25	-0.24	-0.20	-0.18	-0.20	-0.12
Service water usage	-0.01	-0.04	-0.01	-0.06	-0.03	-0.04	0.00	0.01	0.01	-0.01	0.00	0.01	0.02	-0.02	0.00
Indoor heating design supply air humidity ratio	0.06	0.03	-0.02	0.02	0.01	-0.02	0.00	-0.03	0.00	-0.02	0.02	-0.01	0.02	-0.02	0.04
Indoor cooling design supply air humidity ratio	-0.02	-0.04	0.03	-0.05	0.02	-0.08	0.06	0.05	0.01	-0.01	0.03	0.04	-0.05	-0.03	0.02
Weekly operation hours	0.28	0.31	0.29	0.29	0.34	0.34	0.27	0.25	0.30	0.27	0.30	0.24	0.29	0.26	0.15
Electric equipment power density	0.32	0.34	0.31	0.31	0.32	0.34	0.26	0.30	0.21	0.22	0.26	0.21	0.22	0.15	0.17
Ventilation	0.13	0.05	0.04	0.08	0.03	-0.02	0.14	0.03	0.02	0.09	0.08	0.12	0.12	0.18	0.29

3.4. Individualized empirical baselines

Using the key neutral building characteristics in corresponding vintage and climate zone as input variables, this research developed mathematical models to get individualized baselines in a fast way. These 30 well-trained mathematical models were provided in GitHub (Lou 2022a).

To validate these mathematical models, we generated individualized empirical baselines for the rest of 20% of building samples. The MAPE between individualized empirical baselines and their modeled baselines for these 30 mathematical models are listed in Table 8, which are all lower than 5.5%. According to the criteria proposed by Setiawan et al. (Setiawan et al. 2021), if the MAPE value is lower than 10%, the prediction is accurate. Subsection 4.1 will introduce the application of these mathematical models for engineers to generate individualized empirical baselines in a fast way.

Table 8. Performance of mathematical models to generate individualized empirical baselines

Mathematical Model	**MAPE**		**Mathematical Model**	**MAPE**
Pre-1980, 1A	3.6%		Post-1980, 1A	4.7%
Pre-1980, 2A	3.5%		Post-1980, 2A	4.2%
Pre-1980, 2B	3.6%		Post-1980, 2B	4.5%
Pre-1980, 3A	3.8%		Post-1980, 3A	4.4%
Pre-1980, 3B	4.0%		Post-1980, 3B	4.2%
Pre-1980, 3C	3.7%		Post-1980, 3C	3.8%
Pre-1980, 4A	4.3%		Post-1980, 4A	4.3%
Pre-1980, 4B	4.3%		Post-1980, 4B	5.3%
Pre-1980, 4C	4.3%		Post-1980, 4C	4.8%
Pre-1980, 5A	3.8%		Post-1980, 5A	5.0%
Pre-1980, 5B	4.0%		Post-1980, 5B	5.1%
Pre-1980, 6A	4.2%		Post-1980, 6A	4.5%
Pre-1980, 6B	3.8%		Post-1980, 6B	4.7%
Pre-1980, 7	3.7%		Post-1980, 7	4.3%
Pre-1980, 8	4.0%		Post-1980, 8	4.6%

Fig. 9 shows the quantile-quantile plot between individualized empirical baselines and modeled baselines for all building samples. One red dot represents one building sample. If a red dot is close to the black line, it means that the individualized empirical baseline of this building is close to its modeled baseline. The relative errors between the individualized empirical baseline and the modeled baseline for all validated building samples are lower than 25%.

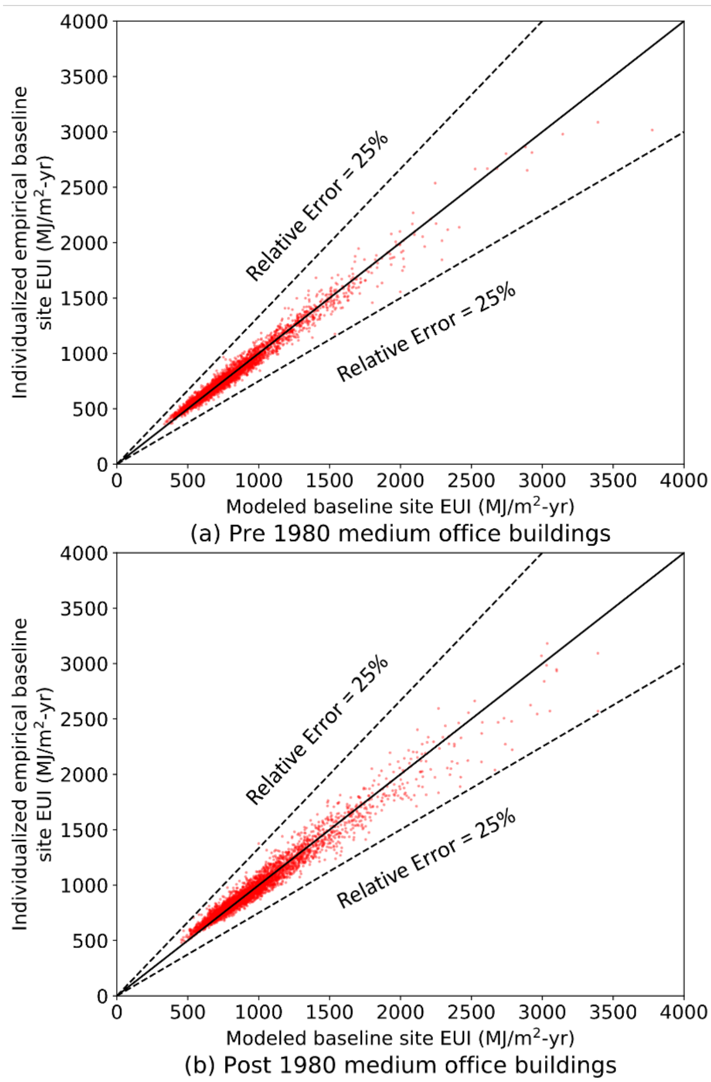

Fig. 9. Quantile-quantile plot between individualized empirical baselines and modeled baselines for the validation of all building samples

4. Discussion

4.1. Application of this research on medium office buildings in the U.S.

The mathematical models developed in this research can be directly used for a medium office building in the U.S. to obtain its individualized empirical baseline following three steps (Fig. 10). After downloading the mathematical models from the GitHub repository (Lou 2022a), the user should first check key neutral building characteristics in corresponding climate zones and vintages in the key_neutral_building_characteristics.csv file. Then, the user needs to add the value of key neutral building characteristics of the candidate building in the input_file.csv file. The final step is running the main.py. The individualized empirical baseline of the candidate building is generated in a few seconds and saved in the individualized_empirical_baseline.txt file.

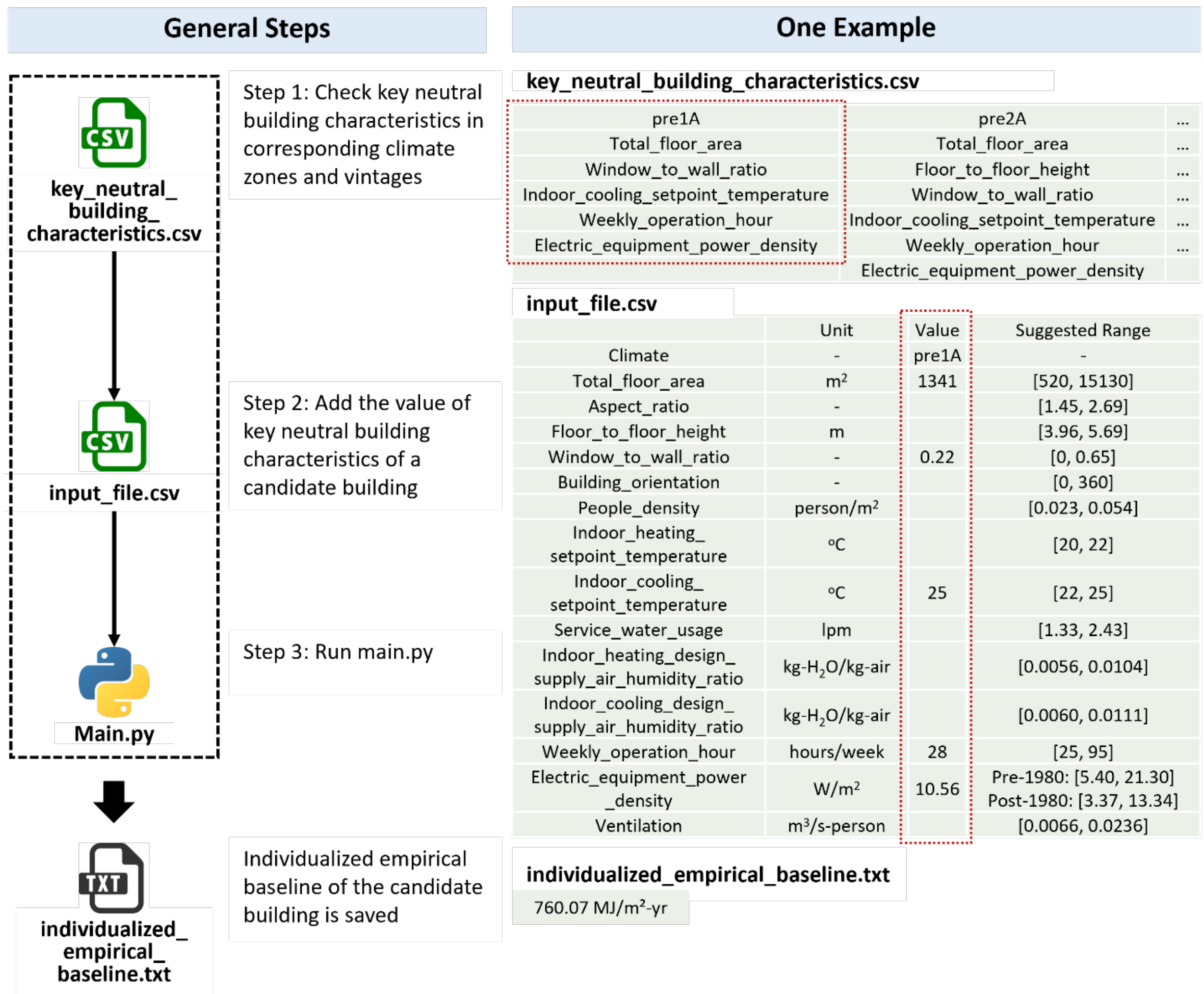

Fig. 10. General steps and one example to get individualized empirical baselines for medium office buildings

For example, an engineer may want to get the individualized empirical baseline for a medium office building in climate zone 1A built before 1980. As illustrated in key_neutral_building_characteristics.csv file, the key neutral building characteristics in pre-1980 climate 1A are total floor area, window-to-wall ratio, indoor cooling setpoint temperature, weekly operation hours, and electric equipment power density. Next, we need to put the value of these five building characteristics of the candidate medium office building in the input_file.csv file. Then, we need to run the main.py, and the individualized empirical baseline of this medium office building is saved in the individualized_empirical_baseline.txt file, which is 760.07 MJ/m^2-yr.

4.2. Application of this research on other types of commercial buildings in the U.S.

To get individualized empirical baselines for other commercial buildings in the U.S., mathematical models should be reconstructed by following the methodology illustrated Fig.1. Common empirical baselines and modeled baselines need to be recreated. Neutral building characteristics and their possible values may need to be reconsidered. For example, gas equipment use density should be considered as a neutral building characteristic for restaurant buildings.

5. Conclusion

This research developed a methodology to get individualized empirical baselines for existing buildings in a fast way. First, empirical baselines are created based on survey data. Then, to get training samples, building energy models for large-scale existing buildings are created and simulated. Finally, based on simulation results, mathematical models to get individualized empirical baselines in a fast way are created. This research used U.S. medium office buildings as an example to demonstrate the method. Empirical baselines for the U.S. medium office buildings in 30 conditions combining two vintages (constructed before 1980 and after 1980) and 15 climate zones were calculated based on the 2012 CBECS data and ASHRAE standard 100. Then, to get training samples, 42,000 building energy models based on the 2012 CBECS data were created and simulated. Finally, based on the simulation results, we developed 30 mathematical models to get individualized empirical baselines for existing buildings. Those mathematical models were accurate because the MAPEs between individualized empirical baselines and modeled baselines are all lower than 5.5%. An engineer can get the individualized empirical baseline for a medium office building in a few seconds by using the open-source tool we developed (Lou 2022a).

The contribution of this study mainly lies in the following two aspects. First, this research developed a methodology to obtain individualized empirical baselines in a fast way, which can be applied to all commercial buildings in the U.S. Second, this research developed prototypical building energy models for medium office buildings based on the latest available survey data. Those models can be used to study the energy and carbon emissions of existing medium office buildings in the U.S.

Acknowledgement

This paper is the outcome of the research project TRP-1771 sponsored by American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). This research was supported by the National Science Foundation under Awards No. CBET- 2217410.

Reference

ASHRAE. 2014. *ASHRAE Guideline 14–2014: Measurement of Energy, Demand and Water Savings*.

ASHRAE. 2015. *Standard 100-2015 -- Energy Efficiency in Existing Buildings*.

ASHRAE. 2017. Standard 100-2015 -- Energy efficiency in existing buildings.

ASHRAE. 2022. Building EQ.

CBECS. 2016. Commercial Buildings Energy Consumption Survey (CBECS).

Chen, J., X. Gao, Y. Hu, Z. Zeng, and Y. Liu. 2019. A meta-model-based optimization approach for fast and reliable calibration of building energy models. *Energy*, *188*:116046.

Chu, Z., J. Yu, and A. Hamdulla. 2021. Throughput prediction based on ExtraTree for stream processing tasks. *Computer Science and Information Systems*, *18*(1):1–22.

CU Boulder. 2022. RMACC Summit Supercomputer.

Deru, M., K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, and M. Rosenberg. 2011. US Department of Energy commercial reference building models of the national building stock.

ENERGY STAR. 2021. ENERGY STAR for Buildings and Plants.

Galarnyk, M. 2018. Understanding boxplots.

Geurts, P., D. Ernst, and L. Wehenkel. 2006. Extremely randomized trees. *Machine Learning*, *63*(1):3–42.

Griffith, B., N. Long, P. Torcellini, R. Judkoff, D. Crawley, and J. Ryan. 2008. *Methodology for modeling building energy performance across the commercial sector*. National Renewable Energy Lab.(NREL), Golden, CO (United States).

Huang, J., and E. Franconi. 1999. Commercial heating and cooling loads component analysis. *LBL-37208, Lawrence Berkeley National Laboratory, Berkeley, CA*.

International Code Council. 2022. 2021 International Energy Conservation Code (IECC).

John, V., Z. Liu, C. Guo, S. Mita, and K. Kidono. 2015. Real-time lane estimation using deep features and extra trees regression. In *Image and Video Technology* (pp. 721–733). Springer.

Kapoor, H. 2020. Random Forest vs Extra Trees.

Lim, H., and Z. J. Zhai. 2017. Comprehensive evaluation of the influence of meta-models on Bayesian calibration. *Energy and Buildings*, *155*:66–75.

Lim, H., and Z. J. Zhai. 2018. Influences of energy data on Bayesian calibration of building energy model. *Applied Energy*, *231*:686–698.

Lou, Y. 2022a. individualized_empirical_baseline.

Lou, Y. 2022b. prototypical-medium-office-building-energy-models-based-on-CBECS-2012.

McKay, M. D., R. J. Beckman, and W. J. Conover. 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *42*(1):55–61.

NREL. 2022a. Building component library.

NREL. 2022b. openstudio-standards.

Schober, P., C. Boer, and L. A. Schwarte. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5):1763–1768.

Scikit-Learn. 2022. Extra Trees Regressor.

Setiawan, I. N., R. Kurniawan, B. Yuniarto, R. E. Caraka, and B. Pardamean. 2021. Parameter optimization of support vector regression using Harris hawks optimization. *Procedia Computer Science*, *179*:17– 24.

Sharp, T. 1996. Energy benchmarking in commercial office buildings. In *Proceedings of the ACEEE* (Vol. 7996, pp. 321–329).

Wang, N, S. Goel, V. Srivastava, and A. Makhmalbaf. 2015. Building energy asset score program overview and technical protocol (version 1.2). *PNNL-22045. Richland, WA, USA: Pacific Northwest National Laboratory*.

Wang, Na, and W. J. Gorrissen. 2013. *Commercial building energy asset score system: program overview and technical protocol (Version 1.1)*. Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Wikipedia. 2022. Decision tree learning.

Winiarski, D. W., M. A. Halverson, and W. Jiang. 2007. *Analysis of building envelope construction in 2003 CBECS*. Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Winiarski, D. W., W. Jiang, and M. A. Halverson. 2006. *Review of pre-and post-1980 buildings in Cbecs-Hvac equipment*. Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Ye, Y., K. Hinkelman, Y. Lou, W. Zuo, G. Wang, and J. Zhang. 2021. Evaluating the energy impact potential of energy efficiency measures for retrofit applications: A case study with US medium office buildings. *Building Simulation*, 1–17.

Ye, Y., Y. Lou, M. Strong, S. Upadhyaya, W. Zuo, and G. Wang. 2020. Development of New Baseline Models for US Medium Office Buildings Based on Commercial Buildings Energy Consumption Survey Data. *Science and Technology for the Built Environment*, (just-accepted):1–19.

Yuan, Y., L. Wu, and X. Zhang. 2021. Gini-Impurity index analysis. *IEEE Transactions on Information Forensics and Security*, *16*:3154–3169.